Dokumentvorlage, Version vom 18.04.2013

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Selexipag (Uptravi®)

Actelion Pharmaceuticals Deutschland GmbH

Modul 4 A

Langzeitbehandlung der PAH der WHO-/NYHA-Klassen II und III, als Kombinationstherapie bei erwachsenen Patienten, deren Erkrankung mit einem ERA und/oder einem PDE-5-I unzureichend kontrolliert ist, oder als Monotherapie bei Patienten, die für diese Therapien nicht infrage kommen

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 13.06.2016

Inhaltsverzeichnis

Seite

Tabellenverzeichnis **4**
Abbildungsverzeichnis **11**
Abkürzungsverzeichnis **13**
4 Modul 4 – allgemeine Informationen **16**
4.1 Zusammenfassung der Inhalte von Modul 4 17
4.2 Methodik 29
4.2.1 Fragestellung 29
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung 40
4.2.3 Informationsbeschaffung 45
4.2.3.1 Studien des pharmazeutischen Unternehmers 45
4.2.3.2 Bibliografische Literaturrecherche 46
4.2.3.3 Suche in Studienregistern 47
4.2.3.4 Selektion relevanter Studien 49
4.2.4 Bewertung der Aussagekraft der Nachweise 50
4.2.5 Informationssynthese und -analyse 52
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien 52
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien 52
4.2.5.3 Meta-Analysen 70
4.2.5.4 Sensitivitätsanalysen 71
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren 73
4.2.5.6 Indirekte Vergleiche 77
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen 79
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel 79
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel 79
4.3.1.1.1 Studien des pharmazeutischen Unternehmers 79
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche 81
4.3.1.1.3 Studien aus der Suche in Studienregistern 83
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel 84
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 85
4.3.1.2.1 Studiendesign und Studienpopulationen 85
4.3.1.2.2 Verzerrungspotential auf Studienebene 101
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien 102
4.3.1.3.1 Ergebnisse patientenrelevanter Endpunkte – RCT 104
4.3.1.3.1.1 Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis 105
4.3.1.3.1.2 Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH 132

4.3.1.3.1.3 Zeit bis zum Tod (Studienschluss) ... 145
4.3.1.3.1.4 Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT) ... 157
4.3.1.3.1.5 Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss) ... 165
4.3.1.3.1.6 Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH ... 172
4.3.1.3.1.7 Symptome der PAH ... 179
4.3.1.3.1.8 Absolute Veränderung der 6-MWD ... 183
4.3.1.3.1.9 Ausbleibende Verschlechterung der WHO-/NYHA-Klasse ... 193
4.3.1.3.1.10 Absolute Veränderung des Borg-Dyspnoe-Index ... 203
4.3.1.3.1.11 Absolute Veränderung des CAMPHOR „Symptome“-Scores ... 210
4.3.1.3.1.12 Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores ... 219
4.3.1.3.1.13 Unerwünschte Ereignisse ... 228
4.3.1.3.1.14 Schwerwiegende unerwünschte Ereignisse ... 242
4.3.1.3.1.15 Therapieabbrüche wegen unerwünschter Ereignisse ... 251
4.3.1.3.1.16 Unerwünschte Ereignisse von besonderem Interesse ... 260
4.3.1.3.1.17 Prostacyclin-assoziierte unerwünschte Ereignisse ... 267
4.3.1.3.2 Subgruppenanalysen – RCT ... 275
4.3.1.3.2.1 Zeit bis zum ersten CEC-bestätigten Morbiditäts-/Mortalitätsereignis ... 278
4.3.1.3.2.2 Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH ... 279
4.3.1.3.2.3 Zeit bis zum Tod (Studienschluss) ... 280
4.3.1.3.2.4 Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT) ... 282
4.3.1.3.2.5 Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss) ... 283
4.3.1.3.2.6 Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH ... 284
4.3.1.3.2.7 Symptome der PAH ... 285
4.3.1.3.2.8 Absolute Veränderung der 6-MWD ... 286
4.3.1.3.2.9 Ausbleibende Verschlechterung der WHO-/NYHA-Klasse ... 287
4.3.1.3.2.10 Absolute Veränderung des Borg-Dyspnoe-Index ... 288
4.3.1.3.2.11 Absolute Veränderung des CAMPHOR "Symptome"-Scores ... 289
4.3.1.3.2.12 Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores ... 290
4.3.1.3.2.13 Unerwünschte Ereignisse ... 291
4.3.1.3.2.14 Schwerwiegende unerwünschte Ereignisse ... 293
4.3.1.3.2.15 Therapieabbrüche wegen unerwünschter Ereignisse ... 294
4.3.1.3.2.16 Unerwünschte Ereignisse von besonderem Interesse ... 296
4.3.1.3.2.17 Prostacyclin-assoziierte unerwünschte Ereignisse ... 300
4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien ... 301
4.3.2 Weitere Unterlagen ... 310
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien ... 310
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche ... 310

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 310
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 311
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT 311
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT 313
4.3.2.2 Nicht randomisierte vergleichende Studien 314
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 314
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien 314
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 315
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien 315
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien 316
4.3.2.3 Weitere Untersuchungen 317
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 317
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 317
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen 317
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen 317
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen 318
4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen 318
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens 320
4.4.1 Beurteilung der Aussagekraft der Nachweise 320
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß 323
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht 332
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte 333
4.5.1 Begründung für die Vorlage indirekter Vergleiche 333
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen 333
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen 333
4.5.4 Verwendung von Surrogatendpunkten 333
4.6 Liste der eingeschlossenen Studien 335
4.7 Referenzliste 335
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche 343
Anhang 4-B : Suchstrategien – Suche in Studienregistern 347
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche) 349
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern) 351
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT 358
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 378

Tabellenverzeichnis

Seite

Tabelle 4-1: Behandlungssituationen im Anwendungsgebiet von Selexipag 30
Tabelle 4-2: Einschlusskriterien für Studien zum Nachweis von Nutzen und Zusatznutzen .. 44
Tabelle 4-3: Ausschlusskriterien für Studien zum Nachweis von Nutzen und Zusatznutzen . 45
Tabelle 4-4: Patientenrelevante Endpunkte in GRIPHON 56
Tabelle 4-5: 6-MWD - Zeitpunkt der Messung und Ersetzungsstrategien in verschiedenen PAH-Studien 64
Tabelle 4-6: Funktionsklassen der pulmonalen Hypertonie nach WHO/NYHA 65
Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 80
Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel 81
Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel 83
Tabelle 4-10: Studienpool – RCT mit dem zu bewertenden Arzneimittel 84
Tabelle 4-11: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 86
Tabelle 4-12: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel 88
Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel 89
Tabelle 4-14: Anteil individueller Erhaltungsdosen in GRIPHON 95
Tabelle 4-15: Verzerrungspotential auf Studienebene – RCT mit dem zu bewertenden Arzneimittel 101
Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel 103
Tabelle 4-17: Operationalisierung von „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ 105
Tabelle 4-18: Bewertung des Verzerrungspotentials für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ in RCT mit dem zu bewertenden Arzneimittel 107
Tabelle 4-19: Ergebnisse für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis-Ereignis“ aus RCT mit dem zu bewertenden Arzneimittel 108
Tabelle 4-20: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ 114
Tabelle 4-21: Ergebnisse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ nach IMD-Kategorien 117

Tabelle 4-22: Gründe für die Zensierung in der Hauptanalyse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ (unter Ausschluss der Ereignisse, die vor dem 16. August 2011 begannen) 119

Tabelle 4-23: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ 121

Tabelle 4-24: Competing-Risk-Analyse für den Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ (Methode nach Gray) 122

Tabelle 4-25: Analyse der Zeit bis zu den Einzelereignissen des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ 124

Tabelle 4-26: Analyse der naiven Raten für den Endpunkt „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ (einschließlich Einzelkomponenten) 126

Tabelle 4-27: Analyse der naiven Raten der Einzelbestandteile der Komponente „Fortschreiten der Erkrankung“ des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ 128

Tabelle 4-28: Analyse des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten 130

Tabelle 4-29: Operationalisierung von „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ 133

Tabelle 4-30: Bewertung des Verzerrungspotentials für „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ in RCT mit dem zu bewertenden Arzneimittel 134

Tabelle 4-31: Ergebnisse für „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ aus RCT mit dem zu bewertenden Arzneimittel 135

Tabelle 4-32: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ 138

Tabelle 4-33: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ 140

Tabelle 4-34: Analyse der naiven Raten für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ (einschließlich Einzelkomponenten) 142

Tabelle 4-35: Analyse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten 143

Tabelle 4-36: Operationalisierung von „Zeit bis zum Tod (Studienschluss)“ 145

Tabelle 4-37: Bewertung des Verzerrungspotentials für „Zeit bis zum Tod (Studienschluss)“ in RCT mit dem zu bewertenden Arzneimittel 146

Tabelle 4-38: Ergebnisse für „Zeit bis zum Tod (Studienschluss)“ aus RCT mit dem zu bewertenden Arzneimittel 147
Tabelle 4-39: RPSFT-Modell: Schätzer für Ψ und für das relative Überleben..... 150
Tabelle 4-40: MSM – Schätzer des Behandlungseffekts (HR) mit 95 % KI, robustem Standardfehler und p-Wert 152
Tabelle 4-41: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum Tod (Studienschluss)“ 152
Tabelle 4-42: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Zeit bis zum Tod (Studienschluss)“ 153
Tabelle 4-43: Analyse der naiven Raten für den Endpunkt „Zeit bis zum Tod (Studienschluss)“..... 155
Tabelle 4-44: Operationalisierung von „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“..... 157
Tabelle 4-45: Bewertung des Verzerrungspotentials für „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ in RCT mit dem zu bewertenden Arzneimittel 158
Tabelle 4-46: Ergebnisse für „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ aus RCT mit dem zu bewertenden Arzneimittel 159
Tabelle 4-47: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ 162
Tabelle 4-48: Analyse der naiven Raten für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“..... 163
Tabelle 4-49: Operationalisierung von „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“..... 165
Tabelle 4-50: Bewertung des Verzerrungspotentials für „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“ in RCT mit dem zu bewertenden Arzneimittel 166
Tabelle 4-51: Ergebnisse für „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“ aus RCT mit dem zu bewertenden Arzneimittel..... 167
Tabelle 4-52: Analyse der naiven Raten für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“..... 170
Tabelle 4-53: Operationalisierung von „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ 172
Tabelle 4-54: Bewertung des Verzerrungspotentials für „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ in RCT mit dem zu bewertenden Arzneimittel .. 173
Tabelle 4-55: Ergebnisse für „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ aus RCT mit dem zu bewertenden Arzneimittel 174
Tabelle 4-56: Analyse der naiven Raten für den Endpunkt „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ 177
Tabelle 4-57: Operationalisierung von „Symptome der PAH“..... 179
Tabelle 4-58: Bewertung des Verzerrungspotentials für „Symptome der PAH“ in RCT mit dem zu bewertenden Arzneimittel..... 180
Tabelle 4-59: Ergebnisse für „Symptome der PAH“ aus RCT mit dem zu bewertenden Arzneimittel..... 181

Tabelle 4-60: Operationalisierung von „Absolute Veränderung der 6-MWD“ 183
Tabelle 4-61: Bewertung des Verzerrungspotentials für „Absolute Veränderung der 6-MWD“ in RCT mit dem zu bewertenden Arzneimittel 184
Tabelle 4-62: Ergebnisse für „Absolute Veränderung der 6-MWD“ aus RCT mit dem zu bewertenden Arzneimittel 185
Tabelle 4-63: Sensitivitätsanalysen für den Endpunkt „Absolute Veränderung der 6-MWD“..... 188
Tabelle 4-64: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Absolute Veränderung der 6-MWD“..... 189
Tabelle 4-65: Analyse des Endpunkts „Absolute Veränderung der 6-MWD“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten 190
Tabelle 4-66: Operationalisierung von „Ausbleibende Verschlechterung der WHO-/ NYHA-Klasse“..... 193
Tabelle 4-67: Bewertung des Verzerrungspotentials für „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ in RCT mit dem zu bewertenden Arzneimittel 193
Tabelle 4-68: Ergebnisse für „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ aus RCT mit dem zu bewertenden Arzneimittel 194
Tabelle 4-69: Sensitivitätsanalysen für den Endpunkt „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ 196
Tabelle 4-70: Stratifizierte Analysen des Endpunkts „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“..... 197
Tabelle 4-71: Analyse des Endpunkts „Ausbleibende Verschlechterung der WHO-/ NYHA-Klasse“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten..... 199
Tabelle 4-72: Operationalisierung von „Absolute Veränderung des Borg-Dyspnoe-Index“ 203
Tabelle 4-73: Bewertung des Verzerrungspotentials für „Absolute Veränderung des Borg-Dyspnoe-Index“ in RCT mit dem zu bewertenden Arzneimittel 204
Tabelle 4-74: Ergebnisse für „Absolute Veränderung des Borg-Dyspnoe-Index“ aus RCT mit dem zu bewertenden Arzneimittel 205
Tabelle 4-75: Analyse des Borg-Dyspnoe-Index zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten..... 207
Tabelle 4-76: Operationalisierung von „Absolute Veränderung des CAMPHOR „Symptome“-Scores“ 210
Tabelle 4-77: Bewertung des Verzerrungspotentials für „Absolute Veränderung des CAMPHOR „Symptome“-Scores“ in RCT mit dem zu bewertenden Arzneimittel 211
Tabelle 4-78: Ergebnisse für „Absolute Veränderung des CAMPHOR ‚Symptome‘-Scores“ aus RCT mit dem zu bewertenden Arzneimittel..... 212
Tabelle 4-79: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Absolute Veränderung des CAMPHOR ‚Symptome‘-Scores“ 215
Tabelle 4-80: Analyse des Endpunkts „Absolute Veränderung des CAMPHOR ‚Symptome‘-Scores“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten .. 216

Tabelle 4-81: Operationalisierung von „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" ... 219

Tabelle 4-82: Bewertung des Verzerrungspotentials für „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" in RCT mit dem zu bewertenden Arzneimittel ... 220

Tabelle 4-83: Ergebnisse für „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" aus RCT mit dem zu bewertenden Arzneimittel ... 221

Tabelle 4-84: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" ... 224

Tabelle 4-85: Analyse des Endpunkts „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten ... 225

Tabelle 4-86: Operationalisierung von „Unerwünschte Ereignisse" ... 228

Tabelle 4-87: Bewertung des Verzerrungspotentials für „Unerwünschte Ereignisse" in RCT mit dem zu bewertenden Arzneimittel ... 230

Tabelle 4-88: Ergebnisse für „Unerwünschte Ereignisse" aus RCT mit dem zu bewertenden Arzneimittel ... 231

Tabelle 4-89: Aufteilung der unerwünschten Ereignisse nach Titrations- und Erhaltungsphase ... 233

Tabelle 4-90: Sensitivitätsanalyse für den Endpunkt „Unerwünschte Ereignisse" unter Ausschluss der Ereignisse des primären Endpunkts in GRIPHON ... 235

Tabelle 4-91: Auflistung der unerwünschten Ereignisse nach bevorzugter Bezeichnung (MedDRA Version 16), die bei mindestens 5 % der Patienten in mindestens einem Studienarm aufgetreten waren (gesamte Behandlungsphase) ... 237

Tabelle 4-92: Auflistung der unerwünschten Ereignisse nach System-Organ-Klasse (MedDRA Version 16) ... 239

Tabelle 4-93: Operationalisierung von „Schwerwiegende unerwünschte Ereignisse" ... 242

Tabelle 4-94: Bewertung des Verzerrungspotentials für „Schwerwiegende unerwünschte Ereignisse" in RCT mit dem zu bewertenden Arzneimittel ... 243

Tabelle 4-95: Ergebnisse für „Schwerwiegende unerwünschte Ereignisse" aus RCT mit dem zu bewertenden Arzneimittel ... 244

Tabelle 4-96: Aufteilung der schwerwiegenden unerwünschten Ereignisse nach Titrations- und Erhaltungsphase ... 245

Tabelle 4-97: Sensitivitätsanalyse für den Endpunkt „Schwerwiegende unerwünschte Ereignisse" unter Ausschluss der Ereignisse des primären Endpunkts in GRIPHON ... 247

Tabelle 4-98: Auflistung der schwerwiegenden unerwünschten Ereignisse nach bevorzugter Bezeichnung (MedDRA Version 16), die bei mindestens 5 Patienten in mindestens einem Studienarm aufgetreten waren (gesamte Behandlungsphase) ... 248

Tabelle 4-99: Auflistung der schwerwiegenden unerwünschten Ereignisse nach System-Organ-Klasse (MedDRA Version 16) ... 249

Tabelle 4-100: Operationalisierung von „Therapieabbrüche wegen unerwünschter Ereignisse“ ... 251
Tabelle 4-101: Bewertung des Verzerrungspotentials für „Therapieabbrüche wegen unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel ... 251
Tabelle 4-102: Ergebnisse für „Therapieabbrüche wegen unerwünschter Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel ... 252
Tabelle 4-103: Aufteilung der Therapieabbrüche wegen eines unerwünschten Ereignisses nach Titrations- und Erhaltungsphase ... 254
Tabelle 4-104: Sensitivitätsanalyse für den Endpunkt „Therapieabbrüche wegen unerwünschter Ereignisse“ unter Ausschluss der Ereignisse des primären Endpunkts in GRIPHON ... 256
Tabelle 4-105: Auflistung der Therapieabbrüche wegen unerwünschter Ereignisse nach bevorzugter Bezeichnung des zugrunde liegenden unerwünschten Ereignisses (MedDRA Version 16), die bei mindestens 5 Patienten in mindestens einem Studienarm aufgetreten waren (gesamte Behandlungsphase) ... 257
Tabelle 4-106: Auflistung der Therapieabbrüche wegen unerwünschter Ereignisse nach System-Organ-Klasse (MedDRA Version 16) ... 258
Tabelle 4-107: Operationalisierung von „Unerwünschte Ereignisse von besonderem Interesse“ ... 260
Tabelle 4-108: Bewertung des Verzerrungspotentials für „Unerwünschte Ereignisse von besonderem Interesse“ in RCT mit dem zu bewertenden Arzneimittel ... 261
Tabelle 4-109: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse“ aus RCT mit dem zu bewertenden Arzneimittel ... 262
Tabelle 4-110: Hypotonie-assoziierte unerwünschte Ereignisse in GRIPHON, aufgeteilt nach bevorzugter Bezeichnung ... 264
Tabelle 4-111: Aufteilung der unerwünschten Ereignisse von besonderem Interesse nach Titrations- und Erhaltungsphase ... 265
Tabelle 4-112: Operationalisierung von „Prostacyclin-assoziierte unerwünschte Ereignisse“ ... 267
Tabelle 4-113: Bewertung des Verzerrungspotentials für „Prostacyclin-assoziierte unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel ... 267
Tabelle 4-114: Ergebnisse für „Prostacyclin-assoziierte unerwünschte Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel ... 269
Tabelle 4-115: Aufteilung der Prostacyclin-assoziierten unerwünschten Ereignisse nach Titrations- und Erhaltungsphase ... 271
Tabelle 4-116: Prostacyclin-assoziierte unerwünschte Ereignisse in GRIPHON, aufgeteilt nach bevorzugter Bezeichnung ... 272
Tabelle 4-117: Prostacyclin-assoziierte schwerwiegende unerwünschte Ereignisse in GRIPHON, aufgeteilt nach bevorzugter Bezeichnung ... 273
Tabelle 4-118: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche ... 311

Tabelle 4-119: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 311
Tabelle 4-120: Operationalisierung von <Endpunkt xxx> 312
Tabelle 4-121: Bewertung des Verzerrungspotentials für <Endpunkt xxx> in RCT für indirekte Vergleiche 312
Tabelle 4-122: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche 313
Tabelle 4-123: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien 315
Tabelle 4-124: Operationalisierung von <Endpunkt xxx> 315
Tabelle 4-125: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien 316
Tabelle 4-126: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen 318
Tabelle 4-127: Zusammenfassung der Ergebnisse zur Bestimmung des Zusatznutzens von Macitentan 325
Tabelle 4-128: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 332
Tabelle 4-129: Liste der eingeschlossenen Studien 335
Tabelle 4-130 (Anhang): Direkter Vergleich - Dokumentation der Recherche in MEDLINE 344
Tabelle 4-131 (Anhang): Direkter Vergleich - Dokumentation der Recherche in EMBASE 345
Tabelle 4-132 (Anhang): Direkter Vergleich - Dokumentation der Recherche bei „The Cochrane Library“ 346
Tabelle 4-133 (Anhang): Dokumentation der Recherche in clinicaltrials.gov 347
Tabelle 4-134 (Anhang): Dokumentation der Recherche in EU-CTR 347
Tabelle 4-135 (Anhang): Dokumentation der Recherche in WHO ICTRP 348
Tabelle 4-136 (Anhang): Dokumentation der Recherche in PharmNet.Bund 348
Tabelle 4-137 (Anhang): Liste der im Volltext gesichteten und ausgeschlossenen Dokumente (bibliografische Literaturrecherche) 349
Tabelle 4-138 (Anhang): Suche nach RCT - Liste der ausgeschlossenen Studien (Suche in Studienregistern) 352
Tabelle 4-139 (Anhang): Studiendesign und -methodik für Studie AC-065A302 – GRIPHON 359
Tabelle 4-140 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie AC-065A302 – GRIPHON 379

Abbildungsverzeichnis

Seite

Abbildung 4-1: Vergleich der Studiendauer von GRIPHON mit Studien zur Untersuchung der Wirksamkeit der Prostanoide ... 39

Abbildung 4-2: Aufbau des kombinierten Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ in GRIPHON ... 59

Abbildung 4-3: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel ... 82

Abbildung 4-4: Studiendesign GRIPHON ... 94

Abbildung 4-5: Kaplan-Meier-Kurve für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ ... 111

Abbildung 4-6: Kaplan-Meier-Kurven für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“, aufgeteilt nach IMD-Kategorien ... 118

Abbildung 4-7: Kaplan-Meier-Kurve der inversen Zensierung für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ (d. h. Zeit bis zum Studienabbruch ohne M/M-Ereignis) ... 120

Abbildung 4-8: Kumulative Inzidenzfunktion für das erste Morbiditäts- oder Mortalitätsereignis bis zu EOT + 7 Tage (Competing-Risk-Analyse nach Gray) ... 122

Abbildung 4-9: Kaplan-Meier-Kurve für „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ ... 137

Abbildung 4-10: Kaplan-Meier-Kurve für „Zeit bis zum Tod (Studienschluss)“ ... 149

Abbildung 4-11: Kaplan-Meier-Kurve für „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ ... 161

Abbildung 4-12: Kaplan-Meier-Kurve für „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“ ... 169

Abbildung 4-13: Kaplan-Meier-Kurve für „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ ... 176

Abbildung 4-14: Absolute Veränderung der 6-MWD (reguläre Untersuchungstermine im Vergleich zum Basiswert) ... 191

Abbildung 4-15: Ausbleibende Verschlechterung der WHO-/NYHA-Klasse (reguläre Untersuchungstermine im Vergleich zum Basiswert) ... 201

Abbildung 4-16: Patienten mit Verbesserung bzw. Verschlechterung der WHO-/NYHA-Klasse zu regulären Untersuchungsterminen ... 201

Abbildung 4-17: Borg-Dyspnoe-Index an regulären Untersuchungsterminen ... 208

Abbildung 4-18: Absolute Veränderung des CAMPHOR „Symptome“-Scores (reguläre Untersuchungstermine im Vergleich zum Basiswert) ... 217

Abbildung 4-19: Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores (reguläre Untersuchungstermine im Vergleich zum Basiswert) ... 226

Abbildung 4-20: Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis – Forest-Plot der Subgruppenanalysen 278
Abbildung 4-21: Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH – Forest-Plot der Subgruppenanalysen 279
Abbildung 4-22: Zeit bis zum Tod (Studienschluss) – Forest-Plot der Subgruppenanalysen 280
Abbildung 4-23: Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT) – Forest-Plot der Subgruppenanalysen 282
Abbildung 4-24: Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss) – Forest-Plot der Subgruppenanalysen 283
Abbildung 4-25: Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH – Forest-Plot der Subgruppenanalysen 284
Abbildung 4-26: Symptome der PAH – Forest-Plot der Subgruppenanalysen 285
Abbildung 4-27: Absolute Veränderung der 6-MWD – Forest-Plot der Subgruppenanalysen 286
Abbildung 4-28: Ausbleibende Verschlechterung der WHO-/NYHA-Klasse – Forest-Plot der Subgruppenanalysen 287
Abbildung 4-29: Absolute Veränderung des Borg-Dyspnoe-Index – Forest-Plot der Subgruppenanalysen 288
Abbildung 4-30: Absolute Veränderung des CAMPHOR „Symptome“-Scores – Forest-Plot der Subgruppenanalysen 289
Abbildung 4-31: Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores – Forest-Plot der Subgruppenanalysen 290
Abbildung 4-32: Unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen 291
Abbildung 4-33: Unerwünschte Ereignisse schwerer Intensität – Forest-Plot der Subgruppenanalysen 292
Abbildung 4-34: Schwerwiegende unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen 293
Abbildung 4-35: Therapieabbrüche wegen unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen 294
Abbildung 4-36: Unerwünschte Ereignisse von besonderem Interesse: Hämorrhagie – Forest-Plot der Subgruppenanalysen 296
Abbildung 4-37: Unerwünschte Ereignisse von besonderem Interesse: Hypotonie – Forest-Plot der Subgruppenanalysen 298
Abbildung 4-38: Unerwünschte Ereignisse von besonderem Interesse: Störungen des Auges – Forest-Plot der Subgruppenanalysen 299
Abbildung 4-39: Prostacyclin-assoziierte unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen 300
Abbildung 4-40: Patientenfluss in Studie AC-065A302 – GRIPHON 376

Abkürzungsverzeichnis

Abkürzung	Bedeutung
%-P.	Prozentpunkte
(S)UE	(schwerwiegendes) unerwünschtes Ereignis
6-MWD	6-Minuten-Gehstrecke
6-MWT	6-Minuten-Gehtest
AIR	Aerosolized Randomized Iloprost study
ANCOVA	Kovarianzanalyse (analysis of covariance)
ATS	American Thoracic Society
BMI	body mass index
CAMPHOR	Cambridge Pulmonary Hypertension Outcome Review
CEC	Critical Event Committee
COMBI	Combination Therapy of Bosentan and aerosolised Iloprost in Idiopathic Pulmonary Arterial Hypertension study
CONSORT	Consolidated Standards of Reporting Trials
CRF	Prüfbogen (case report form)
DIMDI	Deutsches Institut für Medizinische Dokumentation
EG	Europäische Gemeinschaft
EMA	European Medicines Agency
EOS	Studienende (end-of-study)
EOT	Behandlungsende (end-of-treatment)
EOTE	Behandlungsextensionsphase (end-of-treatment-extension)
ERA	Endothelin-Rezeptor-Antagonist
EU-CTR	EU Clinical Trials Register
FAS	full analysis set
FDA	US Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
GRIPHON	Prostacyclin (PGI_2) receptor agonist in pulmonary arterial hypertension
HR	Hazard Ratio
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IMD	individuelle Erhaltungsdosis (individual maintenance dose)

IPCW	Inverse Probability of Censoring Weighting
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
KI	Konfidenzintervall
LOCF	last observation carried forward
LVEDP	linksventrikulärer enddiastolischer Druck (left ventricular end diastolic pressure)
M/M-Ereignis	Ereignis des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“
MedDRA	Medical Dictionary for Regulatory Activities
mPAP	mittlerer pulmonal arterieller Druck (mean pulmonary arterial pressure)
MSM	„Marginal Structural Cox Proportional Hazards“-Modell
MTC	Mixed Treatment Comparison
MTD	patientenindividuell höchste tolerierte Dosis (maximum tolerated dose)
MW	Mittelwert
NT-pro-BNP	N-terminal pro-B type natriuretic peptide
OR	Odds Ratio
PAH	pulmonal arterielle Hypertonie
PCWP	pulmonaler Verschlussdruck (pulmonary capillary wedge pressure)
PDE-5(-I)	Phosphodiesterase-5(-Inhibitor)
PH	pulmonale Hypertonie
PIT	patientenindividuelle Therapie
PPS	per protocol set
PT	Bevorzugte Bezeichnung (preferred term)
PTOCV	post-treatment observation closure visit
PTOP	Beobachtungsphase nach der Behandlungsphase (post-treatment observation period)
PVR	pulmonaler Gefäßwiderstand (pulmonary vascular resistance)
QAS	quality of life analysis set
RCT	Randomized Controlled Trial

RD	(absolute) Risikodifferenz
REVEAL	Registry to Evaluate Early And Long-term PAH Disease Management
RPSFT	Rank Preserving Structural Accelerated Failure Time
RR	Risk Ratio
SAF	Safety (population)
SD	Standardabweichung (standard deviation)
SE	Standardfehler (standard error)
SERAPHIN	Study with endothelin receptor antagonist in pulmonary arterial hypertension to improve clinical outcome
SF-36	Short form (36) Gesundheitsfragebogen
SFU	Sicherheitsnachbeobachtung (safety follow-up)
SGB	Sozialgesetzbuch
SMQ	Standardized MedDRA Query
SOC	System-Organ-Klasse (system organ class)
STE	Surrogate Threshold Effects
STEP	Safety and pilot efficacy Trial in combination with bosentan for Evaluation in Pulmonary arterial hypertension study
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
TTCW	Zeit bis zur klinischen Verschlechterung (time to clinical worsening)
WHO	World Health Organization
WHO ICTRP	WHO International Clinical Trials Registry Platform
zVT	zweckmäßige Vergleichstherapie

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

– Zusammenfassung (Abschnitt 4.1)

– Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)

– Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)

– eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)

– ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

Ziel des vorliegenden Dossiers ist die Darstellung des Ausmaßes des Zusatznutzens für Selexipag im zugelassenen Anwendungsgebiet. Selexipag ist zugelassen zur Langzeitbehandlung der PAH der WHO-/NYHA-Klassen II und III, als Kombinationstherapie bei erwachsenen Patienten, deren Erkrankung mit einem ERA und/oder einem PDE-5-I unzureichend kontrolliert ist, oder als Monotherapie bei Patienten, die für diese Therapien nicht infrage kommen.

Durch die Definition des Anwendungsgebietes ist vorgegeben, dass die Zielpopulation von Selexipag ausschließlich solche Patienten umfasst, die entweder a) bereits mit einer Kombination aus ERA und PDE-5-I behandelt werden oder b) für eine (zusätzliche) Therapie mit den beiden Wirkstoffen nicht infrage kommen. Die Entscheidung, ab wann die Erkrankung eines Patienten durch die Therapie mit einem ERA und/oder einem PDE-5-I nicht mehr ausreichend kontrolliert ist bzw. ab wann er für eine (zusätzliche) Therapie mit einem oder beiden dieser Wirkstoffe nicht infrage kommt, kann hierbei für jeden Einzelfall nur patientenindividuell durch den behandelnden Arzt beantwortet werden. Festgelegte, quantifizierbare Kriterien, an denen der Arzt sich bei dieser Entscheidung orientieren kann, bestehen nicht.

Unabhängig von den Überlegungen zum Anwendungsgebiet geht Actelion aber davon aus, dass fast alle Patienten im Anwendungsgebiet von Selexipag bereits mit einer Kombinationstherapie aus einem ERA <u>und</u> einem PDE-5-I behandelt werden. Die Annahme hierbei ist, dass nur sehr wenige der mit einem ERA <u>oder</u> einem PDE-5-I vorbehandelten bzw. therapienaiven Patienten nicht für eine (zusätzliche) Therapie mit ERA und/oder PDE-5-I infrage kommen.

Von den im Krankheitsbild der PAH zugelassenen Wirkstoffen kann ausschließlich Iloprost zur Inhalation als im Anwendungsgebiet von Selexipag zugelassene und zweckmäßige Therapie angesehen werden. Iloprost ist damit der einzige Wirkstoff, der als mögliche Komponente der innerhalb der vom G-BA festgelegten zVT „patientenindividuelle medikamentöse Therapie nach Maßgabe des Arztes und unter Berücksichtigung des Zulassungsstatus“ in Betracht gezogen werden kann. Zu beachten ist hierbei aber, dass auch die Behandlung mit Iloprost nicht für alle Patienten im Anwendungsgebiet von Selexipag in Betracht kommt, da nach Berücksichtigung des individuellen Erkrankungszustandes bei manchen Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegen kann. Zudem ist Iloprost ausschließlich in WHO-/NYHA-Klasse III zugelassen.

Für diese Patientengruppe steht mit Ausnahme der Behandlung mit Selexipag keine zugelassene medikamentöse Therapie zur Verfügung, sodass das abwartende Vorgehen bis zur weiteren Verschlechterung der Erkrankung die einzige Therapieoption des behandelnden Arztes für diese Teilpopulation des Anwendungsgebietes von Selexipag ist.

Aufgrund der zuvor genannten Gründe und in Abhängigkeit von den zur Verfügung stehenden Therapieoptionen ergeben sich daher für die Nutzenbewertung von Selexipag die folgenden Teilpopulationen:

- Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt **(Teilpopulation a)**.

  Mögliche Therapieoptionen innerhalb der zVT PIT:

  abwartendes Vorgehen bis zur weiteren Verschlechterung der PAH

- Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt **(Teilpopulation b)**.
  - Mögliche Therapieoptionen innerhalb der zVT PIT:

    patientenindividuelle Therapie mit Iloprost (zur Inhalation) nach Maßgabe des Arztes

Es werden ausschließlich patientenrelevante Endpunkte zur Beurteilung herangezogen. Diese sind in Abschnitt 4.2.5.2 detailliert dargestellt.

Ein-/Ausschlusskriterien für Studien

Eingeschlossen wurden Studien im RCT-Design, sofern sie ganz oder teilweise im oben aufgeführten Anwendungsgebiet durchgeführt wurden und Selexipag in der zugelassenen Dosierung angewendet wurde. Weiterhin musste in der Studie ein Vergleich mit der in Abschnitt 4.2.1 konkretisierten zVT durchgeführt worden sein. Aufgrund des Anwendungsgebietes, das sich auf die Langzeitbehandlung der PAH erstreckt, musste die Studiendauer nutzenbewertungsrelevanter Studien in Anlehnung an die EMA-Leitlinie zur klinischen Erprobung medizinischer Produkte in der PAH mindestens 6 Monate betragen. Schließlich mussten relevante Studien mindestens einen patientenrelevanten Endpunkt berichten.

Ausgeschlossen wurden Studien, sofern sie die genannten Einschlusskriterien nicht erfüllten, keine Volltextpublikation[1] zu der Studie vorlag oder die Publikation in einer anderen Sprache als Deutsch oder Englisch verfasst war.

Um für die Nutzenbewertung von Selexipag in Teilpopulation a herangezogen werden zu können, musste die Patientenpopulation einer relevanten Studie zudem aus Patienten bestehen, bei denen das *Risiko* einer Iloprost-Behandlung deren *Nutzen* noch überwog, und es musste in der Studie ein Vergleich von Selexipag mit abwartendem Vorgehen bis zur Verschlechterung der Erkrankung durchgeführt worden sein (dieser konnte in der Studie auch

[1] Studienbericht, Volltextpublikation oder ausführliche Ergebnisdarstellung in Studienregister.

als Placebo-Behandlung operationalisiert sein). Für die Bewertung in Teilpopulation b wiederum wurden nur Studien mit Patienten herangezogen, bei denen der Nutzen der Iloprost-Behandlung deren Risiko überwog und in denen Selexipag mit einer patientenindividuellen Iloprost-Behandlung nach Maßgabe des Arztes verglichen wurde.

Datenquellen

Im vorliegenden Nutzendossier werden die Ergebnisse zum Ausmaß des Zusatznutzens von Selexipag (Uptravi®) in Teilpopulation a anhand der pivotalen Zulassungsstudie GRIPHON und im Vergleich zu Placebo dargestellt. Bei GRIPHON handelt es sich um eine multizentrische, randomisierte, doppelblinde, placebokontrollierte und ereignisgetriebene[2] Phase-III-Studie im Parallelgruppendesign zum Vergleich der Wirkung von Selexipag (zweimal täglich verabreicht in individuellen Dosierungen von 200 bis 1.600 µg nach Auftitration) gegenüber Placebo bei Patienten mit symptomatischer PAH. In GRIPHON wurden 1.156 Patienten im Verhältnis 1:1 auf zwei Behandlungsgruppen randomisiert und über einen medianen Zeitraum von 70,7 Wochen (Selexipag) bzw. 63,7 Wochen (Placebo) beobachtet.

Für Teilpopulation b konnte keine relevante RCT identifiziert werden, in der Wirksamkeit und Sicherheit von Selexipag mit einer patientenindividuellen Therapie mit Iloprost zur Inhalation nach Maßgabe des Arztes verglichen worden wäre. Ein indirekter Vergleich ist für diese Teilpopulation ebenfalls nicht möglich, da die vorliegenden Iloprost-Studien mit 12 Wochen zu kurz sind, um eine verlässliche Aussage zur Langzeitwirksamkeit der Therapie in einer chronischen Erkrankung wie der PAH zu ermöglichen. Zudem bestehen zwischen GRIPHON, der Zulassungsstudie von Selexipag, und den Iloprost-Studien wesentliche methodische Unterschiede, die der notwendigen Voraussetzung der Vergleichbarkeit im indirekten Vergleich entgegenstehen. Eine Quantifizierung des Zusatznutzens von Selexipag in Teilpopulation b ist somit aufgrund der nicht ausreichenden Datengrundlage nicht möglich. Dennoch ergeben sich verschiedene qualitative Vorteile von Selexipag, die einen Zusatznutzen grundsätzlich begründen und die in den folgenden Abschnitten ausführlich beschrieben werden.

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Die Verzerrungsaspekte der Ergebnisse wurden für die eingeschlossenen Studien einzeln beschrieben und bewertet. Die Bewertung der Aspekte erfolgt getrennt auf Ebene der einzelnen RCT sowie für jeden Endpunkt getrennt. Die Ergebnisse wurden in den im Dossier vorgegebenen Bewertungsbögen dokumentiert.

[2] In einer ereignisgetriebenen Studie werden die Ereignisse in den Behandlungsgruppen gezählt, bis eine vordefinierte Anzahl an Ereignissen (GRIPHON: 331) erreicht wird. Entsprechend kann zuvor keine Studiendauer festgelegt werden. Eine ereignisgetriebene Studie wird als positiv angesehen, wenn die statistische Testung einer neuen Therapie gegenüber einem Komparator die Zeit bis zum Eintreten des Ereignisses signifikant verlängert bzw. das zugehörige Risiko für das Eintreten des Ereignisses signifikant vermindert.

Bei der einzigen in die Nutzenbewertung von Selexipag eingeschlossenen Studie GRIPHON handelt es sich um eine randomisierte, kontrollierte Studie (RCT). Sie entspricht daher dem höchsten Evidenzgrad der Evidenzklassifizierung nach § 5 Nr. 6 der AM-NutzenV. Darüber hinaus weist die Studie sowohl endpunktübergreifend als auch bei allen patientenrelevanten Endpunkten (mit Ausnahme der Analyse der Todesfälle und der Analyse der unerwünschten Ereignisse) ein niedriges Verzerrungspotential und somit eine hohe Aussagesicherheit auf.

Es erfolgt eine detaillierte Darstellung der Ergebnisse der GRIPHON-Studie für die in Abschnitt 4.2.5.2 aufgeführten patientenrelevante Endpunkte. Als Datengrundlage dienen insbesondere der Studienbericht samt Appendices sowie weitere Analysen, die *post hoc* für das vorliegende Nutzendossier durchgeführt wurden.

Zur Bestimmung möglicher Effektmodifikatoren wurden ergänzend zu den Ergebnissen der gesamten Studie auch die Ergebnisse der Subgruppenanalysen nach PAH-spezifischer Therapie zum Zeitpunkt der Basiswerterhebung, WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung, Geschlecht, Rasse, Alter zum Zeitpunkt des Screenings, PAH-Ätiologie und geographische Region dargestellt.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Teilpopulation a

Für die Patientengruppe in Teilpopulation a steht mit Ausnahme der Behandlung mit Selexipag keine zugelassene medikamentöse Therapie zur Verfügung, sodass das abwartende Vorgehen bis zur weiteren Verschlechterung der Erkrankung die einzige Option des behandelnden Arztes innerhalb der Teilpopulation ist. Der Placebo-Arm in GRIPHON entspricht durch das gewählte Studiendesign dem abwartenden Vorgehen und spiegelt daher auch die heutige klinische Praxis in dieser Teilpopulation wider, bei denen das Risiko einer Iloprost-Behandlung deren Nutzen überwiegt[3]. Die Ergebnisse der Studie sind damit geeignet, eine Aussage zum Ausmaß des Zusatznutzens von Selexipag in Teilpopulation a zu treffen.

Selexipag erweist sich durch die Ergebnisse der GRIPHON-Studie als sicheres und gut verträgliches Arzneimittel. Sowohl Wirksamkeit als auch Sicherheit, Verträglichkeit und Tolerabilität von Selexipag waren hierbei über alle untersuchten Subgruppen und Sensitivitätsanalysen konsistent.

[3] Wirksamkeit und Sicherheit von Selexipag wurden in der GRIPHON-Studie gegenüber Placebo verglichen. Eine zuvor begonnene patientenindividuelle Therapie mit ERA oder PDE-5-I konnte als PAH-spezifische Begleittherapie fortgeführt werden, eine bereits bestehende Prostanoidtherapie war jedoch Ausschlusskriterium für die Aufnahme in die Studie. Nach einer Verschlechterung des Patienten konnte die Behandlung mit Prostanoiden dennoch jederzeit eingeleitet werden, führte jedoch direkt oder indirekt zur Wertung als primäres Endpunktereignis und beendete damit die verblindete und randomisierte Behandlungsphase für den Patienten. Insofern entspricht das Studiendesign im Placebo-Arm von GRIPHON einem abwartenden Vorgehen bis zum Fortschreiten der Erkrankung, bei Patienten, deren Erkrankung mit einer Kombinationstherapie aus ERA und/oder PDE-5-I unzureichend kontrolliert ist bzw. die für eine Therapie mit ERA und PDE-5-I nicht infrage kommen.

Das Risiko des Auftretens eines Ereignisses des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ im Vergleich zu Placebo wurde durch Selexipag um 40 %, d. h. deutlich sowie statistisch signifikant verringert (HR [99 % KI]: 0,60 [0,46; 0,78], p<0,0001).

Ebenso wurde im Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ das Risiko eines Ereignisses deutlich, d. h. um 30 % reduziert (HR [99 % KI]: 0,70 [0,50; 0,98], p<0,0031). In konsistenter Weise führte die Anwendung von Selexipag zudem zu einer deutlichen Reduktion des Risikos für das Auftreten einer PAH-bedingten Hospitalisierung (HR [95 % KI]: 0,65 [0,50; 0,86], p=0,0022).

Die Auswertung der CEC-bestätigten Todesfälle wegen PAH bis EOT ergab eine leichte numerische, aber nicht statistisch signifikante Erhöhung des relativen Risikos (HR [99 % KI]: 1,16 [0,70; 1,93], p=0,7153). Durch die Problematik der informativen Zensierung in der Auswertung der Todesfälle bis EOT sind die Ergebnisse dieses Endpunkts aber hoch verzerrt.

Das Risiko des Auftretens eines CEC-bestätigten Todesfalles bis Studienschluss ergab eine statistisch nicht signifikante Verringerung des relativen Risikos (HR [99 % KI]: 1,16 [0,70; 1,93], p=0,7153), während das Risiko eines Todesfalles jeglicher Ursache bis zum gleichen Zeitpunkt in beiden Behandlungsgruppen vergleichbar war (HR [99 % KI]: 0,97 [0,68; 1,39], p<0,4214). Wie die Ergebnisse aus RPSFT- und MSM-Analysen bestätigten, sind die Ergebnisse beider Endpunkte aber aller Voraussicht nach durch den erlaubten Therapiewechsel von Selexipag zu Placebo („*Cross over*“-Effekt) zu Ungunsten von Selexipag verzerrt.

Symptome der PAH (nach Definition des G-BA) traten im Selexipag-Arm deutlich seltener auf als im Placebo-Arm (RR [95 % KI]: 0,85 [0,75; 0,96]), der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (p=0,0081).

Die körperliche Leistungsfähigkeit, gemessen anhand der Veränderung der 6-MWD, verbesserte sich im Vergleich zum Basiswert bis Woche 26 in der Selexipag-Gruppe im Median um 4,0 m (Placebo-Gruppe: −9,0 m) und war damit statistisch signifikant besser als in der Placebo-Gruppe (Behandlungseffekt [99 % KI]: 12,0 m [1 m; 24 m], p=0,0027). Aller Voraussicht nach sind die Ergebnisse des Gesamteffektschätzers der 6-MWD aber durch den hohen Anteil vorbehandelter Patienten in der Studie und den damit verbundenen Ceiling-Effekt verzerrt. Darauf deuten auch die Ergebnisse der entsprechenden Subgruppenanalyse hin.

Der Anteil der Patienten, bei denen im Vergleich zum Basiswert bis Woche 26 keine Verschlechterung der WHO-/NYHA-Klasse beobachtet wurde, war in beiden Studienarmen vergleichbar (OR [99 % KI]: 1,16 [0,81; 1,66], p=0,2843). Bemerkenswert ist in diesem Zusammenhang jedoch, dass mit Ausnahme der Messungen zu Woche 26 und Monat 24 von Woche 8 an der Anteil der Patienten ohne Verschlechterung der WHO-/NYHA-Klasse zu allen Zeitpunkten im Selexipag-Arm besser war als im Placebo-Arm.

Ebenso konnte auch in Bezug auf die absolute Veränderung des CAMPHOR „Symptome“-Scores (Behandlungseffekt [99 % KI]: 0,0 [−1,0; 1,0], p=0,2185) bzw. der „Kurzatmigkeit“-Subskala des gleichen Scores (0,0 [−0,4; 0,0], p=0,1700) in Woche 26 kein signifikanter Unterschied zwischen den Behandlungsgruppen festgestellt werden. Zu erwähnen wäre hier noch, dass sich trotz der Titrationsstrategie von Selexipag, bei der das Auftreten prostacyclinspezifischer Nebenwirkungen *de facto* Voraussetzung zur Bestimmung der patientenindividuell höchsten Dosierung war, auch keine Verschlechterung der Lebensqualität eingestellt hatte. Zudem steht der CAMPHOR-Fragebogen nicht in allen notwendigen Übersetzungen zur Verfügung und konnte daher nur bei einem Teil der Patienten (ca. 40 % der Gesamtstudienpopulation) erhoben werden.

Selexipag erwies sich in GRIPHON als sicheres und gut verträgliches Arzneimittel. Die Behandlung mit Selexipag führte im Vergleich zu Placebo zu ausgeglichenen Gesamtraten der unerwünschten Ereignisse (UE) (RR [95 % KI]: 1,01 [1,00; 1,03]). Das Risiko für das Auftreten schwerwiegender UE (RR [95 % KI]: 0,93 [0,82; 1,06]) und für Therapieabbrüche aufgrund UE (RR [95 % KI]: 0,85 [0,73; 1,00]) wurde tendenziell verringert.

Die Untersuchung der UE von besonderem Interesse ergab in keiner der Gruppierungen (Hämorrhagie, Hypotonie, Störung des Auges) ein statistisch signifikant höheres oder niedrigeres Risiko für das Auftreten eines entsprechenden UE.

Wie aufgrund der Titrationsstrategie und des Wirkmechanismus zu erwarten war, traten unter Selexipag insbesondere im Verlauf der Titrationsphase häufiger Prostacyclin-assoziierte UE auf als unter Placebo (RR [95 % KI]: 1,45 [1,35; 1,55]). Bei den Prostacyclin-assoziierten UE handelt es sich aber vorwiegend um Ereignisse mit meist milder oder moderater Intensität, die mit üblichen Therapiemethoden i. d. R. gut zu behandeln sind. Sie betreffen zudem vorwiegend die Verträglichkeit des Arzneimittels und nicht dessen Sicherheit.

Teilpopulation b

Es konnten keine relevanten RCT für die Nutzenbewertung von Selexipag identifiziert werden, die eine Quantifizierung des Zusatznutzens ermöglichen. Indirekte Vergleiche waren ebenfalls nicht möglich, da sämtliche für Iloprost zur Verfügung stehenden RCT mit einer Studiendauer von 12 Wochen zu kurz waren, um den Zusatznutzen einer Therapie in einer chronisch progredienten Erkrankung wie der PAH bewerten zu können. Zudem bestehen wesentliche Unterschiede zwischen den Iloprost-Studien und GRIPHON, der pivotalen Zulassungsstudie von Selexipag, sodass ein indirekter Vergleich schon aus methodischen Gründen ausgeschlossen ist. Insbesondere liegen keine Studien vor, die eine Aussage zum Risiko des Auftretens von kombinierten Morbiditäts- oder Mortalitätsereignissen oder zur Wirksamkeit von Iloprost in Kombination mit ERA und/oder PDE-5-I ermöglichen.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Teilpopulation a

Morbidität

Der Einsatz von Selexipag führt zu einer deutlichen und gegenüber der Vergleichstherapie bisher nicht erreichten Verbesserung des therapierelevanten Nutzens i. S. v. § 5 Abs. 7 Satz 2 der AM-NutzenV in der Behandlung von Patienten im Anwendungsgebiet von Selexipag, bei denen das *Risiko* einer Behandlung mit Iloprost zur Inhalation dessen *Nutzen* noch überwiegt.

Insbesondere führt Selexipag zu einer deutlichen und klinisch relevanten Verringerung des Risikos für das Auftreten eines Ereignisses des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ und damit zu einer für die Patienten spürbaren Linderung der Erkrankung bzw. einer Abschwächung schwerwiegender Symptome (jeweils im Sinne einer Vermeidung der Progression der Erkrankung).[4] Der für den primären Endpunkt beobachtete Unterschied zwischen den Behandlungsgruppen wird hierbei insbesondere durch eine Vermeidung von Hospitalisierungen, aber auch durch die Verhinderung des Fortschreitens der Erkrankung getrieben. Beide Komponenten für sich genommen würden ebenfalls bereits auf eine spürbare Linderung der Erkrankung (Fortschreiten der Erkrankung) bzw. eine Abschwächung schwerwiegender Symptome (indirekt erkennbar an einer geringeren Anzahl Hospitalisierungen) schließen lassen. Sensitivitätsanalysen, bei denen der primäre Endpunkt unter Ausschluss der vom G-BA kritisch bewerteten Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer Prostanoid- oder dauerhaften O_2-Therapie“ ausgewertet wurden, zeigen, dass die Ereignisse der beiden Komponenten nicht alleine ausschlaggebend für das Ergebnis des gesamten kombinierten Endpunkts sind. Auch nach Ausschluss beider Komponenten aus der Analyse verbleibt ein deutlicher und klinisch relevanter Effekt.

In der Auswertung der Zeit bis zur ersten Hospitalisierung aufgrund einer Verschlechterung der PAH oder bis zum Tod wegen PAH wurde zudem eine deutliche Verringerung des kombinierten Risikos für das Auftreten eines der beiden Ereignisse beobachtet. Der Unterschied zwischen den Behandlungsgruppen war hierbei überwiegend auf die Verringerung des Risikos einer PAH-bedingten Hospitalisierung zurückzuführen. Diese Beobachtung ist konsistent zur Auswertung der Zeit bis zur ersten Hospitalisierung wegen PAH, in der ebenfalls ein verringertes Risiko für das Auftreten einer PAH-bedingten Hospitalisierung durch Selexipag beobachtet wurde. Insgesamt gesehen führt Selexipag demnach zu einer deutlichen und klinisch relevanten Verlängerung der Zeit bis zur stationären Aufnahme wegen der PAH, was mit einer Abschwächung schwerwiegender Symptome gleichzusetzen ist.

[4] Die spürbare Linderung der Erkrankung bzw. die Abschwächung schwerwiegender Symptome im Sinne der AM-NutzenV ergibt sich hierbei durch die Vermeidung der Progression, d. h. im Vergleich zum schwereren Krankheitsstand des Patienten, wenn er nicht mit Selexipag behandelt werden würde.

Selexipag führt darüber hinaus zu einer signifikanten Verbesserung der 6-Minuten-Gehstrecke – einem Maß für die körperliche Leistungsfähigkeit des Patienten – und verringert das kombinierte Risiko des Auftretens der PAH-Symptome Dyspnoe, Synkope, Müdigkeit, Schwindel oder Ödeme. Folglich führt Selexipag zu einer relevanten Abschwächung schwerwiegender bzw. einer Verringerung von nicht-schwerwiegenden Symptomen der PAH.

Im Vergleich zu Placebo blieb die Symptomatik der PAH (bestimmt anhand der WHO-/ NYHA-Klasse) bei einem größeren Anteil der Patienten unter Selexipag stabil bzw. verbesserte sich sogar. Der Unterschied zwischen den Behandlungsgruppen war statistich nicht signifikant.

Insgesamt gesehen führt Selexipag zu einer Abschwächung schwerwiegender Symptome sowie zu einer für die Patienten spürbaren Linderung der Erkrankung (ggf. im Sinne einer Vermeidung einer Verschlechterung) und einer Verbesserung der körperlichen Leistungsfähigkeit. Gemäß § 5 Abs. 7 der AM-NutzenV ist der Zusatznutzen von Selexipag in der Endpunktkategorie Morbidität folglich als „beträchtlich“ anzunehmen.

Mortalität

Im Rahmen der GRIPHON-Studie konnte zwischen den Behandlungsgruppen kein wesentlicher Unterschied in Bezug auf die PAH-bedingte Mortalität und die Gesamtmortalität beobachtet werden. Eine leichte numerische Imbalance der PAH-bedingten Mortalität in der Auswertung bis zum individuellen Behandlungsende (EOT) durch Selexipag im Vergleich zu Placebo ist weitgehend auf eine Verzerrung durch informative Zensierung der Daten zurückzuführen. Überdies war der Unterschied statistisch nicht signifikant.

Auch die Analyse der Gesamtheit der vorhandenen Daten zur PAH-bedingten Mortalität, d. h. bis zum Studienschluss bestätigt, dass kein nachteiliger Effekt auf die Mortalität durch Selexipag vorliegt. In dieser Analyse wurde – trotz einer wahrscheinlichen Verzerrung des Effekts zulasten von Selexipag durch den erlaubten Therapiewechsel von Placebo zu Selexipag nach Ende der Behandlungsphase – eine numerische Reduktion des Mortalitätsrisikos beobachtet, die jedoch nicht statistisch signifikant war.

In der Gesamtbetrachtung sind insbesondere die Ergebnisse der Mortalitätsendpunkte als relevant zu erachten, bei denen der Erhebungszeitraum bis zum Studienschluss reichte, und in deren Auswertung somit sämtliche im Studienzeitraum aufgetretenen Todesfälle berücksichtigt werden.

Das Risiko eines Todesfalles jeglicher Ursache bis Studienschluss war in beiden Studienarmen vergleichbar, es ist aber aufgrund der Ergebnisse von Modellen, die die Effekte durch den erlaubten Therapiewechsel von Selexipag zu Placebo korrigieren, davon auszugehen, dass eine Verzerrung der Todesfälle zu Ungunsten von Selexipag vorliegt.

Ein Zusatznutzen/größerer Schaden von Selexipag in der Endpunktkategorie Mortalität ist daher nicht belegt.

Lebensqualität

Selexipag erweist sich als vergleichbar gegenüber Placebo sowohl in der Auswertung des „Symptome"-Scores des CAMPHOR-Fragebogens als auch in der „Kurzatmigkeit"-Subskala des gleichen Scores. Dies ist insofern bemerkenswert, als trotz der gewählten Titrationsstrategie, die eine Dosierung bis zum Auftreten von für den Patienten nicht mehr erträglichen prostacyclinspezifischen Nebenwirkungen vorsah, keine negative Beeinträchtigung der Lebensqualität beobachtet wurde.

Ein Zusatznutzen/größerer Schaden von Selexipag in der Endpunktkategorie Lebensqualität ist ist daher nicht belegt.

Nebenwirkungen

Die Behandlung mit Selexipag erweist sich insgesamt als ähnlich sicher und verträglich wie die Behandlung mit Placebo und führt zu einer im Vergleich zu Placebo ausgeglichenen Gesamtrate der unerwünschten Ereignisse. Das Risiko für das Auftreten schwerwiegender unerwünschter Ereignisse und für Therapieabbrüche aufgrund unerwünschter Ereignisse wurde tendenziell verringert.

Unter Selexipag traten insbesondere im Verlauf der Titrationsphase häufiger Prostacyclin-assoziierte unerwünschte Ereignisse auf als unter Placebo. Aufgrund der Titrationsstrategie, in der das Auftreten entsprechender UE *de facto* Voraussetzung zur Bestimmung der patientenindividuell optimalen Dosis war, und aufgrund des Wirkmechanismus war ein solcher Unterschied zulasten von Selexipag zu erwarten. Bei den Prostacyclin-assoziierten UE handelt es sich zudem vorwiegend um Ereignisse milder oder moderater Intensität, die mit üblichen Therapiemethoden i. d. R. gut zu beherrschen waren. Sie betreffen zudem vorwiegend die Verträglichkeit von Selexipag und nicht dessen Sicherheit.

Für die Endpunktkategorie „Nebenwirkungen" ist ein Zusatznutzen/größerer Schaden somit nicht belegt.

Fazit

Die für die Teilpopulation a vorgelegte Studie GRIPHON besitzt ein niedriges Verzerrungspotential auf Studien- und Endpunktebene und somit eine hohe Aussagekraft. Es haben sich darüber hinaus keine Hinweise ergeben, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext sprechen. Aus GRIPHON lässt sich somit ein Hinweis auf einen Zusatznutzen ableiten.

Die Ergebnisse aus GRIPHON zeigen, dass Selexipag das erste den Prostacyclin-Signalweg adressierende Arzneimittel in der PAH ist, durch dessen Anwendung eine deutliche und klinisch relevante Reduktion patientenrelevanter Langzeitparameter der Erkrankung erreicht werden konnte. In der Endpunktkategorie Morbidität wurde hierbei ein beträchtlicher Zusatznutzen gegenüber Placebo nachgewiesen, in den übrigen Endpunktkategorien Mortalität, Lebensqualität und Nebenwirkungen ist ein Zusatznutzen/größerer Schaden nicht belegt.

In der Gesamtschau der einzelnen Endpunktkategorien ergibt sich eine bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens, die insbesondere auf der Abschwächung schwerwiegender Symptome, einer für den Patienten spürbaren Linderung der Erkrankung und einer Verbesserung der körperlichen Leistungsfähigkeit beruht. Gemäß § 5 Abs. 7 der AM-NutzenV ist das Ausmaß des Zusatznutzens von Selexipag in der bewerteten Teilpopulation a mit „beträchtlich“ anzunehmen.

Teilpopulation b

Wie bereits zuvor beschrieben, konnte für die Teilpopulation b keine relevante RCT identifiziert werden, die eine direkte Bewertung des Zusatznutzens von Selexipag gegenüber einer patientenindividuellen Anwendung von Iloprost zur Inhalation ermöglicht. Ebenso war auch ein indirekter Vergleich nicht möglich, da die zur Verfügung stehenden Iloprost-Studien mit einer Studiendauer von 12 Wochen zu kurz für eine Aussage zum Zusatznutzen in einer chronisch progredienten Erkrankung wie der PAH waren.

Auch wenn damit keine RCT-basierte klinische Evidenz zur Bestimmung des Ausmaßes des Zusatznutzens vorliegt, so lassen sich dennoch verschiedene qualitative Vorteile von Selexipag gegenüber der patientenindividuellen Anwendung von Iloprost zur Inhalation feststellen:

1. Iloprost ist in Deutschland ausschließlich zur inhalativen Anwendung zugelassen. Aufgrund seiner vergleichweise kurzen Halbwertszeit ist es notwendig, dass Iloprost 6- bis 9-mal täglich inhaliert wird, wobei jede Inhalation mit einem Zeitaufwand von 4 bis 10 Minuten verbunden ist. Hinzu kommt die Zeit, die für die Vorbereitung des Inhalators und den anschließenden Reinigungsvorgang aufgewendet werden muss, sodass sich die Gesamtzeit für die Anwendung von Iloprost auf 15-30 Minuten je Inhalation bzw. auf 1,5-4,5 Stunden (bei 6-9 Inhalationen) je Behandlungstag summiert.

   Die inhalative Anwendung von Iloprost ist zudem nicht komplikationslos. Aufgrund der Ablagerung des Wirkstoff-Aerosols im Inhalator kann es zu größeren und möglicherweise variablen Unterschieden zwischen vorbereiteter und tatsächlich inhalierter Dosierung kommen. Weiterhin haben auch die Tröpfchengröße und die Aerosolkonzentration einen bedeutenden Einfluss auf den Ablagerungsort in den Atemwegen und damit auch auf die Aufnahme in die pulmonalen Gefäße.

   Es ist davon auszugehen, dass die Komplikationen und Probleme der inhalativen Anwendung von Iloprost, insbesondere der große Zeitaufwand, der für die Behandlung aufgewendet werden muss, einen starken negativen Einfluss auf die Lebensqualität der Patienten haben und auch zur Beeinträchtigung der Therapieadhärenz der Patienten führen könnten. Durch seine einfache und komplikationsfreie Anwendung als zweimal täglich oral einzunehmende Tablette besitzt Selexipag die genannten Einschränkungen von inhalativem Iloprost nicht, es ist daher auch nicht von einem nachteiligen Effekt auf Lebensqualität und Therapieadhärenz auszugehen. Dies wird auch durch die Ergebnisse der GRIPHON-Studie bestätigt, in der trotz der gewählten Titrationsstrategie, die *de facto* das

Auftreten von unerwünschen Ereignissen voraussetzte, keine negative Beeinträchtigung der Lebensqualität beobachtet wurde.

2. Im Gegensatz zu Selexipag ist das Prostacyclin-Analogon Iloprost kein selektiver IP-Rezeptor-Agonist, sondern bindet (mit unterschiedlicher Affinität) auch andere Rezeptoren der Prostanoid-Rezeptor-Familie. Neben dem IP-Rezeptor aktiviert Iloprost beispielsweise auch die Prostanoid-Rezeptoren EP_1 und EP_3, mit zum Teil vergleichbarer Potenz. Dies kann aber beispielsweise dazu führen, dass der IP-Rezeptor-vermittelte vasodilatative Effekt durch die vasokonstriktive Aktivität des EP_3-Rezeptors negativ beeinflusst und sogar aufgehoben werden kann.

   Weiterhin kann die fehlende Selektivität für den IP-Rezeptor aber auch ungewollte Nebenwirkungen auslösen bzw. deren Schwere und/oder Dauer negativ beeinflussen. Die Aktivierung der EP_1- und EP_3-Rezeptoren durch nicht-selektive Prostacyclin-Analoga kann u. a. eine Kontraktion der glatten Muskulatur im Magen auslösen, was zu Abdominalschmerzen oder Bauchkrämpfen – häufigen Nebenwirkungen der Prostanoidtherapie – führen kann. Versuche, die mit dem Magenfundus von Ratten durchgeführt wurden, konnten in diesem Zusammenhang zeigen, dass Selexipag und ACT-333679 im Gegensatz zu Iloprost nicht zu einer Kontraktion der Magenmuskulatur führen. Durch die gleichzeitige Behandlung mit einem EP_3-Antagonisten konnten die Iloprost-induzierten Kontraktionen signifikant reduziert werden.

   Die Selektivität Selexipags für den IP-Rezeptor ermöglicht es, Patienten mit einer wirksamen Dosierung zu behandeln, ohne dass die Behandlung durch Tolerabilitätsprobleme eingeschränkt würde.

3. Das vorrangige Ziel der Therapie einer chronisch progressiven Erkrankung wie der PAH ist grundsätzlich immer der Langzeittherapieerfolg. Um diesen sicherzustellen, ist es im Erkrankungsbild der PAH nicht ausreichend, die Wirksamkeit eines Wirkstoffs anhand der Veränderung der 6-MWD – einem Endpunkt, der nachweislich keine Aussagekraft für die Langzeitprognose der PAH besitzt – und nur über einen vergleichsweise kurzen Zeitraum von weniger als 6 Monaten zu bestimmen. Vielmehr ist es notwendig, die Wirksamkeit des Wirkstoffs anhand eines klinisch und patientenrelevanten Endpunkts zu bestimmen, der neben der Mortalität die Zeit bis zur klinischen Verschlechterung verlässlich abbildet und gleichzeitig über einen Zeitraum von mehr als 6 Monaten untersucht wird. Dementsprechend unterscheidet die EMA in der PAH auch zwischen Zulassungen zur kurzfristigen Verbesserung der körperlichen Leistungsfähigkeit und Zulassungen zur Langzeitbehandlung.

   Iloprost ist in Europa ausschließlich für die kurzfristige Verbesserung der körperlichen Leistungsfähigkeit und der Symptomatik zugelassen. Für Selexipag hatte die EMA hingegen explizit eine Zulassung für die Langzeittherapie der PAH erteilt. Durch die Überprüfung der Wirksamkeit in der Zulassung ist demnach sichergestellt, dass Selexipag zum Erreichen eines Langzeittherapieerfolgs geeignet ist, während in Bezug auf die Iloprost-Behandlung unklar ist, ob ein über die kurzfristige Verbesserung der

körperlichen Leistungsfähigkeit hinausgehender Effekt in der Langzeittherapie erreicht wird.

Die dargestellten Argumente zeigen deutlich, dass ein Zusatznutzen von Selexipag gegenüber der zweckmäßigen Vergleichstherapie in Teilpopulation b grundsätzlich vorhanden ist. Da keine direkt vergleichenden RCT vorliegen und auch indirekte Vergleiche nicht möglich sind, lässt die wissenschaftliche Datengrundlage derzeit eine Quantifizierung des Zusatznutzens nicht zu. Gemäß § 5 Abs. 7 der AM-NutzenV ist das Ausmaß des Zusatznutzens von Selexipag in der bewerteten Teilpopulation b daher mit „nicht quantifizierbar“ anzunehmen.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

– Patientenpopulation
– Intervention
– Vergleichstherapie
– Endpunkte
– Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Ziel des vorliegenden Dossiers ist die Darstellung des medizinischen Zusatznutzens für Selexipag im zugelassenen Anwendungsgebiet. Selexipag ist zugelassen zur Langzeitbehandlung der PAH der WHO-/NYHA-Klassen II und III, als Kombinationstherapie bei erwachsenen Patienten, deren Erkrankung mit einem ERA und/oder einem PDE-5-I unzureichend kontrolliert ist, oder als Monotherapie bei Patienten, die für diese Therapien nicht infrage kommen.

Patientenpopulation, Intervention und zweckmäßige Vergleichstherapie (zVT) werden im Folgenden näher erläutert, ebenso die patientenrelevanten Endpunkte im Anwendungsgebiet und die ggf. vorgelegten Studientypen.

Patientenpopulation und zweckmäßige Vergleichstherapie

Die Fachinformation zu Uptravi® [1] definiert das Anwendungsgebiet von Selexipag folgendermaßen:

„*Uptravi ist indiziert für die Langzeitbehandlung der pulmonal arteriellen Hypertonie (PAH) bei erwachsenen Patienten der WHO-Funktionsklasse (WHO-FC) II bis III entweder als Kombinationstherapie bei Patienten, deren Erkrankung mit einem Endothelin-Rezeptor-Antagonisten (ERA) und/oder einem Phosphodiesterase-5(PDE-5)-Inhibitor unzureichend*

kontrolliert ist oder als Monotherapie bei Patienten, die für diese Therapien nicht infrage kommen.

Die Wirksamkeit wurde bei PAH, einschließlich idiopathischer und erblicher PAH, PAH in Assoziation mit Bindegewebserkrankungen und PAH in Assoziation mit korrigierten einfachen angeborenen Herzfehlern nachgewiesen."

Das genannte Anwendungsgebiet definiert eine Population von erwachsenen Patienten mit PAH in den WHO-/NYHA-Klassen II und III, bei denen aufgrund des Fortschreitens ihrer Erkrankung der Bedarf nach einer weiteren Behandlungsoption besteht. Durch die Definition des Anwendungsgebietes müssen Patienten in der Zielpopulation von Selexipag hierbei entweder a) bereits mit einer Kombination aus ERA und PDE-5-I behandelt sein oder b) für eine (zusätzliche) Therapie mit den beiden Wirkstoffen nicht infrage kommen (vgl. Tabelle 4-1). Die Behandlung der Patienten in der Zielpopulation mit einem Endothelin-Rezeptor-Antagonisten (ERA) und/oder einem Phosphodiesterase-5-Inhibitor (PDE-5-I) zusätzlich zur eventuell bereits bestehenden Therapie statt der Behandlung mit Selexipag ist damit durch die Zulassung von Selexipag ausgeschlossen.

Tabelle 4-1: Behandlungssituationen im Anwendungsgebiet von Selexipag

Patient bereits behandelt mit …	**Anwendung von Selexipag zugelassen (ja/nein)**
ERA + PDE-5-I, die Erkrankung wird durch die bisherige Therapie unzureichend kontrolliert	Ja
ERA, die Erkrankung wird durch die bisherige Therapie unzureichend kontrolliert	Nur wenn PDE-5-I als zusätzliche Behandlungsoption nicht infrage kommen
PDE-5-I, die Erkrankung wird durch die bisherige Therapie unzureichend kontrolliert	Nur wenn ERA als zusätzliche Behandlungsoption nicht infrage kommen
therapienaiv	Nur wenn die Behandlung mit ERA und PDE-5-I nicht infrage kommt

Zur Entscheidung der Frage, ab wann die Erkrankung eines Patienten durch die Therapie mit einem ERA und/oder einem PDE-5-I nicht mehr ausreichend kontrolliert ist bzw. ab wann er für eine (zusätzliche) Therapie mit einem oder beiden dieser Wirkstoffe nicht infrage kommt, handelt es sich um eine Entscheidung des behandelnden Arztes, die für jeden Einzelfall patientenindividuell getroffen werden muss. Festgelegte, quantifizierbare Kriterien, an denen der Arzt sich bei dieser Entscheidung orientieren kann, bestehen nicht.

Unabhängig von den Überlegungen zum Anwendungsgebiet geht Actelion aber davon aus, dass fast alle Patienten im Anwendungsgebiet von Selexipag bereits mit einer Kombinationstherapie aus einem ERA und einem PDE-5-I behandelt werden. Die Annahme hierbei ist, dass nur sehr wenige der mit einem ERA oder einem PDE-5-I vorbehandelten bzw. therapienaiven Patienten nicht für eine (zusätzliche) Therapie mit ERA und/oder PDE-5-I infrage kommen.

Ausgehend von den Ergebnissen des Beratungsgespräches mit dem G-BA wurde für die Nutzenbewertung von Selexipag folgende zweckmäßige Vergleichstherapie definiert (vgl. Modul 3):

Patientenindividuell optimierte medikamentöse Therapie nach Maßgabe des Arztes und unter Berücksichtigung des jeweiligen Zulassungsstatus (im Weiteren: patientenindividuell optimierte Therapie, PIT).

Im Erkrankungsbild der PAH sind die unten stehenden Wirkstoffe bzw. Wirkstoffklassen zugelassen (vgl. auch Modul 2). Aufgrund der Einschränkungen, die durch das Anwendungsgebiet von Selexipag vorgegeben werden, sowie aufgrund der spezifischen Anwendungsgebiete der einzelnen Wirkstoffe bzw. aufgrund des spezifischen Nutzen/Risiko-Verhältnisses, das mit ihrer Anwendung einhergeht, sind nicht alle Wirkstoffe gleichermaßen im Rahmen der zweckmäßigen Vergleichstherapie PIT in Betracht zu ziehen:

- ERA (Macitentan, Bosentan, Ambrisentan)

  Die Anwendung von Selexipag statt eines ERA ist durch das Anwendungsgebiet von Uptravi® ausgeschlossen. Insofern kann ein ERA zwar Teil einer patienten-individuellen Therapiestrategie sein, die im Zusammenhang mit der Anwendung von Selexipag oder der zweckmäßigen Vergleichstherapie durchgeführt wird, kommt jedoch selbst nicht als möglicher Komparator von Selexipag innerhalb der zweckmäßigen Vergleichstherapie PIT in Betracht.

  Der Grund hierfür ist, dass das Anwendungsgebiet von Uptravi® die Anwendung von Selexipag ausschließlich bei Patienten erlaubt, die entweder für eine Behandlung mit einem ERA nicht infrage kommen oder deren bestehende Therapie bereits einen ERA einschließt (siehe oben).

- PDE-5-I (Sildenafil, Tadalafil)

  Die Anwendung von Selexipag statt eines PDE-5-I ist durch das Anwendungsgebiet von Uptravi® ausgeschlossen. Insofern kann ein PDE-5-I zwar Teil einer patientenindividuellen Therapiestrategie sein, die im Zusammenhang mit der Anwendung von Selexipag oder der zweckmäßigen Vergleichstherapie durchgeführt wird, kommt jedoch selbst nicht als möglicher Komparator von Selexipag innerhalb der zweckmäßigen Vergleichstherapie PIT in Betracht.

  Der Grund hierfür ist, dass das Anwendungsgebiet von Uptravi® die Anwendung von Selexipag ausschließlich bei Patienten erlaubt, die entweder für eine Behandlung mit einem PDE-5-I nicht infrage kommen, oder deren bestehende Therapie bereits einen PDE-5-I einschließt (siehe oben).

- Stimulatoren der löslichen Guanylatcyclase (Riociguat)

  Einziger derzeit zugelassener Vertreter dieser Wirkstoffklasse ist Riociguat. Riociguat kann aufgrund der Hinweise des G-BA zur Festlegung der zweckmäßigen Vergleichstherapie kein Bestandteil der zweckmäßigen Vergleichstherapie PIT sein.

Der Grund hierfür ist, dass Riociguat „*aufgrund der bisher kurzen Marktverfügbarkeit nicht als in der praktischen Anwendung bewährt angesehen und […] von der zweckmäßigen Vergleichstherapie ausgenommen*“ ist [2].[5]

- Inhalierte Prostanoide (Iloprost)

  Iloprost gehört zur Wirkstoffgruppe der Prostanoide, die ähnlich wie der IP-Rezeptor-Agonist Selexipag auf den Prostacyclin-Signalweg wirken. Iloprost ist als mögliche Komponente der zVT PIT grundsätzlich heranzuziehen. Anders als bei den Wirkstoffgruppen der ERA bzw. der PDE-5-I schließt das Anwendungsgebiet von Selexipag eine alternative Behandlung mit Iloprost statt Selexipag nicht aus. Auch aus der Festlegung des G-BA ergeben sich keine Gründe, die für einen Ausschluss von Iloprost aus der zVT sprechen.

  Iloprost ist in Deutschland zur Behandlung der PAH ausschließlich in der inhalativen Darreichungsform zugelassen. Aufgrund seiner relativ kurzen Halbwertszeit muss Iloprost 6- bis 9mal täglich inhaliert werden, wobei jede Inhalation etwa 4–10 Minuten dauert [4]. Hinzu kommt die Zeit, die für die Vorbereitung des Inhalators und den anschließenden Reinigungsvorgang aufgewendet werden muss, sodass die Gesamtzeit, die je Inhalation aufgebracht werden muss, etwa 15-30 Minuten beträgt. Die inhalative Anwendung von Iloprost ist somit als sehr zeitaufwendig für den Patienten anzusehen. Zudem ist sie auch nicht komplikationslos, da es aufgrund der Ablagerung des Wirkstoff-Aerosols im Inhalator zu größeren und möglicherweise variablen Unterschieden zwischen vorbereiteter und tatsächlich inhalierter Dosierung kommen kann. [5]. Weiterhin haben auch die Tröpfchengröße und die Aerosol-konzentration einen bedeutenden Einfluss auf den Ablagerungsort in den Atemwegen und damit auch auf die Aufnahme in die pulmonalen Gefäße [5].

  Iloprost ist in Europa ausschließlich für Patienten in der WHO-/NYHA-Klasse III und damit nur in einem Teil des Anwendungsgebietes von Selexipag zugelassen. Da in der zweckmäßigen Vergleichstherapie nur Arzneimittel berücksichtigt werden können, die eine Zulassung für das jeweilige Anwendungsgebiet besitzen, kann Iloprost daher für Patienten in der WHO-/NYHA-Klasse II kein Bestandteil der zVT PIT sein.

  Weiterhin wird aufgrund der genannten Einschränkungen durch die aufwendige und nicht komplikationslose inhalative Anwendung bei einigen Patienten im Anwendungs-gebiet von Selexipag das Risiko der Iloprost-Behandlung deren Nutzen noch überwiegen. Iloprost wird daher nur bei denjenigen Patienten als Bestandteil der zVT PIT herangezogen, bei denen ein positives Nutzen/Risiko-Verhältnis für die

[5] Der Wortlaut dieses Hinweises ergibt sich aus der Formulierung in der Niederschrift zu Macitentan vom 14.12.2015 i. V. mit der dort vom G-BA getroffenen Aussage, dass „*die zweckmäßige Vergleichstherapie zum Wirkstoff Selexipag analog angepasst wird*“. In der Niederschrift zum Beratungsgespräch vom 03.03.2015 war der entsprechende Hinweis folgendermaßen formuliert: „*Mit Riociguat und Macitentan stehen für das vorliegende Anwendungsgebiet zugelassene, in der Versorgung noch recht neue Behandlungsoptionen zur Verfügung, deren therapeutischer Stellenwert derzeit noch nicht abschließend beurteilbar ist.*“ [3]

Anwendung vorliegt. Insbesondere ist hierbei auch die ausschließliche Zulassung von Iloprost in WHO-/NYHA-Klasse III zu beachten.

- Parenterale Prostanoide (Treprostinil, Epoprostenol)

  Treprostinil und Epoprostenol sind zwei weitere Wirkstoffe aus der Gruppe der Prostanoide, die als mögliche Bestandteile der zVT PIT im Anwendungsgebiet von Selexipag grundsätzlich zur Verfügung stehen. In Europa sind beide Wirkstoffe nur zur Behandlung von Patienten in der WHO-/NYHA-Klasse III bzw. IV (IV nur Epoprostenol) und ausschließlich in ihren parenteralen[6] Darreichungsformen zugelassen [6, 7]. Die Anwendung als subkutane oder intravenöse Injektion ist jedoch mit verschiedenen Risiken und Limitationen verbunden.

  So erfordert die Anwendung des Prostacyclins Epoprostenol, das wegen seiner kurzen Halbwertszeit von 2 bis 6 Minuten (unter physiologischen Bedingungen [7]) kontinuierlich i. v. zugeführt werden muss, die Implantation eines zentralen Venenkatheters, der mit einer stets zu tragenden Pumpe verbunden ist. Zudem sollte die verwendete Epoprostenol-Lösung anhaltend möglichst auf 2-8 C gekühlt werden, da die maximale Verabreichungszeit der Lösung bei Zimmertemperatur ansonsten auf 12 Stunden sinkt [7]. Aufgrund dieser Bedingungen erweist sich die Verabreichung von Epoprostenol letztlich als kompliziert, belastend für den Patienten, verursacht zusätzliche Kosten für das Gesundheitssystem und ist mit zum Teil lebensbedrohlichen Nebenwirkungen wie Katheterinfektionen oder Katheter-assoziierter Sepsis verbunden [8, 9]. Überdies kann es durch Versagen der Pumpe oder Ablösung des zentralen Venenkatheters zu einer lebensbedrohlichen Unterbrechung der Infusion kommen [8].

  Für die intravenöse Darreichungsform von Treprostinil gelten ähnliche Einschränkungen, wie sie auch für Epoprostenol beschrieben wurden. Im Gegensatz zu Epoprostenol steht für die Anwendung von Treprostinil i. v. jedoch eine implantierbare Pumpe zur Verfügung, die für eine etwas einfachere und weniger komplikationsreiche Anwendung im Vergleich zur Epoprostenol-Anwendung sorgt. Dennoch ist auch Treprostinil i. v. mit dem Auftreten von Infektionen der Blutbahn verbunden [6, 10]. Zudem ist Treprostinil i. v. gemäß aktueller Fachinformation insbesondere denjenigen Patienten vorbehalten, die eine Unverträglichkeit gegenüber der subkutanen Darreichung aufweisen und für die die Risiken einer intravenösen Anwendung – hier ist insbesondere die Möglichkeit einer Infektion des Blutkreislaufs ausgehend von der zentralvenösen Injektionsstelle zu nennen – als vertretbar scheinen [6]. Stattdessen sollte bevorzugt die subkutane Infusion von Treprostinil durchgeführt werden, bei der ebenfalls eine tragbare Pumpe zum Einsatz kommt. Bei der Anwendung von Treprostinil kann es wie auch bei der Anwendung von Epoprostenol aufgrund des Versagens der Pumpe zu lebensbedrohlichen Unterbrechungen der

[6] Hier gemeint: intravenös oder subkutan.

Infusion kommen. Bei der subkutanen Infusion treten zudem sehr häufig zum Teil nicht tolerierbare Schmerzen an der Infusionsstelle auf [6, 11].

Aufgrund der Komplikationen und teils lebensbedrohenden Risiken, die mit der parenteralen Anwendung von Treprostinil und Epoprostenol einhergehen, werden beide Wirkstoffe in Deutschland i. d. R. erst bei bereits eingetretener Verschlechterung der Patienten nach WHO-/NYHA-Klasse IV und damit außerhalb des Anwendungsgebietes von Selexipag eingesetzt. Gemäß der deutschen Kommentierung der europäischen Leitlinien von European Respiratory Society (ERS) und European Society for Cardiology (ESC)[7] sollten insbesondere die intravenösen Prostanoide sogar erst dann eingesetzt werden, wenn konkret ein Rechtsherzversagen bei den Patienten droht oder bereits manifest ist [13].

Insofern stellen die parenteralen Prostanoide in dem Sinne keine Alternative zur Therapie mit Selexipag dar, sondern werden erst dann eingesetzt, wenn sich die Patienten trotz maximaler Therapie mit ERA, PDE-5-I und Selexipag bzw. Iloprost in die WHO-/NYHA-Klasse IV verschlechtern und eine weitere Intensivierung der Therapie notwendig ist. Epoprostenol und Treprostinil ersetzen in dieser Therapiesituation die Behandlung mit Selexipag bzw. Iloprost, eine Weiterbehandlung mit Selexipag oder Iloprost wäre für diese Patienten aber keine Option.

Die Wirkstoffe Treprostinil und Epoprostenol werden daher zum Nachweis des Zusatznutzens von Selexipag als Bestandteil der zVT PIT nicht herangezogen.

Nach Betrachtung aller im Krankheitsbild der PAH zugelassenen Wirkstoffe verbleibt damit einzig Iloprost zur Inhalation als möglicher medikamentöser Bestandteil der vom G-BA für Selexipag definierten zweckmäßigen Vergleichstherapie „patientenindividuell optimierte Therapie“. Zu beachten ist hierbei aber, dass auch die Behandlung mit Iloprost nicht für alle Patienten im Anwendungsgebiet von Selexipag in Betracht kommt, da nach Berücksichtigung des individuellen Erkrankungszustandes bei manchen Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegen kann. Auch ist Iloprost nicht für das vollständige Anwendungsgebiet von Selexipag zugelassen, sondern nur zur Behandlung von Patienten in WHO-/NYHA-Klasse III.

Für diese Patientengruppe steht mit Ausnahme der Behandlung mit Selexipag keine zugelassene medikamentöse Therapie zur Verfügung, sodass das abwartende Vorgehen bis zur weiteren Verschlechterung der Erkrankung die einzige Therapieoption des behandelnden Arztes für diese Teilpopulation des Anwendungsgebietes von Selexipag ist.

Zusammenfassend ergeben sich für die Nutzenbewertung von Selexipag aus den zuvor genannten Gründen und in Abhängigkeit von den zur Verfügung stehenden Therapieoptionen die folgenden Teilpopulationen:

---

[7] Die deutsche Kommentierung der Kölner Konsensus-Konferenz von 2011 beruht noch auf der europäischen Leitlinie von 2009. Die europäische Leitlinie wurde zwischenzeitlich aktualisiert, die Empfehlungen zur Anwendung der Prostanoide bei Patienten im Anwendungsgebiet von Selexipag sind im Wesentlichen jedoch gleich geblieben [12].

- Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt.

  Mögliche Therapieoptionen innerhalb der zVT PIT:

  abwartendes Vorgehen bis zur weiteren Verschlechterung der PAH

- Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt.
  - Mögliche Therapieoptionen innerhalb der zVT PIT:

    patientenindividuelle Therapie mit Iloprost (zur Inhalation) nach Maßgabe des Arztes

Intervention

Die vorgelegte Nutzenbewertung erfolgt für den Wirkstoff Selexipag (Uptravi®). Uptravi® ist zugelassen in Kombinationstherapie bei Patienten, deren Erkrankung mit einem ERA und/oder einem PDE-5-I unzureichend kontrolliert ist, sowie als Monotherapie bei Patienten, bei denen eine Monotherapie mit ERA oder PDE-5-I nicht infrage kommt [1]. Uptravi® wird zweimal täglich oral eingenommen; die Dosierung ist patientenindividuell unterschiedlich und wird in einer initialen Titrationsphase bestimmt. In dieser Phase wird Selexipag beginnend mit einer Dosierung von 200 µg zweimal täglich so lange in wöchentlichen Schritten von 200 µg zweimal täglich aufdosiert, bis die Maximaldosierung von 1.600 µg zweimal täglich erreicht wird oder für den Patienten nicht mehr tolerierbare Nebenwirkungen auftreten. In diesem Fall wird die Dosierung wieder auf die zuletzt vertragene Dosis reduziert und im weiteren Verlauf als Erhaltungsdosis weitergeführt.

Patientenrelevante Endpunkte

Im Beratungsgespräch zu Selexipag am 13.05.2015 hatte der G-BA insbesondere folgende Endpunkte genannt, die er als patientenrelevant im Anwendungsgebiet betrachtet [3]:

- Mortalität
- Morbidität, bei der vorliegenden Indikation z. B.
  - Dyspnoe
  - Körperliche Funktionsfähigkeit bzw. Belastbarkeit
  - Hospitalisierung aufgrund von Komplikationen
- (gesundheitsbezogene) Lebensqualität
  - Die Verwendung sowohl eines krankheitsspezifischen als auch eines generischen Fragebogens (bevorzugt SF-36) zur Lebensqualität wird empfohlen.

- Nebenwirkungen:
    - Unerwünschte Ereignisse, differenziert nach Krankheitskonzepten/Organsystemen und Schweregrad
    - Schwerwiegende unerwünschte Ereignisse
    - Therapieabbrüche wegen unerwünschter Ereignisse

Darüber hinaus hatte der G-BA im Beratungsgespräch zu Macitentan vom 02.12.2012 [14] auch noch die Symptome der PAH (Luftnot/Dyspnoe, körperliche Leistungseinschränkung, Schwindel, Synkopen, Müdigkeit und Ödeme) als patientenrelevanten Endpunkt benannt.

Die folgende Auflistung zeigt in einer Übersicht die zur Nutzenbewertung von Selexipag herangezogenen patientenrelevanten Endpunkte sowie ihre Zuordnung zu den Endpunktkategorien Morbidität („Verbesserung des Gesundheitszustands“, „Verkürzung der Krankheitsdauer“), Mortalität („Verlängerung des Überlebens“), Lebensqualität („Verbesserung der Lebensqualität“) und Nebenwirkungen („Verringerung von Nebenwirkungen“) gemäß § 3 Abs. 1 des 5. Kapitels der VerfO des G-BA:

- Mortalität/Morbidität (Kombinationsendpunkte):
    - Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis (zu den Einzelkomponenten des kombinierten Endpunktes vgl. auch Abschnitte 4.2.5.2 und 4.3.1.3.1.1)
    - Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH

  Beide genannten Kombinationsendpunkte werden innerhalb dieser Nutzenbewertung der Endpunktkategorie Morbidität zugerechnet (für eine ausführliche Begründung siehe Abschnitt 4.4.1).
- Mortalität:
    - Zeit bis zum Tod (Studienschluss)
    - Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)
    - Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)
- Morbidität:
    - Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH
    - Symptome der PAH; diese sind nach Definition des G-BA:
        - Dyspnoe
        - Synkope
        - Ödeme
        - Schwindel

        - Müdigkeit
    - Absolute Veränderung der 6-MWD
    - Ausbleibende Verschlechterung der WHO-/NYHA-Klasse
    - Absolute Veränderung des Borg-Dyspnoe-Index
- Lebensqualität:
    - Absolute Veränderung des CAMPHOR „Symptome“-Scores
    - Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores
- Nebenwirkungen:
    - Unerwünschte Ereignisse
    - Schwerwiegende unerwünschte Ereignisse
    - Therapieabbrüche aufgrund unerwünschter Ereignisse
    - Unerwünschte Ereignisse von besonderem Interesse
        - Hämorrhagie
        - Hypotonie
        - Störung des Auges
    - Prostacyclin-assoziierte unerwünschte Ereignisse:

Studiendesign

Die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention liefern randomisierte kontrollierte Studien (RCTs) [15]. In ihrer Leitlinie zur klinischen Erprobung medizinischer Produkte in der PAH geht die EMA zudem davon aus, dass Untersuchungen zu neuen Produkten, bei denen das beabsichtigte Anwendungsgebiet die Langzeitverbesserung von Morbidität und Mortalität des Patienten beinhalten soll, über einen Zeitraum von mindestens 6 Monaten durchgeführt werden sollen [16].

Im Anwendungsgebiet von Selexipag sind zwei Patientenpopulationen zu differenzieren, die sich insbesondere auch durch die Verfügbarkeit therapeutischer Optionen im Rahmen der zVT PIT unterscheiden. Im Folgenden wird darauf eingegangen, welche Daten für die beiden zu differenzierenden Subpopulationen des Anwendungsgebietes vorgelegt werden.

Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegt (Teilpopulation a)

Wie im vorherigen Abschnitt unter „Patientenpopulation und zweckmäßige Vergleichstherapie“ beschrieben, steht im Anwendungsgebiet von Selexipag lediglich Iloprost zur Inhalation als mögliche Arzneimittelbehandlung im Rahmen der zVT PIT zur Verfügung, dies jedoch auch nur bei denjenigen Patienten, bei denen der Nutzen einer Iloprost-Behandlung

deren Risiko überwiegt. Für Patienten, bei denen das Nutzen/Risiko-Verhältnis der Iloprost-Behandlung noch negativ ist, verbleibt als einzige therapeutische Möglichkeit das abwartende Vorgehen bis zum Fortschreiten der Erkrankung.

Im vorliegenden Dossier werden für die Subpopulation der Patienten, bei denen aufgrund des Erkrankungszustandes des Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegt, die Ergebnisse der pivotalen Zulassungsstudie GRIPHON präsentiert. Bei GRIPHON handelt es sich um eine multizentrische, doppelblinde, randomisierte, ereignisgetriebene Phase-III-Studie im Parallelgruppendesign.

Die Wirksamkeit von Selexipag wurde in GRIPHON anhand der Verlängerung der Zeit bis zum ersten, von einem verblindeten „*critical event commitee*“ (CEC) bewerteten kombinierten Morbiditäts- oder Mortalitätsereignis bestimmt. Bei der „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ handelt es sich um einen höchst patientenrelevanten Kombinationsendpunkt, der in enger Abstimmung mit den Zulassungsbehörden und medizinischen Experten entwickelt und in ähnlicher Form bereits in der Zulassungsstudie von Macitentan (SERAPHIN) eingesetzt wurde (vgl. Abschnitt 4.2.5.2). Bisher zugelassene PAH-Arzneimittel wurden hingegen nahezu ausschließlich aufgrund der kurzfristigen Verbesserung der körperlichen Leistungsfähigkeit, gemessen als 6-Minuten-Gehstrecke (*six minute walking distance*, 6-MWD) zugelassen.

Mit 181 teilnehmenden Zentren in 39 Ländern, 1.156 randomisierten Patienten und einem Beobachtungszeitraum von bis zu 4,3 Jahren stellt GRIPHON die derzeit größte und längste RCT dar, die im Bereich der PAH durchgeführt wurde. Die Prostanoid-Arzneimittel wurden mit Ausnahme des in Deutschland nicht zugelassenen Wirkstoffs Beraprost dagegen ausschließlich in Studien mit einer Dauer von 12 bis 16 Wochen untersucht (vgl. Abbildung 4-1).

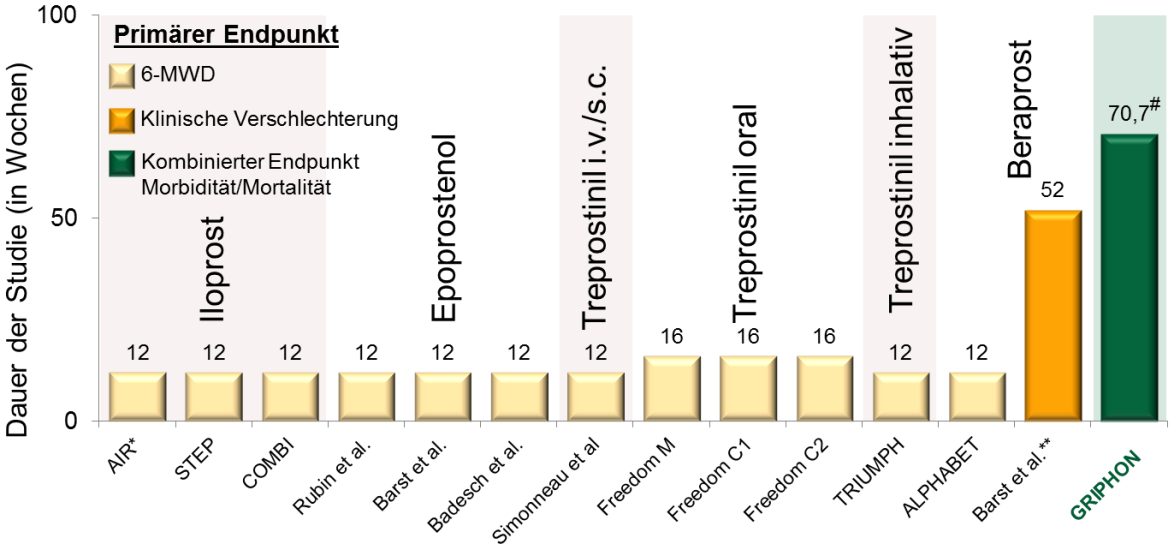


Abbildung 4-1: Vergleich der Studiendauer von GRIPHON mit Studien zur Untersuchung der Wirksamkeit der Prostanoide
Adaptiert nach Galiè et al. 2013 [17]. *primärer Endpunkt: 6-MWD und WHO-/NYHA-Klasse; **primärer Endpunkt definiert als erstes Auftreten von Tod, Lungentransplantation, Initiierung chronischer Therapie mit Epoprostenol i. v. oder chronischer Therapie mit Prostacyclin-Analoga oder eine mindestens 25 %ige Verschlechterung der maximalen Sauerstoffaufnahme im Vergleich zum Basiswert; #mediane Dauer im Selexipag-Arm (maximale Dauer war 217 Wochen)

Wirksamkeit und Sicherheit von Selexipag wurden in der GRIPHON-Studie gegenüber Placebo verglichen. Hierbei ist zu beachten, dass eine zuvor begonnene patientenindividuelle Therapie mit ERA oder PDE-5-I als PAH-spezifische Begleittherapie fortgeführt werden konnte, sofern die Dosierung der Arzneimittel über einen Zeitraum von mindestens 3 Monaten vor Beginn der Studie stabil gehalten worden war. Die Behandlung mit Prostanoiden als PAH-spezifische Begleittherapie war in der Studie nicht erlaubt. Sofern er dies für erforderlich hielt, konnte der Arzt jedoch jederzeit eine Therapie mit einem Prostanoid einleiten, was jedoch direkt (als Initiierung einer intravenösen oder subkutanten Prostanoidtherapie) oder indirekt (als Notwendigkeit zusätzlicher PAH-Therapie im Zusammenhang mit einer Verschlechterung der 6-MWD ≥15 % bei Patienten in den WHO-/NYHA-Klassen III und IV) zur Wertung als Ereignis des primären Endpunkts und damit zum Ende der verblindeten und randomisierten Behandlungsphase führte (*end-of-treatment*, EOT).

Insofern entspricht das Studiendesign im Placebo-Arm von GRIPHON einem abwartenden Vorgehen bis zum Fortschreiten der Erkrankung bei Patienten, deren Erkrankung mit einer Kombinationstherapie aus ERA und/oder PDE-5-I unzureichend kontrolliert ist bzw. die für eine Therapie mit ERA und PDE-5-I nicht infrage kommen. GRIPHON ist somit geeignet, eine Aussage zum Zusatznutzen von Selexipag gegenüber der zVT PIT für diejenigen Patienten im Anwendungsgebiet von Selexipag zu treffen, bei denen aufgrund des Erkrankungszustandes das Risiko einer Behandlung mit Iloprost dessen Nutzen überwiegt.

Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der Nutzen einer Behandlung mit Iloprost deren Risiko überwiegt (Teilpopulation b)

In der Teilpopulation des Anwendungsgebietes von Selexipag, in der der Nutzen der Iloprost-Behandlung deren Risiko überwiegt, ist Iloprost zur Inhalation die einzige mögliche Therapieoption im Rahmen der festgelegten zVT PIT. Infolgedessen muss der Zusatznutzen von Selexipag für diese Patientenpopulation im Vergleich zur patientenindividuellen Therapie mit Iloprost zur Inhalation verglichen werden.

Wie die für dieses Dossier durchgeführte Literaturrecherche ergab (vgl. Abschnitt 4.3.1.1), existieren jedoch keine RCT, die einen direkten Vergleich von Selexipag mit inhalativem Iloprost ermöglichten. Die Durchführung eines indirekten Vergleichs der beiden Wirkstoffe ist ebenfalls nicht möglich, da die Studiendauer der bisher durchgeführten Iloprost-Studien (AIR, STEP, COMBI; vgl. Abbildung 4-1 und Galiè et al. 2015 [18]) mit 12 Wochen zu kurz für die Bewertung des Zusatznutzens einer Langzeittherapie bei einer chronischen Erkrankung wie der PAH ist. Auch aus methodischen Gründen ist ein indirekter Vergleich nicht möglich, da zwischen den Iloprost-Studien und GRIPHON wesentliche Unterschiede im Studiendesign bestehen.

Eine Quantifizierung des Zusatznutzens von Selexipag in Teilpopulation b ist somit aufgrund der nicht ausreichenden Datengrundlage nicht möglich. Dennoch ergeben sich verschiedene qualitative Vorteile von Selexipag, die einen Zusatznutzen grundsätzlich begründen. Die genannten Vorteile sind ausführlich in Abschnitt 4.4.2 beschrieben.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Population

Die Studienpopulation für das Anwendungsgebiet wird gemäß der Fachinformation von Uptravi® [1] wie folgt bestimmt:

„Langzeitbehandlung der PAH der WHO-/NYHA-Klassen II und III, als Kombinationstherapie bei erwachsenen Patienten, deren Erkrankung mit einem ERA und/oder einem PDE-5-I unzureichend kontrolliert ist, oder als Monotherapie bei Patienten, die für diese Therapien nicht infrage kommen“

Grundlegende Voraussetzung zur Berücksichtigung einer Studie bei der vorliegenden Nutzenbewertung ist somit auch, dass Selexipag und die zweckmäßige Vergleichstherapie:

- als zusätzliche Therapie zu einer Zweifachkombination aus ERA und PDE-5-I angewendet wurden;
- als zusätzliche Therapie zu einer Behandlung mit einem ERA angewendet wurden, sofern eine Behandlung mit PDE-5-I für diese Patienten nicht infrage kommt;
- als zusätzliche Therapie zu einer Behandlung mit einem PDE-5-I angewendet wurden, sofern eine Behandlung mit ERA für diese Patienten nicht infrage kommt;
- als Monotherapie angewendet wurden, sofern eine Behandlung mit ERA und PDE-5-I nicht infrage kommt.

Darüber hinaus wird innerhalb der definierten Studienpopulation zwischen den folgenden Teilpopulationen unterschieden (vgl. Abschnitt 4.2.1):

- Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt (Teilpopulation a).
- Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt (Teilpopulation b).

Intervention

Eingeschlossen wurden Studien nur dann, wenn Behandlungsmodus und Dosierung von Selexipag den Angaben in der Fachinformation von Uptravi® [1] entsprachen (vgl. Abschnitt 4.2.1).

Zweckmäßige Vergleichstherapie

Zweckmäßige Vergleichstherapie im Anwendungsgebiet von Selexipag ist die patientenindividuell optimierte medikamentöse Therapie nach Maßgabe des Arztes und unter Berücksichtigung des Zulassungsstatus.

Je nach untersuchter Subpopulation stehen im Rahmen der zVT PIT unterschiedliche Behandlungsoptionen zur Verfügung:

- Für Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt, ist das abwartende Vorgehen bis zur weiteren Verschlechterung der PAH die einzige mögliche Therapieoption im Rahmen der PIT.

  Abwartendes Vorgehen ist hierbei so definiert, dass der Patient mit Ausnahme der Fortführung der bisherigen PAH-spezifischen Therapie keine weitere Behandlung erhält, dass gleichzeitig aber engmaschig überwacht wird, ob sich der Erkrankungszustand des Patienten verändert. Sobald eine Verschlechterung des Erkrankungszustandes beobachtet wird, sollte die Veränderung der Therapie, ggf. durch eine Intensivierung ermöglicht werden.

  Studien konnten nur dann in die Nutzenbewertung eingeschlossen werden, wenn diese Definition erfüllt war. Ein Therapiearm wurde auch dann als abwartendes Vorgehen gewertet, wenn die Patienten im betroffenen Arm Placebo erhalten hatten, ansonsten die genannten Kriterien aber erfüllt waren.

- Für Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt, ist die patientenindividuelle Therapie mit Iloprost (zur Inhalation) nach Maßgabe des Arztes die einzige mögliche Therapieoption im Rahmen der PIT.

  Studien konnten nur dann in die Nutzenbewertung eingeschlossen werden, wenn Behandlungsmodus und Dosierung von Iloprost den Angaben in der Fachinformation entsprachen [4]. Die Behandlung mit Iloprost sollte patientenindividuell optimiert und nach Maßgabe des Arztes durchgeführt werden.

Patientenrelevante Endpunkte

Die Nutzenbewertung von Selexipag erfolgt anhand patientenrelevanter Endpunkte. Um in den Studienpool zur Bewertung des Ausmaßes des Zusatznutzens von Selexipag eingeschlossen werden zu können, muss eine RCT mindestens einen im Erkrankungsbild der PAH als patientenrelevant anzusehenden Endpunkt untersucht haben. Hierzu gehören insbesondere die in Abschnitt 4.2.1 genannten Endpunkte.

Studiendesign

Die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention liefern randomisierte kontrollierte Studien (RCTs) [15]. Das vorliegende Dossier verwendet ausschließlich Daten aus methodisch adäquaten und der jeweiligen Fragestellung angemessenen RCTs. Dieses Vorgehen ermöglicht Aussagen zum medizinischen Zusatznutzen mit hoher Ergebnissicherheit.

Studiendauer

Die Leitlinie der EMA zur klinischen Erprobung medizinischer Produkte in der PAH sieht vor, dass zur Zulassung eines Arzneimittels zur Langzeitbehandlung der PAH Studien mit einem kombinierten Morbiditäts-/Mortalitäts-Endpunkt vorgelegt werden sollen [16]. Die Dauer dieser Studien soll 6 Monate nicht unterschreiten. Da das dem Nutzendossier zugrunde

liegende Anwendungsgebiet die Langzeitbehandlung der PAH umfasst und es sich zudem bei der PAH um eine chronische, unheilbare Erkrankung handelt, wird für die vorliegende Nutzenbewertung ebenfalls eine Mindeststudiendauer von 6 Monaten angenommen.

Sonstige Ausschlusskriterien

Für die Nutzenbewertung wurden Studienberichte, Vollpublikationen sowie – soweit möglich – in Studienregistereinträgen veröffentlichte Ergebnisse berücksichtigt. Konferenzbeiträge wurden nicht berücksichtigt. Publikationen, die nicht in deutscher oder englischer Sprache vorlagen, wurden ebenfalls ausgeschlossen.

Aus den gerade ausführlich beschriebenen Kriterien ergeben sich folgende Ein- und Ausschlusskriterien:

Tabelle 4-2: Einschlusskriterien für Studien zum Nachweis von Nutzen und Zusatznutzen

Kurzbezeichnung	Einschlusskriterium
E1	Erwachsene Patienten mit pulmonal arterieller Hypertonie E1a[1]: Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt (Teilpopulation a) E1b[1]: Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt (Teilpopulation b)
E2	Prüfintervention: Selexipag; zweimal tägliche orale Anwendung von 200 bis 1.600 µg Die Dosierung von Selexipag wird patientenindividuell bestimmt, begonnen wird mit einer Dosierung von 200 µg zweimal täglich. Anschließend wird die Dosierung in wöchentlichen Schritten von jeweils 200 µg zweimal täglich so lange aufdosiert, bis entweder die Maximaldosierung von 1.600 µg zweimal täglich erreicht oder für den Patienten nicht mehr tolerierbare Nebenwirkungen auftreten. Als Erhaltungsdosis wird im Folgenden dann diejenige Dosierung angewendet, die vom Patienten ohne wesentliche Nebenwirkungen toleriert wird.
E3	Vergleichsintervention: E3a[1]: Abwartendes Vorgehen bis zur weiteren Verschlechterung der PAH (gilt nur für Teilpopulation a). Sofern das abwartende Vorgehen ausreichend im Studiendesign abgebildet war, konnte auch ein Placebo-Vergleich als abwartendes Vorgehen gewertet werden. E3b[1]: Patientenindividuelle Therapie mit Iloprost (zur Inhalation) nach Maßgabe des Arztes (gilt nur für Teilpopulation b). 6-9 Einzelinhalationen je Tag, je nach individuellem Bedarf und Verträglichkeit. Die Anwendung sollte mit einer Dosierung von 2,5 µg je Einzelinhalation begonnen werden. Wenn die Dosis von 2,5 µg gut vertragen wurde, war sie im Folgenden auf 5 µg zu steigern und auf diesem Niveau zu halten. Im Falle einer schlechten Verträglichkeit der 5-µg-Dosis sollte diese auf 2,5 µg reduziert werden.
E4	Erhebung mindestens eines patientenrelevanten Endpunktes
E5	Studien im RCT-Design
E6	Studiendauer ≥6 Monate (26 Wochen)

[1]) In der Informationsbeschaffung wurde zunächst eine gemeinsame Literatur- und Studienregistersuche für beide Teilpopulationen durchgeführt. Dadurch wurden Studien identifiziert, die grundsätzlich geeignet sind, einen Zusatznutzen von Selexipag im Erkrankungsbild der PAH zu bestimmen. Die Kriterien E1a und E1b bzw. E3a und E3b wurden anschließend auf den Pool verbliebener und als grundsätzlich relevant erachteter Studien angewendet.

Tabelle 4-3: Ausschlusskriterien für Studien zum Nachweis von Nutzen und Zusatznutzen

Kurzbezeichnung	Ausschlusskriterium
A1	Population nicht wie definiert (E1 nicht erfüllt, E1a bzw. E1b nicht erfüllt[1])
A2	Prüfintervention nicht wie definiert (E2 nicht erfüllt)
A3	Vergleichsintervention nicht wie definiert (E3a bzw. E3b nicht erfüllt[1])
A4	Kein patientenrelevanter Endpunkt wird berichtet (E4 nicht erfüllt)
A5	Keine Studie im RCT-Design (E5 nicht erfüllt)
A6	Studiendauer kürzer als 6 Monate (E6 nicht erfüllt)
A7	Keine Volltextpublikation[1] verfügbar
A8	Publikation in anderer Sprache als Deutsch oder Englisch

[1]) In der Informationsbeschaffung wurde zunächst eine gemeinsame Literatur- und Studienregistersuche für beide Teilpopulationen durchgeführt. Dadurch wurden Studien identifiziert, die grundsätzlich geeignet sind, einen Zusatznutzen von Selexipag im Erkrankungsbild der PAH zu bestimmen. Die Kriterien E1a und E1b bzw. E3a und E3b wurden anschließend auf den Pool verbliebener und als grundsätzlich relevant erachteter Studien angewendet.

[2]) Studienbericht, Volltextpublikation oder ausführliche Ergebnisdarstellung in Studienregister.

Es wurden Studien eingeschlossen, die alle Einschlusskriterien, aber keines der Ausschlusskriterien erfüllen. Die Einschlusskriterien gelten als erfüllt, wenn in der zu bewertenden Studie mindestens 80 % der Patienten die Einschlusskriterien erfüllen [15].

Es konnten aber auch Studien in die Nutzenbewertung eingeschlossen werden, bei denen weniger als 80 % der Patienten die Einschlusskriterien erfüllten, sofern die betreffenden Patienten in Form einer Subgruppenanalyse ausgewertet werden konnten.

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten

4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt. Für MEDLINE und EMBASE erfolgte die Suche über die Oberfläche von Ovid, für die Cochrane-Datenbank wurde die Cochrane-eigene Suchoberfläche verwendet.

Die verwendeten Suchstrategien und Suchbegriffe sind in Anhang 4-A getrennt für jede Datenbank detailliert beschrieben. Die Suche wurde am 03.07.2015 in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt. Eine

Aktualisierungssuche fand am 04.05.2016 statt. Die Aufstellung in Anhang 4-A zeigt die Ergebnisse der aktuellsten Suche.

Die Suchstrategien wurden unter Verwendung eines Pools von verwandten Suchbegriffen in Bezug auf die Einschlusskriterien der jeweiligen Recherche für jede Datenbank angepasst. Die Suchstrategien beinhalten jeweils eine Kombination von Freitext und Begriffen der Medical Subject Headings (MeSH) bzw. verwandter Konzepte in anderen Datenbanken.

Da aufgrund der erst kürzlich erlangten Zulassung von Selexipag nur mit einer geringen Anzahl an Studien mit diesem Wirkstoff zu rechnen war und um die Sensitivität der Recherche nicht unnötig einzuschränken, wurde die Suchstrategie in einem mehrstufigen Verfahren verfeinert. Erster Schritt war die Suche mit Suchbegriffen für die Intervention in Verbindung mit einem RCT-Filter. Nur wenn bei diesem Schritt eine größere Anzahl möglicher Studien identifiziert worden wäre, wäre die Suchstrategie in den nächsten Schritten mit Suchbegriffen für die Vergleichstherapie und für die Indikation verfeinert worden. Dies war jedoch nicht der Fall.

Aufgrund der allgemein niedrigen Trefferzahlen wurden *a priori* keine weiteren Einschränkungen in der Suchstrategie, wie z. B. die Beschränkung auf Humanstudien, vorgenommen.

Die Suche ergab 75 Treffer (nach Entfernung von Dubletten). Von diesen Treffern waren relevant:

- Teilpopulation a

  Ein Treffer; die betreffende Publikation beschreibt die Ergebnisse der GRIPHON-Studie.

- Teilpopulation b

  Kein Treffer.

Zusätzlich zur bibliografischen Literaturrecherche wurde auch Sekundärliteratur in Form von Meta-Analysen, systematischen Reviews oder ähnlichen Publikationstypen ausgewertet, die entweder durch die Literaturrecherche identifiziert wurden oder Actelion anderweitig bekannt waren. Diese Auswertung ergab jedoch keine zusätzlichen Studien.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT

(Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), Klinische Prüfungen PharmNet.Bund (http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: http://apps.who.int/trialsearch/) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Neben der systematischen Suche in medizinischen Datenbanken wurde auch eine Suche in öffentlich zugänglichen Studienregistern durchgeführt.

Das Ziel dieser Suche war die Identifikation aller klinischen Studien mit Selexipag, die zum Nachweis des Zusatznutzens herangezogen werden können. Die Suche in den Registern galt der Identifizierung aller abgeschlossenen, abgebrochenen und laufenden Studien.

Die Suche wurde in den oben angeführten Studienregistern clinicaltrials.gov, WHO International Clinical Trials Registry Plattform (WHO ICTRP), EU Clinical Trials Registry (EU-CTR) und Klinische Prüfungen PharmNet.Bund durchgeführt.

Die Suche wurde am 03.07.2015 durchgeführt; es wurden keine generellen Einschränkungen bei den genannten Recherchen vorgenommen. Eine Aktualisierungssuche fand am 09.05.2016 statt.

Suchstrategien wurden für alle Register und sämtliche Rechercheziele separat entwickelt. Als Suchbegriffe wurden insbesondere verschiedene Abwandlungen der Prüfintervention Selexipag verwendet. Die in den Studien verwendete Sprache wurde nicht eingegrenzt. Anhang 4-B gibt einen Überblick über die verwendeten Suchstrategien sowie die Anzahl der Suchergebnisse für die jeweils aktuellste Recherche.

Die Suche ergab 19 Treffer (nach Entfernung von Dubletten). Von diesen Treffern waren relevant:

- Teilpopulation a

  Zwei Treffer; die betreffenden Studieneinträge sind der Zulassungsstudie GRIPHON- zuzuordnen.

- Teilpopulation b

  Kein Treffer.

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechercheschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Auswahl einzuschließender Studien aus den Treffern in den medizinischen Datenbanken und den Studienregistern wurde nach Entfernung von Dubletten und unter Anwendung der festgelegten Ein- und Ausschlusskriterien (vgl. Abschnitt 4.2.2) im Drei-Schritt-Verfahren durchgeführt. In den ersten beiden Schritten wurden Studien identifiziert, die grundsätzlich geeignet sind, eine Aussage zum Zusatznutzen im Anwendungsgebiet von Selexipag zu treffen. Im dritten Schritt wurde dann überprüft, welche dieser Studien eine Aussage zum Zusatznutzen in den beiden Teilpopulationen des Anwendungsgebietes ermöglichen. Dementsprechend kamen die Ausschlusskriterien A1a/b bzw. A3a/b erst in diesem Schritt zum Tragen.[8]

Im ersten Schritt wurde primär ein Titel-/Abstract-Screening durchgeführt, gleichzeitig konnten Treffer in diesem Schritt aber auch ausgeschlossen werden, wenn klar erkennbar war, dass es sich um Abstracts oder Reviews und nicht um die Vollpublikation einer Studie handelte.

Im zweiten Schritt wurde ein Volltext-Screening mit zuvor nicht ausgeschlossenen Treffern durchgeführt, die Ausschlussgründe hierzu wurden in Anhang 4-C dokumentiert.

[8] Das allgemeiner gefasste Ausschlusskriterium A1 wiederum konnte während aller drei Schritte berücksichtigt werden.

Aufgrund der geringen Anzahl an Studien, die in den ersten beiden Schritten identifiziert wurden (ausschließlich die Zulassungsstudie GRIPHON wurde identifiziert), erfolgte für den dritten Schritt keine separate Dokumentation des Ausschlusses von Studien im Anhang sowie in Modul 5.

Die Vollständigkeit der Literaturrecherche wurde zusätzlich durch die Auswertung von Reviews überprüft, die nur aufgrund von Ausschlusskriterium A5 (Keine Studie im RCT-Design) ausgeschlossen wurden. Diese Reviews wurden dahin gehend gescreent, ob sie zusätzliche, in der Literaturrecherche nicht identifizierte Studien zitierten.

Die einzelnen Schritte zur Selektion der Studien wurden unabhängig von zwei Personen durchgeführt. Sofern sich bezüglich des Ein- und Ausschlusses von Treffern abweichende Beurteilungen ergaben, wurden diese im abschließenden Konsensverfahren durch die beiden am Verfahren beteiligten Personen geklärt.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

<u>A: Verzerrungsaspekte der Ergebnisse auf Studienebene</u>

– Erzeugung der Randomisierungssequenz *(bei randomisierten Studien)*
– Verdeckung der Gruppenzuteilung *(bei randomisierten Studien)*
– zeitliche Parallelität der Gruppen *(bei nicht randomisierten vergleichenden Studien)*
– Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren *(bei nicht randomisierten vergleichenden Studien)*
– Verblindung des Patienten sowie der behandelnden Personen
– ergebnisgesteuerte Berichterstattung
– sonstige Aspekte

<u>B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene</u>

– Verblindung der Endpunkterheber
– Umsetzung des ITT-Prinzips
– ergebnisgesteuerte Berichterstattung
– sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor,

wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Validität der in diesem Dossier präsentierten Evidenz wurde in Übereinstimmung mit den oben beschriebenen Anforderungen und den Kriterien in Anhang 4-F bewertet. Dazu wurden die Verzerrungsaspekte der Ergebnisse der einzigen eingeschlossenen Studie beschrieben. Die Bewertung der einzelnen Aspekte wurde für jeden Endpunkt mithilfe des Bewertungsbogens aus Anhang 4-F dokumentiert.

Die auf Studienebene berücksichtigten Faktoren waren insbesondere: Erstellung einer Randomisierungsreihenfolge, Geheimhaltung der Gruppenzuweisung, Verblindung von Patient und Prüfarzt, Vergleichbarkeit der Eigenschaften der Studienarme bei Studienbeginn, ergebnisorientierte oder selektive Berichterstattung und Berücksichtigung von weiteren für die Prognose relevanten Faktoren. Die auf Endpunktebene berücksichtigten Faktoren waren insbesondere: Verblindung der Endpunkterheber, Umsetzung des ITT-Prinzips, ergebnisorientierte oder selektive Berichterstattung und andere Faktoren, die zur Verzerrung der Ergebnisse beitragen könnten.

Das Verzerrungspotential wurde entweder als niedrig oder als hoch eingestuft. Wenn eine Verzerrung der relevanten Ergebnisse mit hoher Wahrscheinlichkeit ausgeschlossen werden konnte, wurde dies als niedriges Verzerrungspotential angesehen. Wenn von der Verzerrungsquelle eine relevante Beeinflussung der Ergebnisse zu erwarten war, wurde dies als hohes Verzerrungspotential angesehen. Dieser Fall musste auch angenommen werden, wenn keine ausreichenden Informationen zur Bewertung des Verzerrungspotentials vorlagen (z. B. wenn

kein Studienbericht, sondern ausschließlich Publikationen vorlagen). Lagen hingegen zusätzlich Unterlagen – z. B. Informationen von Zulassungsbehörden – vor, wurden diese in die Bewertung des Verzerrungspotentials einbezogen.

Die Verzerrungspotentialbewertung wurde zur Bewertung der Aussagekraft der Nachweise verwendet und führte nicht zum Ausschluss von Daten.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)[9]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[10] bzw. STROBE-Statements[11] folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

In Modul 4 wird zur Nutzenbewertung ausschließlich die GRIPHON-Studie berücksichtigt. Hierbei handelt es sich um eine randomisierte und kontrollierte klinische Studie. Die Darstellung dieser Studie entspricht den in diesem Dossier zitierten Items 2b bis 14 nach CONSORT sowie dem CONSORT-Flow-Chart (vgl. Anhang 4-E).

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten

[9] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[10] Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.
[11] Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Durch die Informationsbeschaffung konnte ausschließlich für Teilpopulation a des Anwendungsgebietes von Selexipag eine relevante RCT identifiziert werden – die Zulassungsstudie GRIPHON (vgl. Abschnitt 4.3.1.1.4). Für Teilpopulation b konnten keine relevanten RCT identifiziert werden.

Sofern in den folgenden Abschnitten konkrete Auswertungen oder Operationalisierungen benannt werden, beziehen sich diese daher – wenn nicht anders angegeben – auf GRIPHON.

Patientencharakteristika

Für die zur Nutzenbewertung herangezogene Studie GRIPHON wurden die unten stehenden demographischen Werte und Basiswert-Charakteristika dargestellt. Zusätzlich wurden die PAH-Diagnose und die PAH-spezifische Medikation zum Zeitpunkt der Basiswerterhebung aufgeführt.

Die Darstellung erfolgte jeweils für das *full analysis set* (FAS-Population).

Demographische Werte

- Geschlecht
- Alter/Altersgruppen
- Gewicht, Größe, *body mass index* (BMI)
- Rasse/Ethnie
- Geographische Region

Basiswert-Charakteristika

- Zeit seit der PAH-Diagnose
- Ätiologie der PAH
- WHO-/NYHA-Klasse

- 6-Minuten-Gehstrecke
- Borg-Dyspnoe-Index
- Systolischer/diastolischer Blutdruck
- Herzfrequenz

Patientenrelevante Endpunkte

Vorbemerkung

Obwohl sich die Prognose der von pulmonal arterieller Hypertonie Betroffenen im Verlauf der letzten Jahrzehnte durch die Einführung verschiedener neuer Arzneimittel zunehmend verbessert hat, handelt es sich bei dieser Erkrankung weiterhin um eine chronisch progressive und schlussendlich nicht heilbare Krankheit. Den bestmöglichen Endpunkt zur Bestimmung der Wirksamkeit neuer Therapien in der PAH stellt entsprechend unzweifelhaft die Verringerung der Mortalität dar [19-21]. Der Einsatz dieses Endpunktes in klinischen Studien ist jedoch sehr schwierig, wenn nicht sogar unmöglich, da die zum Nachweis signifikanter Behandlungseffekte notwendige Größe der Studienpopulation und die notwendige (lange) Studiendauer bei einem seltenen Leiden mit vergleichsweise geringer Krankheitsinzidenz und -prävalenz realistischer Weise nicht erreicht werden können. [20, 22, 23].

Des Weiteren erscheint eine Studie mit Mortalität als primärem Endpunkt ethisch zweifelhaft, da dies implizieren würde, dass bei Patienten bis zum Erreichen des Endpunktes auch bei einer Krankheitsverschlechterung keine Therapieveränderung vorgenommen werden dürfte, obwohl die zur Zeit verfügbaren Behandlungsoptionen dies ermöglichen würden. Hierbei ist anzumerken, dass es in der PAH im Gegensatz zu anderen kardiovaskulären Erkrankungen vergleichsweise selten zum plötzlichen Herztod (Mortalitätsereignis) kommt; i. d. R. geht bei der PAH dem Mortalitäts- ein Morbiditätsereignis voraus. Auf dieses Morbiditätsereignis muss dann reagiert werden, um idealerweise das Mortalitätsereignis zu verhindern. Somit ist festzuhalten, dass dem allgemeinen Anliegen, Mortalitätsendpunktstudien bereitzustellen, bei einer seltenen Erkrankung wie der PAH aufgrund epidemiologischer und ethischer Beweggründe nicht entsprochen werden kann.

Entsprechend wurden alle derzeitigen Arzneimittel in der Therapie der PAH nicht aufgrund ihrer Effekte auf den Endpunkt „Mortalität“ zugelassen; ihre Wirksamkeit wurde stattdessen in Kurzzeitstudien (4-6 Monate) über die Verbesserung des 6-Minuten-Gehtests (6-MWT) bzw. der damit verbundenen Strecke (6-MWD) als vermeintlichem Surrogat für klinisch relevante therapeutische Wirksamkeit bestimmt [18]. Eine Ausnahme hierzu bildet der bereits nutzenbewertete Wirkstoff Macitentan, der aufgrund einer Studie mit einem kombinierten Morbiditäts-/Mortalitätsendpunkt zugelassen wurde (siehe weiter unten).

Im gleichen Maße, wie sich das Verständnis über Entstehung und Pathologie sowie Diagnose und Behandlung der PAH weiterentwickelt hat, sind auch die Erwartungen an die in klinischen Studien verwendeten Endpunkte gestiegen. Die Veränderung der 6-MWD (Δ6-MWD) wird aufgrund nachgewiesener fehlender Relevanz und Verlässlichkeit dieses Endpunktes zur Prognose des (dauerhaften) Therapieerfolgs als ungenügend zur Beurteilung

der Wirksamkeit neuer PAH-Therapien angesehen. So konnte beispielsweise anhand von Auswertungen der SERAPHIN-Daten gezeigt werden, dass zwar die *absolute 6-MWD* zum Zeitpunkt der Basiswerterhebung sowie nach 6 Monaten mit dem Risiko des PAH-assoziierten Todes oder der Hospitalisierung assoziiert ist, die *Veränderung der 6-MWD* über diesen Zeitraum aber keine verlässliche Prognose des Langzeiterfolgs der Therapie ermöglicht [24]. Aufgrund dieser anerkannten Probleme wurde die 6-MWD infolgedessen u. a. auch von Vertretern des G-BA im Rahmen der Nutzenbewertung des Wirkstoffs Riociguat kritisch diskutiert (siehe Anhörung zum Wirkstoff Riociguat vom 08.09.2014 [25]). Im Verlauf der vergangenen Jahre wurde daher ein globaler Konsens für die Verwendung eines kombinierten Endpunkts, der auf klar messbaren, klinisch relevanten Ereignissen basiert, erreicht [26]. Dieser Konsens spiegelt sich in den Empfehlungen des 4. (Dana Point 2009) und des 5. Weltsymposiums (Nizza 2013) zur pulmonalen Hypertonie wider [27, 28]. Nach Ansicht der EMA, dargelegt in ihrer Leitlinie zur Durchführung klinischer Studien in der PAH [16], kann die Veränderung der 6-MWD zudem ausschließlich nur dann als primärer Endpunkt in konfirmatorischen Studien eingesetzt werden, wenn die beantragte Indikation des Arzneimittels auf die kurzfristige Verbesserung der körperlichen Leistungsfähigkeit abzielt bzw. beschränkt wird. Die Beantragung einer Indikation zur Langzeittherapie von Patienten mit PAH ist wiederum nur dann möglich, wenn als primärer Endpunkt ein klinisch relevanter Endpunkt verwendet wird, der neben der Mortalität die Zeit bis zur klinischen Verschlechterung verlässlich abbildet und gleichzeitig über einen Zeitraum von mehr als 6 Monaten (abhängig von der konkreten Zusammensetzung des Endpunktes) untersucht wird.

Folgende Tabelle 4-4 zeigt die patientenrelevanten Endpunkte, die in der GRIPHON-Studie erhoben wurden, sowie ihre Zuordnung zu den Endpunktkategorien Mortalität, Morbidität, Lebensqualität und Nebenwirkungen. Die Patientenrelevanz der einzelnen Endpunkte ist nachfolgend zur Tabelle näher erläutert.

Tabelle 4-4: Patientenrelevante Endpunkte in GRIPHON

Endpunktkategorie	**Endpunkt**
Morbidität/Mortalität	Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis[1]
	Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH[1]
Mortalität	Zeit bis zum Tod (Studienschluss)
	Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)
	Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)
Morbidität	Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH
	Symptome der PAH[2]
	Absolute Veränderung der 6-MWD
	Ausbleibende Verschlechterung der WHO-/NYHA-Klasse
	Absolute Veränderung des Borg-Dyspnoe-Index
Lebensqualität	Absolute Veränderung des CAMPHOR „Symptome“-Scores
	Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores
Nebenwirkungen	Unerwünschte Ereignisse
	Schwerwiegende unerwünschte Ereignisse
	Therapieabbrüche wegen unerwünschter Ereignisse
	Unerwünschte Ereignisse von besonderem Interesse[3]
	Prostacyclin-assoziierte unerwünschte Ereignisse

[1]) Kombinationsendpunkt
[2]) Der Endpunkt wurde *post hoc* auf Grundlage des Beratungsgespräches zu Selexipag [3] eingeführt. Als Symptome der PAH wurden die bevorzugten Bezeichnungen (*preferred terms*, PT) der unerwünschten Ereignisse ausgewertet, die den folgenden PAH-Symptomen zuzuordnen sind: Dyspnoe, Synkope, Ödeme, Schwindel, Müdigkeit.
[3]) Unerwünschte Ereignisse von besonderem Interesse waren: Hämorrhagie, Hypotonie, Störung des Auges

Mortalität

Die Mortalität wird in der PAH als Endpunkt von besonderer Relevanz angesehen, da es sich bei der PAH um eine zurzeit unheilbare und letztlich tödliche Erkrankung handelt. Das Hauptziel der Behandlung dieser Erkrankung stellt somit unstrittig die Verlängerung des

Überlebens dar [16]. Grundsätzlich wäre es sinnvoll, die Mortalität direkt zu bestimmen, da eine Verbesserung der Symptomatik nicht zwangsläufig mit verlängerter Überlebenszeit verbunden sein muss; andererseits konnte ein Zusammenhang zwischen den Morbiditätsereignissen des primären Endpunkts in GRIPHON und dem Überleben beobachtet werden, siehe unten. Der Nachweis einer signifikanten Verbesserung des Überlebens durch ein PAH-Arzneimittel ist zudem mit Herausforderungen verbunden, die zunehmend auch von den regulatorischen Stellen anerkannt werden. Insbesondere gestaltet sich die Rekrutierung einer genügend großen Anzahl Patienten, die an dieser seltenen Erkrankung leiden, auch über einen längeren Zeitraum als sehr schwierig [16]. Als ein Beispiel, welche Anzahl an Patienten notwendig ist, um verlässlich einen Überlebensvorteil im kardiovaskulären Bereich nachweisen zu können, kann die Behandlung der Herzinsuffizienz dienen. In einer kürzlich veröffentlichten Studie von McMurray et al. mussten die Organisatoren fast 10.000 Patienten rekrutieren, um einen Vorteil von LCZ696 gegenüber dem Komparator Enalapril in Bezug auf die Gesamtmortalität nachweisen zu können [29]. Entsprechend wurde bisher in der Indikation PAH keine RCT mit Mortalität jeglicher Ursache als primärem Endpunkt durchgeführt. Die meisten bislang durchgeführten Studien waren zudem von zu kurzer Dauer und wiesen zu niedrige Teilnehmerzahlen auf, um einen Behandlungseffekt auf das Überleben nachweisen zu können [23, 28].

Zusätzlich zu den Herausforderungen in der praktischen Durchführung machen es auch ethische Überlegungen fast unmöglich, einen statistisch signifikanten Effekt auf die Mortalität nachzuweisen. In Anbetracht der oftmals rapiden symptomatischen und klinischen Verschlechterung unbehandelter PAH-Patienten und der gleichzeitigen Verfügbarkeit wirksamer PAH-Arzneimittel wäre es unethisch, den Patienten im Placebo-Arm einer Langzeitstudie die noch verbleibenden PAH-Behandlungsmöglichkeiten vorzuenthalten [30]. Um diesen ethischen Bedenken im Studiendesign zu begegnen, wurde vorgeschlagen, neue PAH-Arzneimittel plus Standardtherapie mit Placebo plus Standardtherapie zu vergleichen [28]. Ein solches Vorgehen würde jedoch zu einer Verringerung des Gruppenunterschiedes führen, insbesondere bei einem Endpunkt wie der Mortalität. Da im Verlauf der Studie relativ wenige Ereignisse (d. h. Tod als erstes Ereignis) erwartet werden, würde dies folglich den Nachweis eines signifikanten Behandlungsunterschiedes deutlich erschweren [31]. Als mögliche Lösung dieser Problematik wurde in den Empfehlungen von Dana Point 2008 die Kombination verschiedener patientenrelevanter Endpunkte einschließlich der Mortalität in Form des Endpunktes *time to clinical worsening* (TTCW) vorgeschlagen [28]. Dieser Endpunkt bzw. der verwandte, jedoch im Gegensatz zum TTCW zentral im Kontext der Langzeitverbesserung zu verstehende Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ werden im folgenden Abschnitt ab S. 58 ff. beschrieben.

In GRIPHON wurde die Mortalität in insgesamt 5 Endpunkte ausgewertet:

- Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis
- Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH

- Zeit bis zum Tod (Studienschluss)
- Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)
- Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)

Bei allen Endpunkten, in denen die PAH-bezogene Mortalität ausgewertet wurde, erfolgte eine Begutachtung der Todesfälle durch das *Critical Event Committee*. Hierbei wurde in verblindeter und standardisierter Weise bewertet, ob der Todesfall tatsächlich in ursächlicher Verbindung zur PAH stand.

Analysiert wurde jeweils die Zeit bis zum Tod, wobei zu beachten ist, dass in den kombinierten Endpunkten das bzw. die jeweils anderen Ereignisse des Endpunkts vor Eintreten des Todesfalles auftreten konnten. Die Auswertung erfolgte entweder innerhalb des Zeitraums von der Randomisierung bis 7 Tage nach Einnahme der letzten Studienmedikation (*end-of-treatment*, EOT) oder von der Randomisierung bis zum Ende der Studie.

Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis

Bei einer unheilbaren und progredienten Erkrankung wie der PAH ist die Mortalität üblicherweise hoch. Fast immer kommt es aber im Verlauf der Erkrankung zunächst zu einer zunehmenden, oft beträchtlichen Morbidität, die eine Verschlechterung des klinischen Zustandes widerspiegelt. Als Beispiele der Morbidität wären hier ungeplante Krankenhauseinweisungen aufgrund der PAH, die Verschlechterung der WHO-/NYHA-Klasse oder auch die Verschlechterung der körperlichen Leistungsfähigkeit zu nennen. Entsprechend sehen sowohl EMA als auch die entsprechende Arbeitsgruppe des 5. PH-Weltsymposiums (Nizza 2013) die Zeit bis zum Eintreten von Ereignissen der klinischen Verschlechterung als validen Endpunkt zur klinischen Überprüfung der Wirksamkeit von Arzneimitteln in der PAH an [16, 27].

Wie im nächsten Abschnitt ab S. 61 ff. detailliert ausgeführt wird, besteht zunehmend Skepsis gegenüber Relevanz, Sensitivität und Interpretation der Ergebnisse des 6-Minuten-Gehtests (6-MWT). Die Arbeitsgruppe „*New Trial Designs and Potential Therapies for PAH*“ des 5. PH-Weltsymposiums kam daher zu der Schlussfolgerung, dass die Erhebung der körperlichen Leistungsfähigkeit in pivotalen PAH-Studien ausschließlich als sekundärer Endpunkt erfolgen sollte [27]. Stattdessen wird für Phase-III-Studien ein kombinierter Endpunkt empfohlen, der als „Zeit bis zur klinischen Verschlechterung“ (*time to clinical worsening*, TTCW) bezeichnet wurde und sowohl Mortalitäts- als auch objektiv bestimmbare „harte“ Morbiditätsereignisse, die im Zusammenhang zur PAH-Progression stehen, enthalten sollte [27]. Innerhalb einer vierstufigen Endpunkthierarchie, die während des 5. Weltsymposiums zur Erstellung von Therapieempfehlungen für die PAH verwendet wurde, werden solche Morbiditäts-/ Mortalitäts-Endpunkte der höchsten Stufe 1 („*true clinical efficacy measure*“) zugerechnet [17]. Im Gegensatz dazu wird die 6-MWD lediglich als nicht-validiertes Surrogat in der PAH angesehen und entsprechend Stufe 3 oder 4 der Endpunkthierarchie zugeordnet.

Der Endpunkt „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“, der als primärer Wirksamkeitsendpunkt in der GRIPHON-Studie gewählt wurde (dargestellt

in Abbildung 4-2), entspricht weitestgehend den Empfehlungen des 4. PH-Weltsymposiums in Dana Point 2008, die im 5. PH-Weltsymposium in Nizza 2013 bestätigt wurden [27, 28]. Im Vergleich zu den Dana-Point-Empfehlungen geht die Endpunktdefinition in GRIPHON sogar noch einen Schritt weiter, indem die Komponente „Fortschreiten der Erkrankung“ durch ein zweites Kriterium ergänzt wird, das insbesondere für Patienten in den WHO-/NYHA-Klassen III und IV relevant ist (Verschlechterung der 6-MWD in Kombination mit der Notwendigkeit zusätzlicher PAH-Therapie). Die in Dana Point 2008 formulierten Empfehlungen sahen hierzu nur ein Kriterium vor (Verschlechterung der 6-MWD in Kombination mit einer Verschlechterung der WHO-/NYHA-Klasse), das jedoch bei Patienten in der WHO-/NYHA-Klasse IV nicht erreicht werden kann.

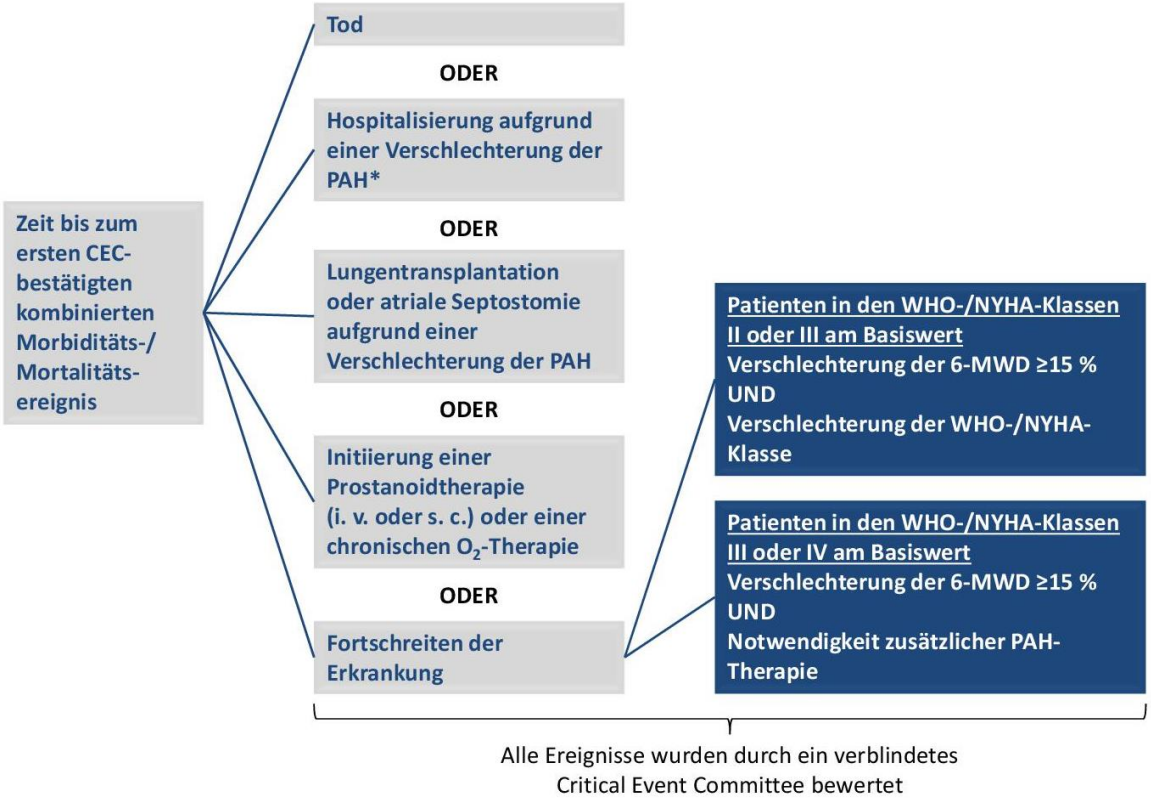


Abbildung 4-2: Aufbau des kombinierten Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ in GRIPHON
*Hospitalisierung aufgrund einer Verschlechterung der PAH: Jeglicher unfreiwillige Krankenhausaufenthalt (≥24 Stunden) aufgrund einer Verschlechterung der PAH; Verschlechterung der PAH schließt Anzeichen und Symptome des Rechtsherzversagens (z. B. (Prä-)Synkope, Zyanose, Verschlechterung der Kurzatmigkeit, klinisch relevante Verschlechterung der körperlichen Leistungsfähigkeit, Abfall der arteriellen Sauerstoffsättigung, verschlimmerte Ödeme, Hepatomegalie und Ascites) mit ein.

Der in GRIPHON verwendete Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ besitzt überdies eine grundlegende Ähnlichkeit zum

primären Endpunkt „Zeit bis zum Erreichen des ersten Morbiditäts- oder Mortalitätsereignisses“ in SERAPHIN, der Zulassungsstudie von Macitentan, einem ebenfalls zur Langzeitbehandlung von erwachsenen Patienten mit PAH zugelassenem Wirkstoff. Ein wesentlicher Unterschied zwischen den beiden Endpunkten besteht jedoch darin, dass im GRIPHON-Endpunkt – im Gegensatz zu SERAPHIN – zusätzlich die harte und zweifelsfrei patientenrelevante Komponente „Hospitalisierung aufgrund einer Verschlechterung der PAH“ als Ereignis eingeschlossen wurde. Der zusätzliche Einschluss dieser Komponente in die Analyse des primären Endpunkts wurde auch durch den G-BA im Beratungsgespräch zu Selexipag ausdrücklich begrüßt [3].

Mithilfe von Landmark-Analysen, die im Zusammenhang mit dem primären Endpunkt in GRIPHON durchgeführt wurden, konnte nachgewiesen werden, dass das Risiko, bis zum Ende der Studie zu versterben, für Patienten, bei denen vor Monat 6 ein CEC-bestätigtes Morbiditätsereignis beobachtet wurde, deutlich höher war als bei Patienten ohne solches Ereignis (HR [95 % KI]: 3,19 [2,00; 5,10], p<0,0001) [32]. Dies bestätigt die Relevanz des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ zur Langzeitprognose des klinischen Verlaufs der PAH-Patienten.

Primär ausgewertet wurde in GRIPHON die Zeit bis zum ersten Auftreten eines Morbiditäts- oder Mortalitätsereignisses im Zeitraum von der Randomisierung bis zu 7 Tage nach der letzten Einnahme der Studienmedikation. Weiterhin wurden verschiedene Sensitivitätsanalysen durchgeführt, um die Robustheit der Ergebnisse des primären Endpunkts zu überprüfen (vgl. auch Abschnitt 4.2.5.4). Hierzu gehörten insbesondere auch Analysen unter Ausschluss der Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer Prostanoidtherapie (i. v. oder s. c.) oder einer dauerhaften O_2-Therapie“, bei denen seitens des G-BA Zweifel an der Patientenrelevanz bestehen [3]. Unabhängig von diesem Vorgehen werden die genannten Komponenten jedoch auch weiterhin von Actelion als patientenrelevant angesehen.

Hospitalisierungen

Nach den allgemeinen Methoden des IQWiG wird als patientenrelevant verstanden, „wie ein Patient fühlt, seine Funktionen und Aktivitäten wahrnehmen kann oder ob er überlebt“ [15]. Entsprechend ist die Auswertung von Hospitalisierungen als patientenrelevanter Endpunkt im Sinne dieser Definition zu verstehen, da Patienten durch die Hospitalisierung in bedeutsamem Ausmaß an der Wahrnehmung ihrer Funktionen und Aktivitäten behindert werden. Die Patientenrelevanz der Hospitalisierungen aufgrund PAH-bedingter Komplikationen wurde darüber hinaus auch durch den G-BA im Beratungsgespräch zu Selexipag bestätigt [3].

In GRIPHON wurden Hospitalisierungen wegen PAH in insgesamt 3 Endpunkte ausgewertet. Zum einen wurden Hospitalisierungen als Bestandteil der kombinierten Endpunkte „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ und „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ analysiert, zum anderen erfolgte auch eine separate Auswertung der Hospitalisierungen wegen PAH als eigenständigem

Endpunkt. Auf die Darstellung der Hospitalisierungen jeglicher Ursache wurde verzichtet, da ausweislich der Niederschrift des Beratungsgespräches nur die Hospitalisierungen aufgrund PAH-bedingter Komplikationen vom G-BA als patientenrelevant angesehen werden [3].

Sämtliche Hospitalisierungen wegen PAH wurden vor dem Einschluss in die Auswertung durch das verblindete CEC dahin gehend begutachtet, ob tatsächlich von einem ursächlichen Zusammenhang zur PAH auszugehen ist.

Analysiert wurde jeweils die Zeit bis zur Hospitalisierung, wobei zu beachten ist, dass in den kombinierten Endpunkten das jeweils andere Ereignis bzw. die jeweils anderen Ereignisse des Endpunkts vor Eintreten des Todesfalles auftreten konnten. Die Auswertung erfolgte innerhalb des Zeitraums von der Randomisierung bis 7 Tage nach Einnahme der letzten Studienmedikation (EOT).

Bei der Ereigniszeitanalyse, die für den Endpunkt „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ dargestellt wurde, handelt es sich um eine nach Erstellung des Studienberichts durchgeführte Post-hoc-Analyse.

Symptome der PAH (nach Definition des G-BA)

Bereits im Beratungsgespräch zur Nutzenbewertung von Macitentan vom 02.12.2012 [14] hatte der G-BA ausgeführt, dass für ihn die symptomatische Krankheitsverschlechterung entscheidend sei. Unter Symptome der PAH zählt der G-BA Luftnot (Dyspnoe), körperliche Leistungseinschränkung bzw. -fähigkeit, Schwindel, Synkopen, Müdigkeit (Erschöpfung) und Ödeme. Dyspnoe und körperliche Leistungsfähigkeit wurden vom G-BA zudem auch im Beratungsgespräch zu Selexipag vom 13.05.2015 [3] als patientenrelevante Endpunkte in der Endpunktkategorie Morbidität genannt.

Mit Ausnahme der körperlichen Leistungsfähigkeit, die anhand der Veränderung der 6-MWD bestimmt wurde, waren die vom G-BA genannten Symptome der PAH nicht prospektiv als eigenständige Endpunkte in GRIPHON definiert. Eine Aussage darüber ist dennoch möglich, da die Symptome Dyspnoe, Schwindel, Synkopen, Erschöpfung und Ödeme als „bevorzugte Bezeichnungen“ (*preferred terms*, PT) innerhalb der unerwünschten Ereignisse erfasst wurden. Es wurde daher *post hoc* eine Analyse des Anteils der Patienten mit mindestens einem der vom G-BA definierten Symptome durchgeführt. Analysezeitraum war die Zeit von der Randomisierung bis zu 7 Tage nach der letzten Einnahme der Studienmedikation. Darüber hinaus wurde das Auftreten der einzelnen Symptome (jeweils für sich genommen) im Endpunkt „Unerwünschte Ereignisse“ in der Auflistung der einzelnen UE nach PT aufgeführt.

Veränderung der 6-Minuten-Gehstrecke (6-MWD)

Bei der Untersuchung der Gehstrecke, die ein Patient in sechs Minuten zurücklegen kann (6-Minuten-Gehstrecke, 6-MWD), bzw. der Veränderung über den Studienverlauf handelt es sich um den bisher am häufigsten in pivotalen PAH-Studien eingesetzten primären Endpunkt [28, 33]. Der 6-Minuten-Gehtest (6-MWT) ist ein einfach durchzuführender, günstiger und reproduzierbarer Test, der von den regulatorischen Behörden als Instrument zur Bestimmung der körperlichen Leistungsfähigkeit anerkannt ist [34].

In der Diskussion, ob sich die 6-MWD als patientenrelevanter Endpunkt oder als Surrogat für andere patientenrelevante Endpunkte in einer schwerwiegenden Erkrankung wie der PAH eignet, ist eine differenzierte Betrachtung der folgenden Messwerte erforderlich: des absoluten Werts zum Zeitpunkt der Basiswerterhebung, des absoluten Werts zu einem vorher festgelegten Zeitpunkt nach Beginn der Therapie und der Veränderung der 6-MWD vom Basiswert zu diesem vorher festgelegten Zeitpunkt (Δ6-MWD).

Wie die Ergebnisse verschiedener Register und Kohortenstudien nahelegen, korreliert der *absolute Wert* der 6-MWD zum Zeitpunkt der Basiswerterhebung mit kardialer Funktion und Überleben [35-39], wenngleich einige neuere Studien keine statistisch signifikante Assoziation nachweisen konnten [40, 41]. Ebenso scheint auch der *absolute Wert* der 6-MWD, der am Ende einer vorbestimmten Behandlungsdauer bestimmt wird, eine gewisse prognostische Relevanz zu besitzen [42], ein wünschenswerter Zielwert der 6-MWD ist jedoch weiterhin unklar [43].

Während also davon auszugehen ist, dass der 6-MWT-*Basiswert* – und mit Einschränkungen auch der absolute Wert am Ende einer vorbestimmten Behandlungsdauer – prognostischen Wert für die Mortalität und andere patientenrelevante Ereignisse besitzt, scheint die *Verbesserung* dieses Wertes durch eine medizinische Intervention nicht mit einer besseren Prognose der PAH zu korrelieren. Bestätigung hierfür findet sich in einer 2010 durchgeführten Meta-Analyse, die eine kumulative Verringerung der Mortalität um 39 % durch den Einsatz verschiedener Vasodilatatoren über einen mittleren Zeitraum der untersuchten klinischen Studien von 14 Wochen, aber keinen statistischen Zusammenhang dieser Verringerung mit der *Verbesserung* im 6-MWT feststellen konnte [44]. Ebenso konnten Fritz und Kollegen zwar erneut bestätigen, dass der *Basiswert* im 6-MWT mit der 2-Jahres-Überlebensrate korreliert, gleichzeitig konnten sie jedoch auch nachweisen, dass die erneute Bestimmung der 6-MWD *während der Behandlung* über einen Zeitraum von zwölf Wochen nur einen sehr geringen prognostischen Zusatzwert zur Vorhersage der 2-Jahres-Überlebensrate gegenüber dem 6-MWT-*Basiswert* besitzt [45]. Vielmehr zeigte die *Verbesserung* der Gehstrecke unter Therapie (Δ6-MWD) keine Korrelation mit der 2-Jahres-Überlebenswahrscheinlichkeit. Zum gleichen Schluss kommt auch eine Meta-Analyse von Gabler et al., in der die Ergebnisse aus zehn RCTs ausgewertet wurden. Hier zeigt sich, dass der Effekt der PAH-Behandlung auf klinisch relevante Endpunkte (Tod, Lungentransplantation, atriale Septostomie, Hospitalisierung aufgrund sich verschlechternder PAH, Abbruch wegen zunehmenden Rechtsherzversagens, Hinzufügen weiterer PAH-Medikation) nur zu 22 % durch die Δ6-MWD erklärt werden kann, die damit laut Aussage der Autoren einen „inadäquaten Surrogatendpunkt“[12] darstellt, der „möglicherweise nicht ausreichend ist, um alleine verwendet zu werden“[13] [46].

[12] Originalzitat: “However, Δ6-MWD remains an inadequate surrogate end point given the modest degree of mediation of the treatment effect.”
[13] Originalzitat: “Thus, the finding of true but modest mediation by 6-MWD suggests that it may not be sufficient to use on its own.”

Umgekehrt ermöglichte die *Verschlechterung* der 6-MWD jedoch die Vorhersage der Mortalität bei Patienten mit PAH; eine durch wiederholtes Testen an unterschiedlichen Tagen bestätigte *Verschlechterung* im 6-MWT von 15 % stellt einen Verlust der körperlichen Leistungsfähigkeit dar, von dem gezeigt werden konnte, dass er indikativ für das Fortschreiten der PAH ist [47]. Ebenso konnte 2015 durch Farber und Kollegen in einer Studie gezeigt werden, dass die *Verschlechterung* der 6-MWD (in 15 %-Inkrementen) von Patienten, die in das US-amerikanische REVEAL-Register eingeschlossen worden waren, stark mit deren 1-Jahres-Überlebensrate korreliert. Schließlich hat auch die Dana-Point-*Task Force on End Points and Clinical Trial Design* die *Verschlechterung* der 6-MWD um 15 % als geeignetes Instrument zur Bestimmung der klinischen Verschlechterung der PAH anerkannt. Dies setzt jetzt jedoch voraus, dass die *Verschlechterung* der 6-MWD – wie in GRIPHON im kombinierten Endpunkt „Zeit bis ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ geschehen – mit anderen klinischen Erhebungen, wie z. B. der Verschlechterung der WHO-/NYHA-Klasse oder der Notwendigkeit zusätzlicher PAH-Therapie, kombiniert wird [28].

Wenngleich es sich bei der *Verbesserung* der 6-MWD nicht um ein Surrogat für den patientenrelevanten Endpunkt Mortalität handelt, so ist sie doch als ein Maß für die aktuelle körperliche Leistungsfähigkeit der Patienten von Bedeutung und steht als solche in Verbindung mit den täglichen Aktivitäten der Patienten. Entsprechend wurde die 6-MWD in den Nutzenbewertungen von Macitentan und Riociguat als patientenrelevanter Endpunkt anerkannt [48, 49] und wird daher auch in der Nutzenbewertung von Selexipag zur Bestimmung des Ausmaßes des medizinischen Zusatznutzens in der Endpunktkategorie Morbidität angeführt.

In GRIPHON wurde die absolute Veränderung der 6-MWD in Woche 26 in Vergleich zum Basiswert als sekundärer Endpunkt erhoben. Weitere Zeitpunkte außer Woche 26 (Woche 8 bis Monat 36; *End-Of-Study* (EOS)-Untersuchung) wurden als explorative Endpunkte erhoben und sind ebenfalls im Dossier dargestellt. Die Durchführung des 6-MWT erfolgte anhand eines standardisierten Actelion-Leitfadens (dargestellt in Appendix 1 des Studienprotokolls von GRIPHON), der wiederum eine Modifikation des entsprechenden Leitfadens der American Thoracic Society (ATS) ist. Nach Beginn der Behandlung mit der Studienmedikation wurden sämtliche Erhebungen der 6-MWD zum Zeitpunkt des Talspiegels der Wirkstoffkonzentration, d. h. 12 Stunden nach der letzten Einnahme der Studienmedikation erhoben.

Sehr wichtig in der Interpretation der Ergebnisse zur Δ6-MWD – insbesondere auch im Vergleich zu den Ergebnissen aus Studien mit anderen PAH-spezifischen Arzneimitteln – ist die Kenntnis über die in den verschiedenen Studien gewählten Ersetzungsstrategien sowie den Zeitpunkt der Messung relativ zum Plasmaspiegel des jeweiligen Wirkstoffes (vgl. Tabelle 4-5). Hierbei zeigt sich, dass sowohl die Ersetzungsstrategie als auch der Zeitpunkt der Messung in GRIPHON äußerst konservativ gewählt wurden und man somit den Behandlungseffekt im Vergleich zu den meisten anderen Studien tendenziell unterschätzt.

Tabelle 4-5: 6-MWD - Zeitpunkt der Messung und Ersetzungsstrategien in verschiedenen PAH-Studien

	Selexipag (GRIPHON)	**Macitentan (SERAPHIN)**	**Riociguat (PATENT-1)**	**Tadalafil (PHIRST)**	**Sildenafil (SUPER-1)**
Ersetzungsstrategie bei …					
Tod	0 m	0 m	0 m	Schlechtester beobachteter Rang	LOCF
Abbruch wegen Verschlechterung der PAH	0 m	Unteres Quartil	0 m	Schlechtester beobachteter Rang	LOCF
Abbruch wegen UE	zweit-schlechtester Wert (10 m)	LOCF	LOCF	Basiswert	LOCF
Erhebungszeitpunkt relativ zum Plasmaspiegel					
Zeitpunkt	Talspiegel	Talspiegel	Spitzenspiegel	Spitzenspiegel	Spitzenspiegel

Quellen: Studienbericht zu GRIPHON, Publikationen zu SERAPHIN [50], PATENT-1 [51], PHIRST [52], SUPER-1 [53]

UE – unerwünschtes Ereignis, LOCF – *last observation carried forward* (Fortschreibung des letzten bekannten Wertes)

Verschlechterung/ausbleibende Verschlechterung der WHO-/NYHA-Klasse

Das Standardsystem zur Klassifizierung der Schwere der PAH-Erkrankung ist die in Tabelle 4-6 dargestellte Einteilung in funktionale Klassen nach der World Health Organization (WHO) bzw. der New York Heart Association (NYHA) [54, 55]. Die Einteilung in WHO-/NYHA-Klassen wird als wichtiges Instrument zur Erfassung des klinischen Status und der klinischen Verbesserung der PAH-Patienten angesehen. In klinischen Studien, Beobachtungsstudien und Registeranalysen konnte wiederholt eine Korrelation zwischen der WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung und der Mortalität beobachtet werden [36, 42, 56-60]. Zudem ist eine Verbesserung aus höheren Klassen in die WHO-/NYHA-Klassen I und II, d.h. aus Klassen mit schwerwiegender Symptomatik in Klassen mit vergleichsweise weniger schwerwiegender Symptomatik, prädiktiv für verbessertes Überleben [23], während die Verschlechterung der WHO-/NYHA-Klasse mit einer erhöhten Mortalität korreliert [39].

Die gegenwärtigen Behandlungsleitlinien der PAH definieren die Verbesserung bzw. Stabilisierung der WHO-/NYHA-Klasse als eines der vorrangigen Ziele der PAH-Behandlung [17, 61, 62]; gleichzeitig sehen die regulatorischen Behörden in der Langzeitverbesserung der WHO-/NYHA-Klasse einen klinisch signifikanten Endpunkt [16]. Die funktionelle Klassifizierung nach WHO/NYHA ermöglicht eine direkte Aussage über den klinischen Status und insbesondere auch die Symptome des Patienten [16] und steht somit in unmittelbarem Zusammenhang mit seinen täglichen Aktivitäten. Es handelt sich beim Endpunkt „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ daher um einen patientenrelevanten Endpunkt in der Endpunktkategorie Morbidität.

Tabelle 4-6: Funktionsklassen der pulmonalen Hypertonie nach WHO/NYHA

Funktionsklasse	Definition
I	Patienten mit pulmonaler Hypertonie, aber ohne daraus resultierende Einschränkungen der körperlichen Aktivitäten. Normale körperliche Aktivitäten verursachen keine übermäßige Dyspnoe, Erschöpfung, Brustschmerzen oder Synkopen.
II	Patienten mit pulmonaler Hypertonie mit daraus resultierenden leichten Einschränkungen der körperlichen Aktivitäten. Im Ruhezustand liegen keine Probleme vor. Normale körperliche Aktivitäten verursachen übermäßige Dyspnoe, Erschöpfung, Brustschmerzen oder Synkopen.
III	Patienten mit pulmonaler Hypertonie mit daraus resultierenden deutlichen Einschränkungen der körperlichen Aktivitäten. Im Ruhezustand liegen keine Probleme vor. Weniger als normale körperliche Aktivitäten verursachen übermäßige Dyspnoe, Erschöpfung, Brustschmerzen oder Synkopen.
IV	Patienten mit pulmonaler Hypertonie mit Unvermögen zur Ausführung jeglicher körperlicher Aktivitäten, ohne dass sich Symptome einstellen. Es zeigen sich Anzeichen des Rechts-Herz-Versagens. Dyspnoe und Erschöpfung können sich bereits in Ruhe einstellen. Die Beschwerden werden durch körperliche Aktivität weiter gesteigert.
Quelle: Empfehlungen zur Behandlung der PAH, Nizza 2013 [17]; Kommentierung derselben durch das PH-DACH-Symposium 2014 [61]	

Eine Verschlechterung der WHO-/NYHA-Klasse war in GRIPHON definiert als Verschlechterung im Vergleich zur Beurteilung zum Zeitpunkt der Basiswerterhebung. Eine ausbleibende Verschlechterung war dementsprechend als gleiche oder bessere WHO-/NYHA-Klasse zu verstehen. Patienten, die zum Beginn der Studie in WHO-/NYHA-Klasse IV waren, waren von der Auswertung ausgeschlossen, da bei ihnen per Definition keine weitere Verschlechterung mehr auftreten konnte.

In GRIPHON wurde die ausbleibende Verschlechterung der WHO-/NYHA-Klasse als sekundärer Endpunkt vom Zeitpunkt der Basiswerterhebung bis zu Woche 26 erhoben. Neben Woche 26 wurden auch weitere Zeitpunkte (Woche 4 bis Monat 36; EOS-Untersuchung) als explorative Endpunkte ausgewertet. Diese sind als zusätzliche Informationen ebenfalls im Dossier dargestellt.

Veränderung des Borg-Dyspnoe-Index

Im Kontext der PAH stellt die Dyspnoe des Patienten eines der zentralen Symptome dar. Zur Erfassung dieses Symptoms wurde daher während des 4. PH-Weltsymposiums der Einsatz des Borg-Dyspnoe-Index vorgeschlagen [28]. Der Borg-Dyspnoe-Index besteht aus einer Frage (vereinfacht: „Wie stark ist Ihre Atemfähigkeit derzeit eingeschränkt?“), die anhand einer zwölfteiligen Skala mit Ausprägungen von 0 (keine Dyspnoe) bis 10 (schwerste bisher erlebte Dyspnoe) beantwortet wird. Die Erfassung des Borg-Dyspnoe-Index kann entweder im Patient-Arzt-Interview stattfinden oder per Fragebogen durch den Patienten selbst vorgenommen werden.

In GRIPHON wurde der Borg-Dyspnoe-Index in einem Patient-Arzt-Interview nach einem von Actelion vorgegebenen Protokoll erhoben. Die Bestimmung erfolgte im unmittelbaren Anschluss an die Durchführung des 6-MWT und damit zum Talspiegel der Wirkstoffkonzentration. Aufgrund der Verknüpfung mit dem 6-MWT wurde der Borg-Dyspnoe-Index zudem zu den gleichen Zeitpunkten (Woche 8 bis Monat 36, EOS-Untersuchung) wie die 6-MWD erhoben.

Wie der G-BA in der Nutzenbewertung von Macitentan anmerkt, liegt eine Validierung des Borg-Dyspnoe-Index als Surrogat für einen patientenrelevanten Endpunkt bei Patienten mit PAH bisher nicht vor [63]. Die Ergebnisse dieses Endpunktes werden in der Nutzenbewertung von Selexipag daher lediglich ergänzend dargestellt.

Veränderung der Lebensqualität

Neben der Verlängerung des Überlebens und der Verbesserung der Symptome ist gemäß den Empfehlungen des 5. PH-Weltsymposiums ein weiteres Hauptziel der PAH-Behandlung die Verbesserung der gesundheitsbezogenen Lebensqualität [28]. Diese Forderung wird unter anderem auch durch die EMA gestützt, die in ihrer Leitlinie zur klinischen Untersuchung von Arzneimitteln in der PAH die gesundheitsbezogene Lebensqualität als wichtigen sekundären Endpunkt ansieht, da PAH-Arzneimittel aufgrund ihrer Darreichungsform (z. B. als belastende kontinuierliche intravenöse Infusion) oder wegen assoziierter Nebenwirkungen negative Auswirkungen auf die Lebensqualität der Patienten haben könnten [16].

Unter den Instrumenten, die bisher zur Bestimmung der Lebensqualität in PAH-Studien eingesetzt wurden, finden sich sowohl generische Fragebögen wie der Medical Outcomes Study Short Form-36 (SF-36) als auch krankheitsspezifische (aber nicht PAH-spezifische) Fragebögen wie der Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) [28, 64].

Ein grundsätzliches Bedenken gegenüber dem Einsatz generischer Fragebögen, die nicht spezifisch für den Einsatz bei Patienten mit PAH validiert wurden, ist, dass diese möglicherweise für die Patienten wichtige Probleme/Symptome nicht erfassen. Veränderungen im Gesundheitszustand führen daher nicht unbedingt zu einer Veränderung der Lebensqualität, gemessen anhand dieser Fragebögen [65]. Bis zum jetzigen Zeitpunkt liegt jedoch kein abschließend validiertes, spezifisch für die PAH entwickeltes Instrument zur Erhebung der gesundheitsbezogenen Lebensqualität vor. Daher wurden in vergangenen Studien häufig Instrumente eingesetzt, die spezifisch für Herzerkrankungen oder respiratorische Erkrankungen sind, in manchen Fällen mit geringfügigen Adaptionen für pulmonalen Bluthochdruck (z. B. die MHLF-PH-Adaption) [66, 67].

Der Fragebogen, der zum Zeitpunkt des Designs der GRIPHON-Studie einem PAH-spezifischen Instrument zur Erhebung der PAH möglicherweise am nächsten kommt, ist der bereits erwähnte CAMPHOR-Fragebogen. Beim CAMPHOR handelt es sich laut G-BA um „*das [derzeit] einzige etablierte krankheitsspezifische Instrument zur Erfassung der Lebensqualität bei Patienten mit PH*“ [63]. Erstmalig veröffentlicht wurde der Fragebogen 2006 [68] und 2008 zunächst für die USA validiert [69]. Derzeit liegt der Fragebogen für

insgesamt 15 Länder bzw. Sprachgruppen vor, darunter auch für Deutschland [70, 71]. Im Zusammenhang mit dem CAMPHOR-Fragebogen ist zu beachten, dass die ausgewertete Validierungspopulation nicht ausschließlich auf Patienten mit PAH beschränkt war, sondern dass auch PH-Patienten miteingeschlossen wurden.

Der CAMPHOR-Fragebogen besteht aus insgesamt 3 unterschiedlichen Skalen („Symptome“, „Aktivitäten“ und „Lebensqualität“) mit insgesamt 65 Items. Die Studienplanung von GRIPHON sah vor, dass von den insgesamt verfügbaren Items die „Symptome“-Skala und die dazugehörige „Kurzatmigkeit“-Subskala ausgewertet werden sollten. Die „Symptome“-Skala besteht aus 25 negativ gewichteten Items des CAMPHOR-Fragebogens, die sich auf die 3 Subskalen „Energie“, „Kurzatmigkeit“ und „Stimmung“ aufteilen. Alle Items der „Symptome“-Skala sind so gestaltet, dass die Patienten auf einfache Weise angeben können, ob die jeweilige Aussage in der derzeitigen Situation auf sie zutrifft oder nicht. Eine „Ja“-Antwort wird hierbei mit 1, eine „Nein“-Antwort mit 0 gewertet. Die Item-Scores werden anschließend zum Gesamt-Score addiert, der für die „Symptome“-Skala des Fragebogens von 0 (schlecht) bis 25 (gut) und für die „Kurzatmigkeit“-Subskala von 0 bis 8 reichen kann.

Da der CAMPHOR-Fragebogen wie beschrieben nur zur Verwendung in einer begrenzten Anzahl von Ländern verfügbar ist, musste die Auswertung auf das „*quality of life analysis set*“ (QAS) beschränkt werden. Die QAS-Population beinhaltet daher nur Patienten aus den folgenden Ländern: Australien, Vereinigtes Königreich und USA (Englisch), Österreich und Deutschland (Deutsch), Kanada (Französisch und Englisch), Schweiz (Französisch und Deutsch), Spanien (Spanisch), Schweden (Schwedisch), Frankreich (Französisch), Ungarn (Ungarisch), Israel (Hebräisch), Italien (Italienisch) und Polen (Polnisch).

In GRIPHON wurde die absolute Veränderung im „Symptome“-Score und in der „Kurzatmigkeit“-Subskala des CAMPHOR-Fragebogens als sekundäre Endpunkte vom Zeitpunkt der Basiswerterhebung bis zu Woche 26 erhoben. Neben Woche 26 wurden auch weitere Zeitpunkte (Woche 16 bis Monat 24; EOS-Untersuchung) als explorative Endpunkte ausgewertet. Diese sind als zusätzliche Information ebenfalls im Dossier dargestellt.

Unerwünschte Ereignisse

Bei der Bewertung einer medikamentösen Behandlung im Rahmen einer klinischen Studie ist die Erfassung unerwünschter Ereignisse von größter Bedeutung für die Sicherheitsbewertung des Arzneimittels. Daher werden in diesem Nutzendossier alle unerwünschten Ereignisse angegeben. Schwerwiegende unerwünschte Ereignisse im Sinne dieses Nutzendossiers waren wie in der üblichen Definition Ereignisse, die zum Tod führten, lebensbedrohlich waren, zu Einweisung oder verlängertem Aufenthalt im Krankenhaus führten, Behinderungen nach sich zogen oder Geburtsfehler bzw. angeborene Anomalie sind. Zusätzlich wurden Therapieabbrüche aufgrund von unerwünschten Ereignissen zur Bewertung der Verträglichkeit der Studienmedikation angegeben.

Ausgewertet wurde der Anteil der Patienten, bei denen in der Zeit von der Randomisierung bis zu 7 Tage nach Ende der Behandlung mindestens ein unerwünschtes Ereignis beobachtet wurde. Zusätzlich erfolgt eine Darstellung der UE in Abhängigkeit von der Behandlungsphase

(Titration oder Erhaltung), in die jedoch nur diejenigen Patienten eingeschlossen wurden, die an beiden Phasen teilgenommen hatten.

Da bestimmte Ereignisse des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ zeitgleich auch als UE gewertet werden konnten, ist eine Doppelbewertung dieser Ereignisse in zwei Endpunktkategorien möglich. Zur Überprüfung dieses Effekts wurden Sensitivitätsanalysen der UE durchgeführt, indem die Ereignisse des primären Endpunkts aus der Analyse der UE ausgeschlossen wurden. Es wurden hierbei solche UE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen, da ein Ausschluss dieser Ereignisse auf Ebene der bevorzugten Bezeichnungen nicht möglich ist.

Unerwünschte Ereignisse von besonderem Interesse

Ausgehend von präklinischen Erfahrungen oder vorangegangenen Erfahrungen mit anderen Wirkstoffen, die auf den Prostacyclin-Signalweg wirken, wurden prospektiv zur näheren Evaluation möglicher Sicherheitsrisiken für GRIPHON verschiedene unerwünschte Ereignisse von besonderem Interesse definiert und ausgewertet. Die Auswertung der unerwünschten Ereignisse von besonderem Interesse erfolgte analog zur Auswertung der unerwünschten Ereignisse (d. h. primäre Analysemethode war die Ereigniszeitanalyse).

Hämorrhagie

Vor Beginn der Studie war bereits bekannt, dass Prostacyclin-Rezeptor-Agonisten die Thrombozytenaggregation hemmen können [72]. Gleichzeitig wurde in In-vitro-Studien beobachtet, dass Selexipag und sein aktiver Metabolit einen inhibitorischen Effekt auf die menschliche Thrombozytenaggregation haben [72]. Andererseits zeigte sich in einer Phase-I-Studie (AC-065-101) mit gesunden Probanden kein wesentlicher Effekt von Selexipag auf die Thrombozytenaggregation [73]. In GRIPHON sollte dennoch untersucht werden, ob die Anwendung von Selexipag unter physiologischen Bedingungen zu einer erhöhten Anzahl von Blutungsereignissen führt. Es wurde jedoch kein maßgeblicher Unterschied in der Inzidenz von Blutungsereignissen während der Studie beobachtet.

Die Auswertung der Blutungsereignisse erfolgte anhand der *Standardized MedDRA Queries* (SMQ) „Hämorrhagie“ (ausschließlich von Laborwerten) und „Gastrointestinale Hämorrhagie“. Ausreichend war jeweils mindestens ein UE, das als PT in einer der beiden SMQ gelistet war.

Hypotonie

Hypotonie ist ein möglicher Klasseneffekt der Prostacyclin-Rezeptor-Agonisten und wurde daher ebenfalls als unerwünschtes Ereignis von besonderem Interesse in GRIPHON ausgewertet [72]. Die Analyse der Hypotonie-Ereignisse lassen jedoch höchstens auf einen geringen Effekt durch Selexipag auf den systemischen Blutdruck schließen. Die Tatsache, dass im Selexipag-Arm zudem weniger Synkopen aufgetreten waren (sowohl als UE wie auch als SUE), bestätigt diese Beobachtung.

Die Auswertung der Hypotonie-Ereignisse erfolgte anhand einer vorab definierten Liste von UE mit Hypotonie-assoziierten PT (siehe Tabelle 4-107). Ausreichend war jeweils mindestens ein UE der vorab definierten Liste.

Störung des Auges

Störungen des Auges wurden aufgrund von Beobachtungen in präklinischen Studien als unerwünschtes Ereignis von besonderem Interesse aufgenommen, in denen am Ende einer zweijährigen Karzinogenität-Studie[14] Tortuosität und Dilatation retinaler Blutgefäße bei Ratten beobachtet worden waren [72]. Auch wenn die Übertragbarkeit dieser Beobachtung auf den Menschen als unwahrscheinlich angesehen wurde, wurden Störungen des Auges dennoch als besonderer Sicherheitsschwerpunkt in GRIPHON untersucht. Die Ergebnisse von GRIPHON können indes die ursprüngliche Befürchtung nicht bestätigen.

Die Auswertung der Störungen des Auges erfolgte anhand der System-Organ-Klasse (*system organ class*, SOC) „Augenerkrankungen“ und der SMQ „Retinale Störungen“. Ausreichend war jeweils mindestens ein UE, das als PT in der SOC oder der SMQ gelistet war.

Prostacyclin-assoziierte unerwünschte Ereignisse

Da Selexipag ebenso wie die Prostanoide zu einer Aktivierung des Prostacyclin-Signalweges führt, sollte in GRIPHON zusätzlich untersucht werden, inwiefern die Anwendung von Selexipag mit dem Auftreten unerwünschter Ereignisse verbunden ist, die üblicherweise mit der Aktivierung dieses Signalweges assoziiert werden. Zu diesen Prostacyclin-assoziierten UE gehören insbesondere Kieferschmerzen, Flush, Übelkeit und Erbrechen, Durchfall, Kopfschmerz und Schwindel. Aufgrund des Titrationsschemas von Selexipag, das die Titration des Wirkstoffs bis zu der individuellen Dosis vorsah, an dem Nebenwirkungen der Behandlung nicht mehr durch adäquate Behandlung handhabbar waren, war davon auszugehen, dass die Wahrscheinlichkeit Prostacyclin-assoziierter UE im Verlauf der Titrationsphase höher ist. Die Prostacyclin-assoziierten UE wurden daher getrennt für die Titrationsphase und die Erhaltungsphase ausgewertet.

Die Auswertung der Prostacyclin-assoziierten UE erfolgte anhand einer vorgegebenen Liste von UE mit Prostacyclin-assoziierten PT (vgl. Tabelle 4-112). Ausreichend war jeweils mindestens ein UE der vorab definierten Liste.

[14] Die durchschnittliche Lebenserwartung einer Ratte beträgt zwei Jahre, die beobachteten Ereignisse traten somit in einem für Ratten sehr hohen Lebensalter auf.

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten[15] erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet[16] werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität[17] erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

---

[15] *DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-188.*
[16] *Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2008. S. 243-296.*
[17] *Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.*

Für keine der beiden untersuchten Teilpopulationen liegt mehr als eine Studie für die Nutzenbewertung vor. Die Notwendigkeit einer Meta-Analyse entfällt daher.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Teilpopulation a

Sensitivitätsanalysen als sekundäre Analysen zu patientenrelevanten Endpunkten

In GRIPHON wurden zur Untersuchung der Robustheit des primären Endpunkts folgende Sensitivitätsanalysen durchgeführt:

- Analyse unter Ausschluss der Endpunkt-Ereignisse vor dem 16. August 2011

  Um auszuschließen, dass es sich bei den Änderungen im 4. globalen Protokoll-Amendment um informierte Änderungen handelt, wurden sämtliche M/M-Ereignisse, die vor dem 16. August 2011 (Datum des Amendments) auftraten, aus der Analyse ausgeschlossen. Im Studienbericht war diese Sensitivitätsanalyse als primäre Analyse dargestellt. Aufgrund der hohen Übereinstimmung mit der Analyse ohne Ausschluss der Ereignisse wurde anschließend jedoch entschieden, sämtliche weiteren Sensitivitätsanalysen und Subgruppenanalysen unter Einschluss aller Ereignisse, also der Daten aller teilnehmenden Patienten durchzuführen.

- *Per-protocol-Analyse*

- Analysen zur Untersuchung möglicher informativer Zensierung

- Analyse des Endpunktes bei den ersten 670 randomisierten Patienten und bis zum Auftreten von 202 CEC-bestätigten Ereignissen (Zeitpunkt der Interimsanalyse)

  Die Analyse entspricht der ursprünglichen Studienplanung vor dem 5. globalen Protokoll-Amendment.

- Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden[18]
- Analyse unter Einschluss der vom Prüfarzt gemeldeten Ereignisse, unabhängig von ihrer Bestätigung durch das CEC
- Analyse unter Ausschluss von Ereignissen der Komponente „Fortschreiten der Erkrankung“, wenn kein oder nur ein 6-MWT durchgeführt wurde
- Analysen unter Ausschluss der Komponenten „Fortschreiten der Erkrankung“ und/oder „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“

  Im Beratungsgespräch zu Selexipag hatte der G-BA angegeben, dass er – ähnlich wie bereits in der Nutzenbewertung von Macitentan – die Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“ als von unterschiedlicher Schwere im Vergleich zu den übrigen Komponenten des Endpunkts ansehe [3]. Zudem sei aus seiner Sicht die Relevanz der „Fortschreiten der Erkrankung“ kritisch zu sehen. Actelion hatte infolgedessen die genannten Sensitivitätsanalysen vorgeschlagen, ihre Darstellung im Nutzendossier wurde vom G-BA empfohlen [3].

Ebenso wie für den primären Endpunkt wurden auch für verschiedene sekundäre Endpunkte Sensitivitätsanalysen zur Untersuchung der Robustheit des Ergebnisses durchgeführt:

- Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH:
  - *Per-protocol-Analyse*
  - Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden
- Zeit bis zum Tod (Studienschluss):
  - Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden
- Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT):
  - Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden
- Absolute Veränderung der 6-MWD:

[18] Die Sicherheit der PAH-Diagnose wurde durch zwei verblindete PAH-Experten sichergestellt. Patienten mit „typischer PAH“ hatten keine Komorbidität, die die Ätiologie der PAH infrage stellten, bei Patienten mit „atypischer PAH“ war eine solche Komorbidität vorhanden. Patienten mit „konsistenter PAH“ waren Patienten, die weder in die typische noch die atypische Kategorie fielen.

  - Analyse unter Ausschluss der Patienten mit nicht am Talspiegel durchgeführtem 6-MWT
  - *Per-protocol-Analyse*
  - Analysen mit alternativen Imputationsalgorithmen
  - Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden
- Ausbleibende Verschlechterung der WHO-/NYHA-Klasse:
  - Analyse ohne Ersetzung fehlender Werte
  - *Per-protocol-Analyse*

Einzelheiten zu den verschiedenen Sensitivitätsanalysen sind in Abschnitt 4.3.1.3.1 innerhalb der Beschreibung der entsprechenden Endpunkte aufgeführt.

Sensitivitätsanalysen im Rahmen von Meta-Analysen und indirekten Vergleichen

Die Nutzenbewertung von Selexipag erfolgt anhand der RCT GRIPHON. Meta-Analysen oder indirekte Vergleiche wurden nicht durchgeführt. Daher entfällt auch die Notwendigkeit zu entsprechenden Sensitivitätsanalysen.

Teilpopulation b

Es konnte keine relevante RCT für die Nutzenbewertung von Selexipag in dieser Teilpopulation identifiziert werden. Die Notwendigkeit von Sensitivitätsanalysen entfällt daher.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

– Geschlecht
– Alter

– Krankheitsschwere bzw. –stadium
– Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Durchführung von Subgruppenanalysen dient der Identifikation möglicher Effekte, die zur Modifikation der Ergebnisse in Teilpopulationen führen könnten; es soll hiermit insbesondere die Fragestellung geklärt werden, ob für eine Teilpopulation ein anders zu bewertender Zusatznutzen vorliegt [15]. Überprüft wird die Ähnlichkeit der Subgruppen durch den Vergleich der jeweiligen Effekte in den Subgruppen und mithilfe eines Interaktionstests. Hierbei wird ein zum Niveau $\alpha=0{,}05$ signifikantes Ergebnis als Beleg für einen Unterschied zwischen den Subgruppen gewertet.

Bei der Testung einer größeren Anzahl statistischer Hypothesen ist weiterhin zu beachten, dass eine gewisse Anzahl der durchgeführten Tests aufgrund der angenommenen Irrtumswahrscheinlichkeit, also per Definition, zu einem falsch positiven Ergebnis führen. So ist zum Beispiel bei den für dieses Dossier durchgeführten 140 Testungen (20 Forest Plots à 7 Subgruppen) rein statistisch mit 7 falsch positiven Belegen ($\alpha<0{,}05$) für eine Effektmodifikation zu rechnen. Ein positiver Interaktionstest ($p<0{,}05$) wird in diesem Dossier daher lediglich als Ausgangspunkt für die Überprüfung einer möglichen Effektmodifikation herangezogen. Liegt ein positiver Interaktionstest vor ($p<0{,}05$), werden daher zusätzlich die folgenden Kriterien geprüft, die aus dem Entwurf der EMA-Leitlinien zur Untersuchung von Subgruppen in konfirmatorischen klinischen Studien sowie aus einem Gutachten des medizinischen Dienstes der Spitzenverbände der Krankenkassen zu Subgruppenanalysen abgeleitet wurden [74, 75]:

- Besteht eine medizinisch, biologisch oder physiologisch plausible Erklärung für die mögliche Effektmodifikation?
- Liegen konsistente Ergebnisse zur möglichen Effektmodifikation über mehrere Endpunkte vor?

- Liegen zur möglichen Effektmodifikation bereits ähnliche Ergebnisse aus anderen Studien vor?

Teilpopulation a

Gemäß den Anforderungen des Nutzendossiers sind zur Bestimmung möglicher Effektmodifikatoren des medizinischen Zusatznutzens *a priori*, d. h. im Studienprotokoll festgelegte Subgruppenanalysen immer zu berichten. Im Studienprotokoll von GRIPHON sind lediglich für den primären Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ Subgruppen definiert. Für das Nutzendossier werden jedoch für alle patientenrelevanten Endpunkte Subgruppenanalysen vorgelegt. Hierbei werden die folgenden Subgruppen analysiert:

- PAH-spezifische Therapie zum Zeitpunkt der Basiswerterhebung (ERA-Monotherapie, PDE-5-I-Monotherapie, ERA- und PDE-5-I-Kombinationstherapie, keine PAH-spezifische Therapie)
- WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I/II, III/IV)
- Geschlecht (männlich, weiblich)
- Rasse (kaukasisch/hispanisch, asiatisch, schwarz)
- Alter zum Zeitpunkt des Screenings (<65 Jahre, ≥65 Jahre)[19]
- PAH-Ätiologie zum Zeitpunkt der Basiswerterhebung (IPAH/HPAH/HIV-assoziiert/Drogen- oder Toxin-induziert, assoziiert mit Bindegewebserkrankungen, assoziiert mit angeborener Herzschwäche)
- Geographische Region (Nordamerika, Westeuropa/Australien, Osteuropa, Asien, Lateinamerika)[20]

Die in der Nutzendossiervorlage geforderte Subgruppenanalyse nach Krankheitsschwere bzw. -stadium entspricht im Anwendungsgebiet der Einteilung der Patienten nach WHO-/NYHA-Klasse. Die Entscheidung zur Unterteilung der Studienpopulation in zwei Subgruppen, die auf der einen Seite Patienten der WHO-/NYHA-Klassen I/II und auf der anderen Seite Patienten der WHO-/NYHA-Klassen III/IV zusammenfassen, folgt vor allem den Therapiezielen der aktuellen Leitlinien. Als eines der Ziele wird hierbei angesehen, dass Patienten die WHO-/NYHA-Klassen I oder II erreichen und in diesen stabilisiert werden sollen, während für Patienten in den WHO-/NYHA-Klassen III und IV eine Anpassung der Therapie empfohlen wird [12, 17]. Auch ist der mögliche Therapieerfolg in den Klassen I und II möglicherweise durch den Ceiling-Effekt beschränkt, was in den Klassen III und IV nicht der Fall ist. Da sich überdies nur ein geringer Anteil der Patienten in GRIPHON in den

[19] Für die Darstellung der demographischen Werte wurde zusätzlich noch die Gruppe der Patienten, die 75 Jahre oder älter waren, ausgewertet.

[20] Da der CAMPHOR-Fragebogen nur für eine begrenzte Anzahl von Ländern verfügbar ist, musste die Auswertung der entsprechenden Endpunkte auf die QAS-Population eingeschränkt werden. Diese Population beinhaltet jedoch keine Patienten in Asien oder Lateinamerika, sodass auch keine CAMPHOR-Subgruppenergebnisse für diese Regionen vorliegen.

WHO-/NYHA-Klassen I und IV befand, erscheint eine individuelle Auswertung dieser Schweregrade nur wenig sinnvoll.

Für die Subgruppe Alter wurden die Trennpunkte so gewählt, dass sie die Altersgruppen Erwachsene (18-64 Jahre) und Senioren (≥65 Jahre) abdecken.

Die Subgruppenanalyse wurde in den folgenden zwei Schritten durchgeführt. Im ersten Schritt wurde jede einzelne Subgruppe mit der gleichen statistischen Methode analysiert wie die statistische Analyse der vollständigen FAS des gleichen Endpunkts. Dies waren z. B.:

- Log-Rang-Test und Proportional-Hazard-Modell für Ereigniszeitendpunkte
- Cochran-Mantel-Haenszel-Test oder Barnards exakter Test für binäre Endpunkte
- (nicht-parametrische) Kovarianzanalyse (*analysis of covariance*, ANCOVA) für kontinuierliche Endpunkte

Im zweiten Schritt wurde ein modifiziertes statistisches Modell angewandt, das die Intergration von „Behandlung × Subgruppe“-Interaktionstermen ermöglicht. Dies waren z. B.:

- Proportional-Hazard-Modell mit Studienbehandlung, Subgruppe und „Behandlung × Subgruppe“-Interaktionsterm für Ereigniszeitendpunkte
- Logistisches-Regressions-Modell mit Studienbehandlung, Subgruppe und „Behandlung × Subgruppe“-Interaktionsterm für binäre Endpunkte
- (nicht-parametrische) ANCOVA mit Studienbehandlung, Subgruppe und „Behandlung × Subgruppe“-Interaktionsterm für kontinuierliche Endpunkte

Die Darstellung der Ergebnisse erfolgt in Form von Forest-Plots und unter Angabe des Punktschätzers sowie des dazugehörigen 95 %-Konfidenzintervalls je Subgruppe. Signifikanztests zum Behandlungsunterschied zwischen den Therapiearmen wurden in den Subgruppen nicht durchgeführt, jedoch lässt sich die Signifikanz des Ergebnisses indirekt aus der Lage des Konfidenzintervalls ableiten. Von einem signifikanten Ergebnis wird hierbei dann ausgegangen, wenn das Konfidenzintervall den Nulleffekt nicht miteinschließt.

Sofern sich mehrere Zeitpunkte für die Analyse eines Endpunkts anbieten, wurde derjenige Zeitpunkt für die Subgruppenanalysen angewendet, der auch bereits für die jeweilige Hauptanalyse herangezogen worden war. Beispielsweise erfolgte die Subgruppenanalyse für die absolute Veränderung der 6-MWD als Analyse der Veränderung in Woche 26 im Vergleich zum Basiswert.

Teilpopulation b

Es konnte keine relevante RCT für die Nutzenbewertung von Selexipag in dieser Teilpopulation identifiziert werden. Die Notwendigkeit von Subgruppenanalysen entfällt daher.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich[21]. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“[22], „Multiple-Treatment-Meta-Analysen“[23] oder auch „Netzwerk-Meta-Analysen“[24], sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist[25].

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vorliegen. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen[26].

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*

---

[21] *Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D‘Amico R et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-148.*
[22] *Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.*
[23] *Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005; 331(7521): 897-900.*
[24] *Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. Stat Methods Med Res 2008;17(3): 279-301.*
[25] *B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*
[26] *Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 2009; 338: b1147.*

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*

- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*

- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*

- *Art und Umfang von Sensitivitätsanalysen.*

Teilpopulation a

Für Teilpopulation a ist eine Aussage zu Wahrscheinlichkeit und Ausmaß des Zusatznutzens von Selexipag über die direkt vergleichende Studie GRIPHON möglich. Ein indirekter Vergleich wurde nicht durchgeführt.

Teilpopulation b

Es konnten keine RCT identifiziert werden, die einen direkten Vergleich von Selexipag mit der patientenindividuellen Anwendung von Iloprost zur Inhalation ermöglichen (vgl. Abschnitt 4.3.1.1). Die Durchführung eines indirekten Vergleichs der beiden Wirkstoffe ist ebenfalls nicht möglich, da die Studiendauer der bisher durchgeführten Iloprost-Studien (vgl. Abbildung 4-1) mit 12 Wochen zu kurz für die Bewertung des Zusatznutzens einer Langzeit-therapie bei einer dauerhaften Erkrankung wie der PAH ist. Auch aus methodischen Gründen ist ein indirekter Vergleich nicht möglich, da zwischen den Iloprost-Studien und GRIPHON wesentliche Unterschiede im Studiendesign bestehen.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

– Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
– Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
– Weitere Untersuchungen (Abschnitt 4.3.2.3)

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*

- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungs-studie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen/ abgebrochen/ laufend)**	**Studiendauer**	**Therapie-arme**
GRIPHON (AC-065A302)	ja	ja	abgeschlossen	Ereignisgetriebene Studie. Die Behandlungsphase begann mit der Randomisierung und endete (EOT) mit dem Eintreten eines Ereignisses des primären Endpunkts, dem vorzeitigen Abbruch der Therapie oder dem Studienschluss. Der Studienschluss trat nach Erreichen von 331 CEC-bestätigten Ereignissen des primären Endpunkts ein. Die mediane Behandlungs-dauer betrug 70,7 Wochen im Selexipag-Arm bzw. 63,7 Wochen im Placebo-Arm.	Placebo[1] Selexipag[1]
NS-304-02	ja	ja	abgeschlossen	21 Wochen (Wirksamkeits-erhebung nach 17 Wochen)	Placebo[2] Selexipag[2]

[1]) Während der Studie konnte eine bereits begonnene PAH-spezifische Therapie mit ERA und PDE-5-I fortgeführt werden, sofern ihre Dosierung innerhalb der letzten 3 Monate vor der Basiswerterhebung stabil war. Eine Änderung der Therapie war nach dem Auftreten eines M/M-Ereignisses war erlaubt. Insofern entspricht die Behandlung in GRIPHON einer patientenindividuell optimierten medikamentösen Therapie.

[2]) Während der Studie waren ERA und PDE-5-I als PAH-spezifische Therapie erlaubt. Die Dosierung sollte vom Zeitpunkt des Screenings bis zum Ende der doppelblinden Behandlungsphase stabil gehalten werden.

Geben Sie an, welchen Stand die Information in Tabelle 4-7 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information: 09.05.2016

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-7 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
NS-304-02	A5 – Studiendauer kürzer als 6 Monate (hier: Erhebungszeitraum der Wirksamkeit: 17 Wochen)

4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

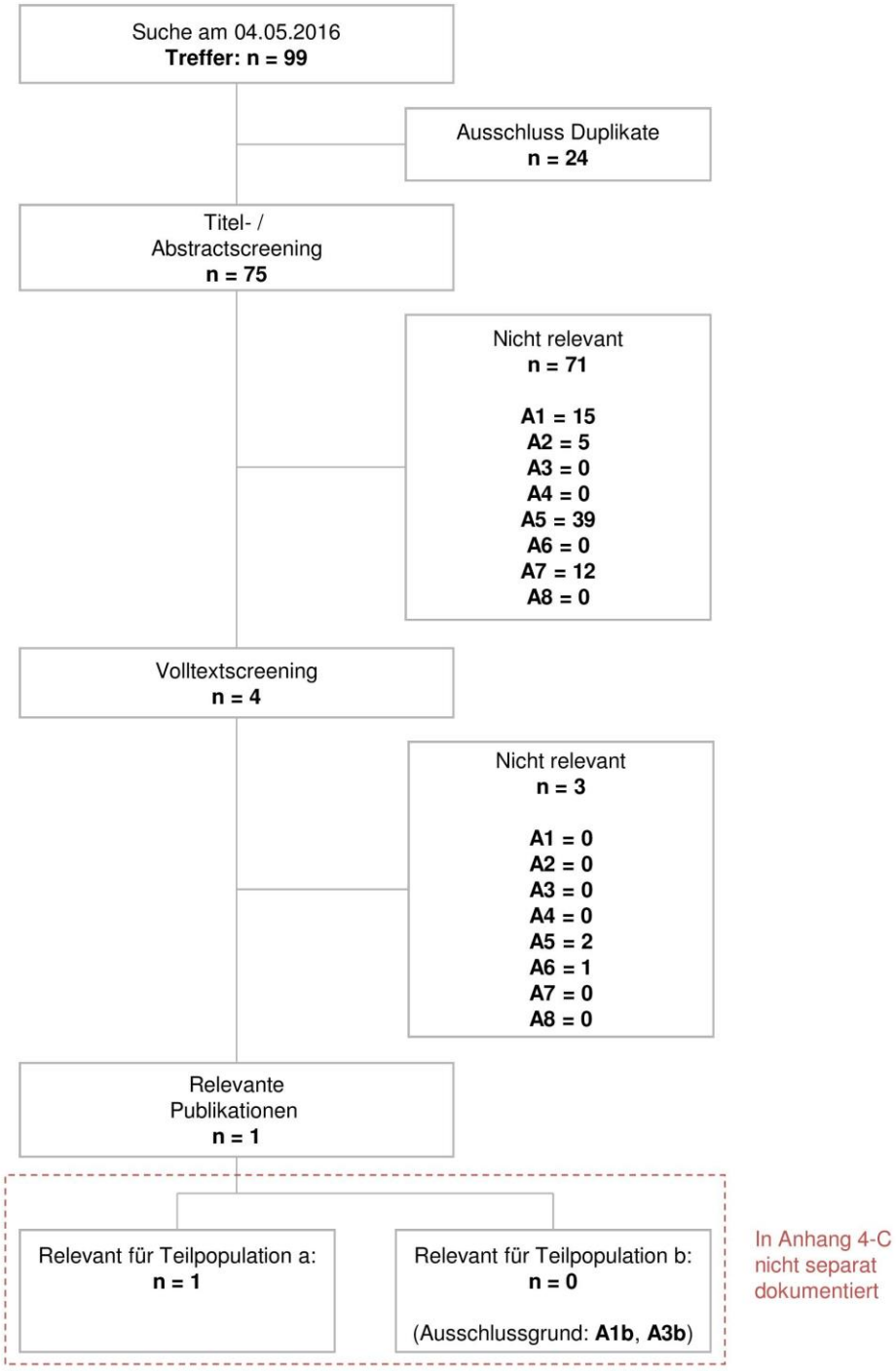


Abbildung 4-3: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	**Identifikationsorte (Name des Studienregisters und Angabe der Zitate[a])**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literatur-recherche identifiziert (ja/nein)**	**Status (abgeschlossen/ abgebrochen/ laufend)**
Teilpopulation a – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungs-zustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt				
GRIPHON (AC-065A302)	Clinicaltrials.gov: NCT01106014 [76] EU-CTR: 2009-014490-41 [77]	ja	ja	abgeschlossen
Teilpopulation b – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungs-zustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt				
keine Studie identifiziert				
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

Geben Sie an, welchen Stand die Information in Tabelle 4-9 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information: 09.05.2016

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 0) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilen-überschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-10: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	**Studienkategorie**			**Verfügbare Quellen[a]**		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	**gesponserte Studie[b] (ja/nein)**	**Studie Dritter (ja/nein)**	**Studien-bericht (ja/nein [Zitat])**	**Register-eintrag[c] (ja/nein [Zitat])**	**Publikation (ja/nein [Zitat])**
Teilpopulation a – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungs-zustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt						
Komparator: abwartendes Vorgehen bis zur weiteren Verschlechterung der PAH; ERA und PDE-5-I (auch in Kombination) waren als begleitende PAH-spezifische Therapie erlaubt						
GRIPHON (AC-065A302)	ja	ja	nein	ja [78] (Post-hoc-Analysen [79])	ja [76, 77]	ja [80]
Teilpopulation b – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungs-zustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt						
keine Studie identifiziert						

a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.

b: Studie, für die der Unternehmer Sponsor war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Tabelle 4-11: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	**Studiendesign**	**Population**	**Interventionen (Zahl der randomisierten Patienten[2])**	**Studiendauer**	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
Teilpopulation a – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegt						
GRIPHON	RCT, ereignis-getrieben, placebo-kontrolliert[1], doppelblind, parallel, multizentrisch	Erwachsene Patienten mit PAH, 6-MWD zwischen 50 und 450 m zum Zeitpunkt des Screenings, eingeschlossene Ätiologien: IPAH, HPAH, durch Medikamente oder Toxine verursachte PAH, assoziierte PAH in Verbindung mit Bindegewebs-erkrankungen, angeborenen systemisch-pulmonalen Shunts (mindestens 1 Jahr nach chirurgischem Verschluss) oder einer HIV-Infektion	Placebo[3] (N= 574) Selexipag (N= 582)	Ereignisgetriebene Studie. Die Be-handlungsphase begann mit der Randomisierung und endete (EOT) mit dem Eintreten eines Ereignisses des primären End-punkts, dem vorzeitigen Ab-bruch der Therapie oder dem Studienschluss. Der Studienschluss trat nach Erreichen von 331 CEC-bestätigten Ereignissen des primären End-punkts ein. Die mediane Behandlungsdauer betrug 70,7 Wochen im Selexi-pag-Arm bzw. 63,7 Wochen im Placebo-Arm.	Asien (China, Indien, Südkorea, Malaysia, Singapur, Taiwan, Thailand, Türkei) Australien Europa (Österreich, Weißrussland, Belgien, Tschechien, Dänemark, Frankreich, Deutschland, Griechenland, Ungarn, Irland, Israel, Italien, Niederlande, Polen, Rumänien, Russische Föderation, Serbien, Slowakei, Spanien, Schweiz, Ukraine, Vereinigtes Königreich) Lateinamerika (Argentinien, Chile, Kolumbien, Peru) Nordamerika (Kanada, Mexiko, USA)	Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH; Zeit bis zum Tod (Studienschluss); Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT); Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss); Absolute Veränderung der 6-MWD; Ausbleibende Verschlechterung der WHO-/NYHA-Klasse; Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH; Absolute Veränderung des Borg-Dyspnoe-Index; Absolute Veränderung des CAMPHOR „Symptome“-Scores; Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores; Unerwünschte Ereignisse; Schwerwiegende unerwünschte Ereignisse; Therapieabbrüche wegen

	12/2009 – 04/2014	unerwünschter Ereignisse; Unerwünschte Ereignisse von besonderem Interesse (Hämorrhagie, Hypotonie, Störung des Auges); Prostacyclin-assoziierte unerwünschte Ereignisse (aufgeteilt nach Titrations- und Erhaltungsphase)
Teilpopulation b – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt		
Keine Studie identifiziert		
[1]) PAH-spezifische Begleittherapie (z. B. ERA, PDE-5-I) war erlaubt [2]) Entspricht in GRIPHON der FAS-Population [3]) Die Behandlung mit Placebo entspricht im Studiendesign von GRIPHON einem abwartenden Vorgehen bis zur Verschlechterung der Erkrankung IPAH – idiopatische PAH, HPAH – angeborene PAH (*hereditary PAH*), FAS – *full analysis set*		

Tabelle 4-12: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	**Selexipag**	**Placebo**[1]
Teilpopulation a – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegt		
GRIPHON	Selexipag, Filmtabletten nach patientenindividueller Dosierung, zweimal täglich + ggf. bestehende PAH-spezifische Begleittherapie	Placebo, Filmtabletten nach patientenindividueller Dosierung, zweimal täglich + ggf. bestehende PAH-spezifische Begleittherapie
	<u>Titration und Anwendungszeitpunkt</u> Patienten erhielten am Abend des ersten Tages die erste Dosierung (200 µg Selexipag bzw. entsprechende Tablettenanzahl Placebo). Beginnend von Tag 2 sollte der Patient das Arzneimittel zweimal täglich und im Abstand von etwa 12 Stunden einnehmen. Sofern die erste Dosierung (200 µg zweimal täglich) gut vertragen wurde, sollte der Patient die Dosierung in wöchentlichen Schritten von 200 µg zweimal täglich so lange steigern, bis die maximal tolerierte Dosis (oder eine Maximaldosierung von 1.600 µg zweimal täglich) erreicht wurde. Der Prüfarzt konnte sich bei jedem Titrationsschritt gegen eine weitere Auftitration entscheiden, sofern nach seiner medizinischen Einschätzung Auftreten und Schwere der typischen Nebeneffekte einer Behandlung mit Agonisten des IP-Rezeptors wie z. B. Kopfschmerzen, Durchfall, Kieferschmerz, Myalgie, Erröten oder Übelkeit für den Patienten nicht mehr tolerierbar waren. In solchen Fällen sollte die Dosierung des Arzneimittels auf die vorherige Dosierung reduziert und als individuelle höchste tolerierte Dosis des Patienten beibehalten werden. Um die Wahrscheinlichkeit des Auftretens typischer pharmakologischer Nebeneffekte der IP-Rezeptorantagonisten-Behandlung zum Zeitpunkt der Auftitration zu vermeiden, sollten die Patienten die erste gesteigerte Dosis am Abend einnehmen. Die individuell höchste tolerierte Dosis des Patienten sollte von Woche 12 bis Woche 26 stabil gehalten werden. Nach Woche 26 waren weitere Steigerungen der Dosis durch den Prüfarzt in Schritten von 200 µg zweimal täglich bis zur Maximaldosierung von 1.600 µg zweimal täglich während der Untersuchungen erlaubt. Reduktionen der Dosierung konnten in diesem Zeitraum bei Verträglichkeitsproblemen jederzeit durchgeführt werden. Die Patienten wurden instruiert, vom Morgen der zweiten Untersuchung an und nachfolgend an jedem weiteren Untersuchungstermin die morgendliche Dosis erst nach Durchführung des 6-MWT (sofern vorgesehen), des CAMPHOR (sofern vorgesehen), der EKG-Untersuchung oder Blutprobenentnahme einzunehmen.	
Teilpopulation b – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt		
Keine Studie identifiziert		

[1]) Die Behandlung mit Placebo entspricht im Studiendesign von GRIPHON einem abwartenden Vorgehen bis zur Verschlechterung der Erkrankung.

Tabelle 4-13: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**		
Merkmal	**Selexipag**	**Placebo**	**Gesamt**
Demographische Werte			
N	**574**	**582**	**1156**
Geschlecht, n (%)			
Männlich	117 (20,4)	116 (19,9)	233 (20,2)
Weiblich	457 (79,6)	466 (80,1)	923 (79,8)
Alter (Jahre)			
Mittelwert (Standardabweichung)	48,2 (15,19)	47,9 (15,55)	48,1 (15,37)
Median (Min; Max)	49,0 (18,0; 78,0)	49,0 (18,0; 80,0)	49,0 (18,0; 80,0)
Altersgruppen, n (%)			
<65 Jahre	475 (82,8)	474 (81,4)	949 (82,1)
65-74 Jahre	91 (15,9)	103 (17,7)	194 (16,8)
≥75 Jahre	8 (1,4)	5 (0,9)	13 (1,1)
Gewicht (kg)			
Fehlende Werte	1	0	1
Mittelwert (Standardabweichung)	71,6 (18,60)	70,7 (17,72)	71,2 (18,16)
Median (Min; Max)	69,4 (39,1; 150,0)	68,0 (38,9; 134,5)	68,1 (38,9; 150,0)
Größe (cm)			
Mittelwert (Standardabweichung)	163,0 (8,75)	162,6 (8,80)	162,8 (8,78)
Median (Min; Max)	162,0 (137,7; 193,8)	162,0 (138,4; 192,0)	162,0 (137,7; 193,8)
BMI (kg/m^2)			
Fehlende Werte	1	0	1
Mittelwert (Standardabweichung)	26,9 (6,40)	26,7 (6,13)	26,8 (6,26)
Median (Min; Max)	25,6 (17,2; 56,1)	25,5 (17,5; 50,4)	25,5 (17,2; 56,1)
BMI-Gruppen, n (%)			
Fehlende Werte	1	0	1
<25 kg/m^2	271 (47,3)	270 (46,4)	541 (46,8)
25-30 kg/m^2	139 (24,3)	163 (28,0)	302 (26,1)
>30 kg/m^2	163 (28,4)	149 (25,6)	312 (27,0)
Rasse/Ethnie, n (%)			
Kaukasisch/weiß oder hispanisch	427 (74,4)	438 (75,3)	865 (74,8)
Kaukasisch/weiß	*376 (65,5)*	*375 (64,4)*	*751 (65,0)*
Hispanisch	*51 (8,9)*	*63 (10,8)*	*114 (9,9)*
Asiatisch	125 (21,8)	120 (20,6)	245 (21,2)
Schwarz	13 (2,3)	14 (2,4)	27 (2,3)
Andere	9 (1,6)	10 (1,7)	19 (1,6)

Geographische Region, n (%)			
Asien	115 (20,0)	113 (19,4)	228 (19,7)
Osteuropa	149 (26,0)	155 (26,6)	304 (26,3)
Lateinamerika	54 (9,4)	56 (9,6)	110 (9,5)
Nordamerika	95 (16,6)	98 (16,8)	193 (16,7)
Westeuropa/Australien	161 (28,0)	160 (27,5)	321 (27,8)
Basiswerte			
Zeit seit der PAH-Diagnose (Jahre)			
Mittelwert (Standardabweichung)	2,3 (3,49)	2,5 (3,75)	2,4 (3,62)
Median (Min; Max)	0,9 (0,0; 37,3)	1,1 (0,0; 38,9)	1,0 (0,0; 38,9)
Ätiologie der PAH, n (%)			
Idiopathisch	312 (54,4)	337 (57,9)	649 (56,1)
Angeboren	13 (2,3)	13 (2,2)	26 (2,2)
Drogen/Medikamente oder Toxine	17 (3,0)	10 (1,7)	27 (2,3)
Bindegewebserkrankung	167 (29,1)	167 (28,7)	334 (28,9)
Angeborener Herzfehler	60 (10,5)	50 (8,6)	110 (9,5)
HIV-Infektion	5 (0,9)	5 (0,9)	10 (0,9)
WHO-/NYHA-Klasse, n (%)			
I	4 (0,7)	5 (0,9)	9 (0,8)
II	274 (47,7)	255 (43,8)	529 (45,8)
III	293 (51,0)	314 (54,0)	607 (52,5)
IV	3 (0,5)	8 (1,4)	11 (1,0)
6-Minuten-Gehstrecke (m)			
Mittelwert (Standardabweichung)	358,5 (76,31)	348,0 (83,23)	353,2 (80,01)
Median (Min; Max)	376,0 (90,0; 482,0)	369,0 (50,0; 515,0)	372,0 (50,0; 515,0)
Borg-Dyspnoe-Index			
Mittelwert (Standardabweichung)	3,6 (2,12)	3,7 (2,11)	3,7 (2,11)
Median (Min; Max)	3,0 (0,0; 10,0)	3,0 (0,0; 10,0)	3,0 (0,0; 10,0)
Systolischer Blutdruck (mmHg)			
Mittelwert (Standardabweichung)	115,0 (16,19)	114,1 (15,41)	114,5 (15,80)
Median (Min; Max)	112,0 (85,0; 186,0)	110,0 (80,0; 170,0)	112,0 (80,0; 186,0)
Diastolischer Blutdruck (mmHg)			
Mittelwert (Standardabweichung)	72,3 (10,67)	71,9 (10,40)	72,1 (10,53)
Median (Min; Max)	70,0 (45,0; 110,0)	70,0 (38,0; 106,0)	70,0 (38,0; 110,0)
Herzfrequenz (Schläge/min)			
Mittelwert (Standardabweichung)	77,3 (12,30)	77,1 (11,83)	77,2 (12,06)
Median (Min; Max)	76,0 (46,0; 115,0)	78,0 (50,0; 120,0)	77,0 (46,0; 120,0)

PAH-Diagnose			
PAH-Diagnose, n (%)			
Typische PAH	439 (76,5)	444 (76,3)	883 (76,4)
Konsistente PAH	97 (16,9)	99 (17,0)	196 (17,0)
Atypische PAH	38 (6,6)	39 (6,7)	77 (6,7)
PAH-spezifische Medikation zum Zeitpunkt der Basiswerterhebung			
Medikation, n (%)			
Keine PAH-spezifische Medikation	112 (19,5)	124 (21,3)	
PAH-spezifische Medikation	462 (80,5)	458 (78,7)	
ERA-Monotherapie	94 (16,4)	76 (13,1)	
Ambrisentan	*26 (4,5)*	*20 (3,4)*	
Bosentan	*67 (11,7)*	*56 (9,6)*	
Sitaxentan	*1 (0,2)*	*0 (0,0)*	
PDE-5-I-Monotherapie	189 (32,9)	185 (31,8)	
Sildenafil	*164 (28,6)*	*169 (29,0)*	
Tadalafil	*23 (4,0)*	*14 (2,4)*	
Vardenafil	*2 (0,3)*	*2 (0,3)*	
Kombination aus ERA- und PDE-5-I-Therapie	179 (31,2)	197 (33,8)	
Ambrisentan-Sildenafil	*44 (7,7)*	*40 (6,9)*	
Ambrisentan-Tadalafil	*13 (2,3)*	*12 (2,1)*	
Bosentan-Sildenafil	*101 (17,6)*	*115 (19,8)*	
Bosentan-Tadalafil	*17 (3,0)*	*22 (3,8)*	
Bosentan-Vardenafil	*2 (0,3)*	*4 (0,7)*	
Sitaxentan-Sildenafil	*2 (0,3)*	*4 (0,7)*	

Min – Minimalwert, Max – Maximalwert, BMI – *body mass index*, FAS – *full analysis set*, ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor

Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Teilpopulation a

Studiendesign

Bei GRIPHON handelt es sich um eine multizentrische, randomisierte, doppelblinde, placebokontrollierte und ereignisgetriebene Phase-3-Studie im Parallelgruppendesign zum

Vergleich der Wirkung (Wirksamkeit, Sicherheit, Verträglichkeit, Pharmakokinetik, Pharmakodynamik) von Selexipag (zweimal täglich verabreicht in individuellen Dosierungen von 200 bis 1.600 µg nach Auftitration) gegenüber Placebo bei Patienten mit symptomatischer PAH.

Ebenso wie die Patienten im Selexipag-Arm konnten auch die Patienten im Placebo-Arm ihre vor dem Studienbeginn begonnene PAH-spezifische Begleittherapie aus ERA und/oder PDE-5-I weiterführen, solange die Dosierung dieser Arzneimittel in den letzten 3 Monaten stabil war. Patienten, die bereits vor Studienbeginn mit Prostanoiden (z. B. Iloprost zur Inhalation, parenterale Anwendung von Treprostinil oder Epoprostenol) behandelt worden waren, konnten an der Studie nicht teilnehmen. Nach Studienbeginn hingegen war die Initiation einer solchen Therapie bei Vorliegen einer Verschlechterung der Erkrankung jederzeit erlaubt, wurde aber direkt (Initiierung einer i. v. oder s. c. Prostanoidtherapie) oder indirekt (in Kombination mit einer Verschlechterung der WHO-/NYHA-Klasse bei Patienten in WHO-/NYHA-Klasse III oder IV) als Ereignis des primären Endpunkts gewertet und beendete damit die Behandlungsphase für den jeweiligen Patienten. Insofern ist das Vorgehen im Placebo-Arm von GRIPHON einem abwartenden Vorgehen bis zur Verschlechterung der Erkrankung gleichzusetzen.

Patienten konnten in die Studie eingeschlossen werden, wenn sie zum Zeitpunkt des Einschluss 18-75 Jahre alt waren und eine mittels Rechtsherzkatheteruntersuchung bestätigte Diagnose der PAH vorlag, die einer der folgenden Subgruppen aus Gruppe 1 des aktualisierten klinischen Klassifikationssystems aus Dana Point 2008 zuzuordnen war:

- Idiopathische PAH
- Angeborene PAH
- Durch Medikamente oder Toxine verursachte PAH
- Assoziierte PAH in Verbindung mit:
    - Bindegewebserkrankungen
    - angeborenen systemisch pulmonalen Shunts (mindestens 1 Jahr nach chirurgischem Verschluss)
    - einer HIV-Infektion

Alle weiteren Einschlusskriterien finden sich in Anhang 4-E.

Die Studie bestand aus den folgenden Phasen (vgl. auch Abbildung 4-4):

- Screening-Phase: Screening-Untersuchungen wurden bis höchstens 28 Tage vor der Basiswert-Untersuchung durchgeführt.
- Behandlungsphase: vom Tag der Basiswertuntersuchung (Untersuchung 1) bis zur Untersuchung am Ende der Studie (EOS-Untersuchung).

Die Patienten wurden instruiert, am Abend des ersten Untersuchungstermins (Untersuchung 1) mit der Einnahme der ersten Studienmedikation (200 µg Selexipag oder Placebo) zu beginnen.

Die Studienmedikation sollte in wöchentlichen Schritten um zweimal täglich 200 µg so lange gesteigert werden, bis die individuell höchste tolerierte Dosis des jeweiligen Patienten oder die Maximaldosis von 1.600 µg in Woche 12 erreicht wurde.

Die Behandlungsphase wurde spätestens mit der EOS-Untersuchung zum Zeitpunkt der Bekanntgabe des Studienschlusses abgeschlossen (d. h. sobald die Gesamtzahl von 331 CEC-bestätigten M/M-Ereignissen erreicht wurde, wobei nur solche Ereignisse gezählt wurden, die spätestens innerhalb von 7 Tagen nach der letzten Einnahme der Studienmedikation begannen[27]). Die EOS-Untersuchung sollte innerhalb von 4 Wochen nach Bekanntgabe des Studienschlusses durchgeführt werden. Bei Patienten, bei denen ein CEC-bestätigtes M/M-Ereignis auftrat oder die die Behandlung mit der Studienmedikation vor dem Studienschluss abgebrochen hatten, wurde die EOS-Untersuchung nach dem Auftreten des Morbiditätsereignisses oder nach dem vorzeitigen Behandlungsabbruch durchgeführt.

- Beobachtungsphase nach der Behandlungsphase (*post-treatment observation period*, PTOP): Patienten, die (unabhängig vom Auftreten eines M/M-Ereignisses) die Behandlung vor Bekanntgabe des Studienschlusses beendeten, waren aufgefordert an einer Beobachtungsphase nach der Behandlungsphase teilzunehmen, um zusätzliche klinische Daten zu erfassen. Diese Phase war definiert als der Zeitraum nach der EOS-Untersuchung aufgrund des vorzeitigen Behandlungsabbruches bis zur Abschlussbehandlung der PTOP im Anschluss an die Bekanntgabe des Studienschlusses durch Actelion.
- Sicherheitsnachbeobachtung nach der Behandlungsphase zur Erkennung von SUE: bei den Patienten wurde jegliches Auftreten von SUE bis 30 Tage nach der letzten Einnahme der Studienmedikation oder bis zur ersten Anwendung von Selexipag in der Studie AC-065A303/GRIPHON OL (je nachdem, was zuerst auftrat) dokumentiert.
- Zum Zeitpunkt des Studienschlusses wurden sämtliche Patienten (mit Ausnahme derer, die ihre Zustimmung zu allen Studienkomponenten zurückgezogen hatten) kontaktiert, um ihren Vitalstatus zu erheben (telefonische Nachverfolgung des Überlebens).
- Behandlungsextensionsphase: Patienten, die zum Zeitpunkt des Studienschlusses noch ihre Medikation aus der Studie AC-065A302/GRIPHON erhalten hatten, wurde eine Behandlungsextensionsphase angeboten. Die Extensionsphase war geplant über einen zeitlichen Verlauf von bis zu 3 Monaten von der EOS-Untersuchung bis zur Untersuchung am Ende der Extensionsphase, die an die Entblindung der AC-065A302/GRIPHON-Datenbank anschloss. Während dieser Phase erhielten die Patienten weiterhin die AC-065A302/GRIPHON-Studienmedikation in doppelverblindeter Weise.

[27] I. d. R. war das auftretende Ereignis aber der Grund für den Abbruch mit der Studienmedikation.

Patienten, die nach Auftreten eines Morbiditätsereignisses die EOS-Untersuchung absolviert hatten, waren anschließend berechtigt, an der Extensionsstudie AC-065A303/GRIPHON OL teilzunehmen, sofern diese von der zuständigen nationalen Gesundheitsbehörde im jeweiligen Land genehmigt worden war.

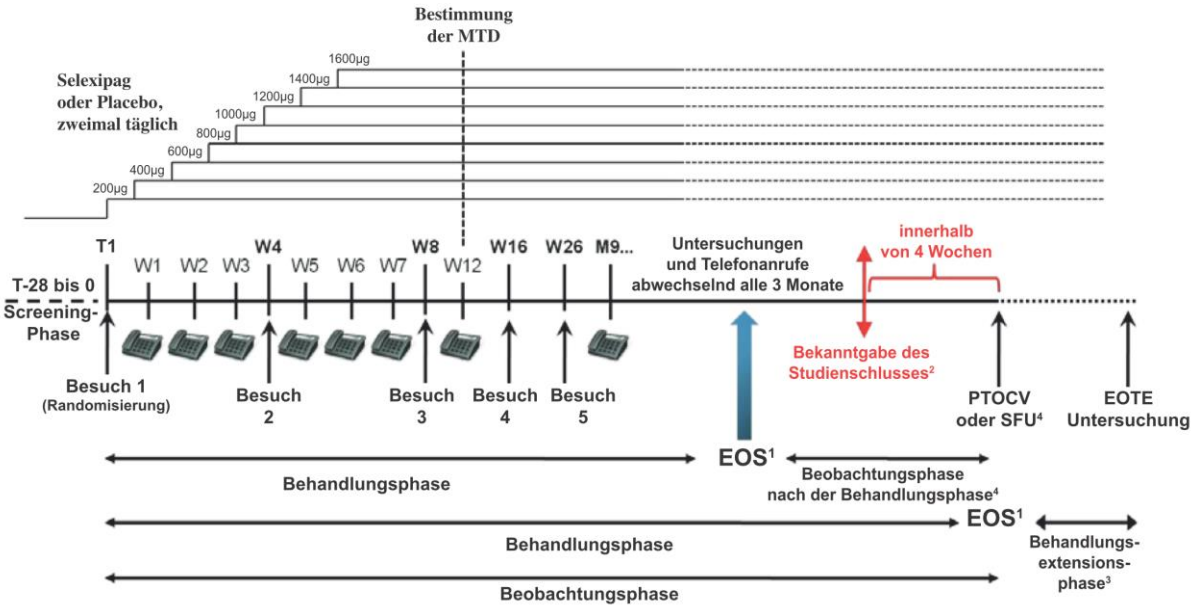


Abbildung 4-4: Studiendesign GRIPHON

[1]) Die EOS-Untersuchung sollte innerhalb von 4 Wochen nach Ankündigung des Studienschlusses durchgeführt werden. Bei Patienten, bei denen ein CEC-bestätigtes M/M-Ereignis aufgetreten war oder die die Behandlung mit der Studienmedikation vorzeitig abgebrochen hatten, wurde die EOS-Untersuchung im Anschluss an das M/M-Ereignis oder den vorzeitigen Abbruch durchgeführt. Bei allen Patienten, die die Behandlung vorzeitig abgebrochen hatten, wurde nachfolgend ein Anruf zur Erhebung sicherheitsrelevanter Informationen durchgeführt.

[2]) Der Studienschluss sollte nach Erreichen der Zielzahl CEC-bestätigter Ereignisse bekanntgegeben werden.

[3]) In den Ländern, in denen AC-065A303/GRIPHON OL durch die zuständigen nationalen Behörden genehmigt worden war, mussten Patienten, die zum Zeitpunkt des Studienschlusses noch mit der Studienmedikation behandelt wurden und in diese Extensionsstudie nach der Feststellung eines positiven Risiko/Nutzen-Verhältnisses von Selexipag in GRIPHON aufgenommen werden wollten, die Behandlungsextensionsphase absolvieren.

[4]) Patienten, die die Studienbehandlung in GRIPHON vor Studienschluss beendeten, konnten an einer Beobachtungsphase nach der Behandlungsphase teilnehmen. Dies war unabhängig von ihrer Teilnahme an der Extensionsstudie GRIPHON OL. Nach Ankündigung des Studienschlusses sollte innerhalb von 4 Wochen eine Abschlussuntersuchung (*post-treatment observation closure visit*, PTOCV) für diese Phase durchgeführt werden. Hierzu wurden alle Patienten (mit Ausnahme derer, die ihre Einwilligung vollständig zurückgezogen hatten) zur Überprüfung ihres Vitalstatus telefonisch kontaktiert.

EOS – Studienende (*end-of-study*, die Abkürzung EOS bezieht sich ausschließlich auf die EOS-Untersuchung, die nicht mit dem eigentlichen Studienschluss gleichzusetzen ist), EOT – Behandlungsende (*end-of-treatment*), EOTE – Ende der Behandlungsextensionsphase (*end-of-treatment-extension*), PTOCV – *post-treatment observation closure visit*, MTD – patientenindividuell

höchste tolerierte Dosis (*maximum tolerated dose*), SFU – Sicherheitsnachbeobachtung (*safety follow-up*)

Individuelle Erhaltungsdosis

Tabelle 4-14 zeigt die Verteilung der verschiedenen individuellen Erhaltungsdosen (*individual maintenance dose*, IMD) in GRIPHON. Als IMD wird hierbei diejenige zweimal tägliche Selexipag- oder Placebodosierung gezählt, der der Patient in der Erhaltungsphase insgesamt am längsten ausgesetzt war. Bei Patienten, die vor Erreichen der Erhaltungsphase aus der Studie ausgeschieden sind, wurde die höchste tolerierte Dosierung in der Titrationsphase als IMD gewertet. Für eine Subgruppenanalyse des primären Endpunkts wurden die Patienten entsprechend ihrer IMD in den Kategorien 200-400 µg zweimal täglich, 600-1.000 µg zweimal täglich und 1.200-1.600 µg zweimal täglich ausgewertet. Es handelt sich hierbei um präspezifizierte Gruppen. Entsprechend dieser Gruppierung waren im Selexipag-Arm 23,1 % in der niedrigsten, 31,2 % in der mittleren und 42,8 % in der höchsten IMD-Kategorie, während es im Placebo-Arm 5,7 %, 11,6 % bzw. 80,4 % waren.

Tabelle 4-14: Anteil individueller Erhaltungsdosen in GRIPHON

Studie: GRIPHON	**Ausgewertete Population: FAS**			
Intervention	Selexipag (N=574)		Placebo (N=582)	
Individuelle Dosierungen				
Zweimal tägliche Dosierung (oder Placebo-Äquivalent)	n	%	n	%
0 µg[1]	14	2,4	9	1,5
200 µg	68	11,8	15	2,6
400 µg	65	11,3	18	3,1
600 µg	62	10,8	20	3,4
800 µg	82	14,3	21	3,6
1.000 µg	35	6,1	27	4,6
1.200 µg	42	7,3	20	3,4
1.400 µg	41	7,1	55	9,5
1.600 µg	163	28,4	393	67,5
Andere[2]	2	0,3	4	0,7

[1]) Patienten, die die initiale Dosierung erhalten hatten, aber auf dieser Dosierung in der Titrationsphase die Behandlung abbrachen

[2]) Patienten, die mit einem anderen Therapieregimen als zweimal täglich behandelt wurden

Patientencharakteristika

Insgesamt wurden in der Studie 1.156 Patienten randomisiert, 574 Patienten davon in der Selexipag-Gruppe und 582 Patienten in der Placebo-Gruppe.

Der Großteil der Teilnehmer der Studie waren weibliche Patienten (79,8 %), der Anteil männlicher Patienten betrug 20,2 %. Im Durchschnitt waren die Patienten in der Studie 48,1±15,37 Jahre alt, 71,2±18,16 kg schwer und 162,8±8,78 cm groß. Gewicht und Größe entsprechen einem durchschnittlichen BMI von 26,8±62,6 kg/m^2. Mit 74,8 % der Gesamstudienpopulation waren hauptsächlich kaukasische/weiße und hispanische Patienten in der Studie vertreten, einen geringeren Anteil bildeten asiatische Patienten (21,2 %). Schwarze Patienten und Patienten anderer Rassen/Ethnien als der bereits genannten, waren nur in sehr geringer Anzahl vertreten (2,3 % bzw. 1,6 %). Die Studienteilnehmer in GRIPHON stammten aus Westeuropa/Australien (28.0 %), Osteuropa (26,0 %), Asien (20,0 %), Nordamerika (16,6 %) und Lateinamerika (9,4 %).

Zum Zeitpunkt der Basiswerterhebung waren im Schnitt 2,4±3,62 Jahre seit der PAH-Diagnose vergangen. Beim Großteil der Patienten in der Studie (56,1 %) war die Ursache der PAH zum Zeitpunkt der Basiswerterhebung nicht bekannt (idiopathische PAH), bei 29,1 % der Patienten war eine Bindegewebserkrankung, bei 10,5 % ein angeborener Herzfehler die Ursache. Weitere PAH-Ätiologien (angeborene PAH, PAH assoziiert mit Drogen/ Medikamenten oder Toxinen oder HIV-Infektionen) waren mit einem Anteil von 0,9 bis 2,2 % vertreten. Bei der PAH-Diagnose[28] zeigten 76,4 % der Patienten eine „typische“, 17,0 % der Patienten eine „konsistente“ und 6,7 % eine „atypische“ PAH.

Bei einem überwiegenden Anteil der Patienten lag die körperliche Einschränkung durch die PAH (erfasst als WHO-/NYHA-Klasse) zu Beginn der Studie im leichten (45,8 % in WHO-/ NYHA-Klasse II) oder deutlichen Bereich (54,0 % in WHO-/NYHA-Klasse III). Nur ein geringer Anteil der Patienten zeigte minimale körperliche Einschränkungen durch die PAH (0,8 % in WHO-/NYHA-Klasse I), ein ebenfalls geringer Anteil war zu Studienbeginn so stark eingeschränkt, dass eine Ausführung körperlicher Aktivitäten ohne Auftreten von Symptomen unmöglich war (1,0 % in WHO-/NYHA-Klasse IV). Somit wurden in GRIPHON in Bezug auf die WHO-/NYHA-Klasse auch Patienten eingeschlossen, die nicht von der Zulassung von Selexipag umfasst sind. Der Anteil der Studienteilnehmer, die aufgrund dessen außerhalb der Zulassung liegen, ist aber so gering (1,8 %), dass eine separate Auswertung nur der Patienten in den WHO-/NYHA-Klassen II und III nicht notwendig ist. Im Durchschnitt waren die Patienten zum Zeitpunkt der Basiswerterhebung in der Lage, in 6 Minuten eine Gehstrecke von 353,2±80,01 m zu bewältigen, und erreichten im Borg-Dyspnoe-Index einen Wert von 3,7±2,11.

Von den Patienten in GRIPHON erhielten zum Zeitpunkt der Basiswerterhebung lediglich 19,5 % (Selexipag) bzw. 21,3 % (Placebo) keine PAH-spezifische Begleittherapie. Der weit überwiegende Teil der Patienten wurde jedoch bereits vor Beginn der Studie zusätzlich mit einem ERA (16,4 % in der Selexipag-Gruppe bzw. 13,1 % in der Placebo-Gruppe), einem PDE-5-I (32,9 % bzw. 31,8 %) oder sogar mit beiden Wirkstoffen, einem ERA und einem

[28] Die Sicherheit der PAH-Diagnose wurde durch zwei verblindete PAH-Experten sichergestellt. Patienten mit „typischer PAH“ hatten keine Komorbidität, die die Ätiologie der PAH infrage stellten, bei Patienten mit „atypischer PAH“ war eine solche Komorbidität vorhanden. Patienten mit „konsistenter PAH“ waren Patienten, die weder in die typische noch in die atypische Kategorie fielen.

PDE-5-I (31,2 % bzw. 33,8 %) behandelt. Diese Behandlung konnte dann in der Studie weitergeführt werden. Betrachtet man die insgesamt 61 Patienten in den deutschen Zentren, so zeigt sich dort, dass alle Patienten zum Zeitpunkt des Studieneinschlusses bereits mit mindestens einer PAH-spezifischen Begleittherapie behandelt wurden [81]. Der Großteil der deutschen Patienten erhielt hierbei eine Kombinationstherapie aus ERA und PDE-5-I (62,3 %), der Anteil an Patienten mit ERA- oder PDE-5-I-Monotherapie war etwa gleich (ERA: 16,4 %, PDE-5-I: 21,3 %).

In keinem der hier dargestellten Charakteristika (demographische Werte, Basiswerte, PAH-Diagnose oder PAH-spezifische Medikation zum Zeitpunkt der Basiswerterhebung) bestanden bemerkenswerte Unterschiede zwischen den Behandlungsgruppen.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Es ist daher von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

Hinweise des G-BA zur zweckmäßigen Vergleichstherapie

Im Rahmen des Beratungsgespräches zu Selexipag vom 03.03.2015, ergänzt durch die Änderungen im Beratungsgespräch von Macitentan vom 14.12.2016, hatte der G-BA die zweckmäßige Vergleichstherapie durch die unten stehenden Hinweise konkretisiert [2, 3]. Im folgenden Abschnitt wird geprüft, inwiefern die Anforderungen, die sich aus den Hinweisen des G-BA ableiten, im Studiendesign von GRIPHON erfüllt sind.

1. „*Es wird vorausgesetzt, dass die Patienten in der vorliegenden Therapiesituation entweder nicht ausreichend auf eine Therapie mit Calciumkanalblocker angesprochen haben oder nicht vasoreaktiv waren.*“

   Im Beratungsgespräch zu Macitentan vom 14.12.2015 [2] hatte der G-BA den vorgenannten Hinweis auf Nachfrage von Actelion näher erläutert und ausgeführt, dass damit nicht eine Voraussetzung beschrieben wird, die im Studiendesign einer relevanten Studie abgebildet sein muss. Vielmehr konkretisiert der Hinweis die zweckmäßige Vergleichstherapie dahin gehend, dass Calciumkanalblocker nicht als mögliche Komparatoren im Rahmen der patientenindividuellen Therapie infrage

kommen. Der G-BA erklärt weiterhin, dass die Therapiesituation, in der PAH-spezifische Arzneimittel (wie z. B. ERA, PDE-5-I, Prostanoide) zur Anwendung kommen, „*der Therapiesituation von Patienten [entspricht], die nicht ausreichend auf eine Therapie mit Calciumkanalblockern angesprochen haben oder nicht vasoreaktiv sind*“.

Insofern erfüllt GRIPHON den vorgenannten Hinweis, da a) kein Vergleich von Selexipag mit Calciumkanalblockern durchgeführt wurde und b) die Anwendung von ERA und PDE-5-I bei den Patienten eine Therapiesituation anzeigt, in der die Patienten nicht ausreichend auf eine Therapie mit Calciumkanalblockern angesprochen haben oder nicht vasoreaktiv waren.

2. „*Es wird davon ausgegangen, dass für therapienaive Patienten eine Kombinationstherapie in der Regel nicht angezeigt ist.*“

   Im Rahmen von GRIPHON war die Anwendung von Selexipag (oder Placebo) entweder als Monotherapie oder in Kombination zu einer bereits bestehenden Therapie aus ERA und/oder PDE-5-I erlaubt. Die Fortführung der bereits bestehenden Therapie konnte jedoch nur dann fortgeführt werden, wenn die Dosierung der verwendeten Arzneimittel für einen Zeitraum von mindestens 3 Monaten vor Eintritt in die Studie stabil geblieben war. Der Beginn einer Therapie mit anderen PAH-spezifischen Arzneimitteln als Selexipag (oder Placebo) bei therapienaiven Patienten als initiale Kombinationstherapie war somit durch die Einschränkungen zur erlaubten Begleitmedikation ausgeschlossen.

3. „*Es wird vorausgesetzt, dass im Interventions- sowie im Vergleichsarm vergleichbare Therapieregime eingesetzt werden.*“

   Der Vergleich von Selexipag und Placebo in GRIPHON erfolgte entweder als Monotherapie oder als sequentielle Kombinationstherapie auf Grundlage einer PAH-spezifischen Begleittherapie aus ERA und/oder PDE-5-I. Die Definition der PAH-spezifischen Begleittherapie erfolgte vor Beginn der Studie individuell für jeden Patienten und nach Maßgabe des Arztes.

   Zwischen den Behandlungsarmen bestanden keine unterschiedlichen Vorgaben in Bezug darauf, welche Arzneimittel oder Arzneimittelkombinationen als Teil der PAH-spezifischen Begleittherapie berücksichtigt werden konnten. Darüber hinaus zeigt sich auch in der Betrachtung der PAH-spezifischen Medikation zum Zeitpunkt der Basiswerterhebung kein systematischer Unterschied zwischen den Behandlungsarmen in Bezug auf die angewendeten Therapieregime (vgl. Tabelle 4-13).

   Es ist daher davon auszugehen, dass der genannte Hinweis, dass im Interventions- sowie im Vergleichsarm vergleichbare Therapieregime angewendet werden, in GRIPHON erfüllt ist.

4. „*Der Wirkstoff Riociguat wird aufgrund der bisher kurzen Marktverfügbarkeit nicht als in der praktischen Anwendung bewährt angesehen und ist von der zweckmäßigen Vergleichstherapie ausgenommen.*“[29]

   In GRIPHON erfolgte ein Vergleich von Selexipag mit Placebo, wobei das Vorgehen im Placebo-Arm letztlich einem abwartenden Vorgehen bis zur Verschlechterung der Erkrankung gleichzusetzen ist. Ein Vergleich gegenüber Riociguat erfolgte nicht.

   Zwar wäre die Behandlung mit Riociguat als Teil der PAH-spezifischen Begleittherapie im Studiendesign von GRIPHON grundsätzlich erlaubt gewesen, tatsächlich wurde aber kein Patient in der Studie mit Riociguat behandelt.

   Insofern ist der Hinweis, dass der Wirkstoff Riociguat aus der zweckmäßigen Vergleichstherapie ausgenommen ist, im Studiendesign von GRIPHON erfüllt.

Auswertungspopulation

Full analysis set(FAS)-Population

Die FAS-Population beinhaltet alle Patienten, die in GRIPHON randomisiert wurden. Um das *Intention to treat*-Prinzip bestmöglich zu erfüllen, wurden alle Patienten in der Therapiegruppe ausgewertet, auf die sie randomisiert wurden (die tatsächlich eingenommene Studienmedikation konnte hiervon abweichen), und es wurden alle verfügbaren Wirksamkeitsdaten verwendet. Die FAS-Population stellt die Hauptpopulation für die Wirksamkeitsanalyse dar.

Für die Analyse des Endpunkts „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ wurde eine leicht modifizierte FAS-Population verwendet. In dieser modifizierten Population wurden Patienten, die zum Studienbeginn in WHO-/NYHA-Klasse IV waren, nicht in die Analyse eingeschlossen, da sie sich nicht weiter hätten verschlechtern können und daher fälschlicherweise als Responder gewertet worden wären.

Per Protocol Set(PPS)-Population

Die PPS-Population beinhaltete alle Patienten der FAS-Population, die:

- zum Zeitpunkt des Screenings 18 Jahre oder älter waren;
- einer der Subgruppen der aktualisierten klinischen Klassifikation der PAH Gruppe 1 aus Dana Point angehörten;
- zumindest eine Anwendung der Studienmedikation erhalten hatten;

[29] Der Wortlaut dieses Hinweises ergibt sich aus der Formulierung in der Niederschrift zu Macitentan vom 14.12.2015 i. V. mit der dort vom G-BA getroffenen Aussage, dass „*die zweckmäßige Vergleichstherapie zum Wirkstoff Selexipag analog angepasst wird*“. In der Niederschrift zum Beratungsgespräch vom 03.03.2015 war der entsprechende Hinweis folgendermaßen formuliert: „*Mit Riociguat und Macitentan stehen für das vorliegende Anwendungsgebiet zugelassene, in der Versorgung noch recht neue Behandlungsoptionen zur Verfügung, deren therapeutischer Stellenwert derzeit noch nicht abschließend beurteilbar ist.*“

- die Einnahme ihrer Studienmedikation für einen Zeitraum von mindestens 50 % der Zeit zwischen erster und letzter Einnahme nicht unterbrochen hatten;
- im Zeitraum von erster bis letzter Einnahme der Studienmedikation nicht länger als 4 Wochen ohne Behandlung waren;
- einen gültigen 6-MWT am ersten Untersuchungstermin aufwiesen, wobei als gültig angesehen wurde, wenn der 6-MWT:
    - vor der ersten Einnahme der Studienmedikation durchgeführt worden war,
    - nicht aus anderen Gründen als der PAH unterbrochen worden war.

Die PPS-Population enthielt zudem zusätzlich sämtliche Daten bis zum Auftreten eines der folgenden Ereignisse:

- Initiierung einer Therapie mit PDE-5-I, ERA, Prostacyclin oder Prostacyclin-Analoga nach der Randomisierung für mehr als 7 Tage, ohne dass ein CEC-bestätigtes M/M-Ereignis aufgetreten war oder vor dessen Aufzeichnungszeitpunkt (ein Wechsel innerhalb der Wirkstoffklasse führte nicht zum Ausschluss)
- Einnahme einer anderen Studienmedikation als der, auf die der Patient randomisiert wurde
- Individuelle Entblindung der Gruppenzugehörigkeit durch den Prüfarzt

Die Patienten in der PPS-Population wurden entsprechend der Gruppenzugehörigkeit analysiert, auf die sie randomisiert wurden.

Die Analyse der PPS-Population wurde zusätzlich zur Analyse der FAS-Population zur Überprüfung der Robustheit der Wirksamkeit bei ausgewählten Endpunkten durchgeführt.

Quality of life analysis set(QAS)-Population

Die QAS-Population beinhaltete sämtliche Patienten der FAS-Population, für die eine geeignete Übersetzung des CAMPHOR-Fragebogens im Studienzentrum vorlag.

Sicherheitspopulation (SAF)

Die SAF-Population beinhaltete alle randomisierten Patienten, die zumindest eine Anwendung der Studienmedikation erhalten hatten. Die Auswertung erfolgte auf Grundlage der tatsächlich eingenommenen Studienmedikation (die sich von der Therapiegruppe unterscheiden konnte, auf die der Patient randomisiert worden war). Hierbei galt, dass Patienten, die mindestens eine Dosis Selexipag erhalten hatten, in der Selexipag-Gruppe ausgewertet wurden, während Patienten, die nur Placebo erhalten hatten, in der Placebo-Gruppe ausgewertet wurden.

Mit der SAF-Population wurden sämtliche Analysen der Sicherheitsendpunkte durchgeführt.

Teilpopulation b

Es konnte keine relevante RCT für die Nutzenbewertung von Selexipag in dieser Teilpopulation identifiziert werden.

4.3.1.2.2 Verzerrungspotential auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-15: Verzerrungspotential auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotential auf Studienebene
			Patient	Behandelnde Personen			
Teilpopulation a – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegt							
GRIPHON	ja	ja	ja	ja	ja	ja	niedrig
Teilpopulation b – Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt							
Keine Studie identifiziert							

Begründen Sie für jede Studie die abschließende Einschätzung.

Teilpopulation a

Die vorliegende Bewertung des Verzerrungspotentials begründet sich auf der Auswertung der Studienmethodik in Anhang 4-E sowie auf den in Anhang 4-F dargestellten Bögen zur Bewertung des Verzerrungspotentials.

Die Erzeugung der Randomisierungssequenz wurde in GRIPHON adäquat durchgeführt, die Gruppenzuteilung fand verdeckt statt. Bei GRIPHON handelt es sich um eine doppelt verblindete Studie, d.h. sowohl Behandler als auch Patient waren verblindet. Die Ereignisse des primären Endpunkts wurden zudem durch ein verblindetes Critical Event Committee dahin gehend begutachtet, ob ein direkter/unmittelbarer Zusammenhang zur PAH besteht.

Hinweise auf eine ergebnisabhängige Berichterstattung sind nicht zu beobachten, und es liegen auch keine sonstigen endpunktübergreifenden Aspekte vor, die zu einer Verzerrung der Ergebnisse führen könnten.

Das Verzerrungspotential in der dargestellten Studie ist somit als niedrig anzusehen.

Teilpopulation b

Es konnte keine relevante RCT für die Nutzenbewertung von Selexipag in dieser Teilpopulation identifiziert werden.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Durch die in Abschnitt 4.3.1.1 beschriebene Informationsbeschaffung konnte lediglich für die Teilpopulation a eine relevante RCT – GRIPHON – identifiziert werden. Für Teilpopulation b liegt keine relevante RCT vor. Die Darstellung im folgenden Abschnitt 4.3.1.3 wird daher auf die Ergebnisse zu Wahrscheinlichkeit und Ausmaß des Zusatznutzens in Teilpopulation a beschränkt.

Tabelle 4-16: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Endpunkt	**Studie**
	GRIPHON
Kombinierte Endpunkte mit Anteilen aus verschiedenen Endpunktkomplexen	
Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis	ja
Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH	ja
Mortalität	
Zeit bis zum Tod (Studienschluss)	ja
Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)	ja
Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)	ja
Morbidität	
Zeit bis zur ersten Hospitalisierung wegen PAH	ja
Symptome der PAH	ja
Absolute Veränderung der 6-MWD	ja
Ausbleibende Verschlechterung der WHO-/NYHA-Klasse	ja
Absolute Veränderung des Borg-Dyspnoe-Index	ja
Lebensqualität	
Absolute Veränderung des CAMPHOR „Symptome“-Scores	ja
Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores	ja
Nebenwirkungen	
Unerwünschte Ereignisse	ja
Schwerwiegende unerwünschte Ereignisse	ja
Therapieabbrüche wegen unerwünschter Ereignisse	ja
Unerwünschte Ereignisse von besonderem Interesse	ja
Prostacyclin-assoziierte unerwünschte Ereignisse	ja

4.3.1.3.1 Ergebnisse patientenrelevanter Endpunkte – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

– Ergebnisse der ITT-Analyse
– Zahl der Patienten, die in die Analyse eingegangen sind
– dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
– bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
– bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
– entsprechende Maße bei weiteren Messniveaus
– Effektschätzer mit zugehörigem Standardfehler
– Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 0). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

4.3.1.3.1.1 Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-17: Operationalisierung von „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“

Studie	Operationalisierung

GRIPHON	*Definition des Endpunkts*
	Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis.
	Es handelt sich um einen kombinierten Endpunkt. Als Morbiditäts- oder Mortalitätsereignis wurden die folgenden Ereignisse gezählt:
	• Tod (jeglicher Ursache)
	oder
	• Hospitalisierung aufgrund einer Verschlechterung der PAH[1]
	oder
	• Lungentransplantation oder atriale Septostomie aufgrund einer Verschlechterung der PAH
	oder
	• Initiierung einer Prostanoidtherapie (i. v. oder s. c.) oder einer dauerhaften O_2-Therapie
	oder
	• Fortschreiten der Erkrankung:
	  ○ Für Patienten in den WHO-/NYHA-Klassen II-III zum Zeitpunkt der Basiswerterhebung:
	  Verschlechterung der 6-MWD ≥15 % im Vergleich zum Basiswert (bestätigt durch 2 unabhängige Tests an unterschiedlichen Tagen in einem Zeitraum von 2 Wochen)
	  und
	  Verschlechterung der WHO-/NYHA-Klasse
	  ○ Für Patienten in den WHO-/NYHA-Klassen III-IV zum Zeitpunkt der Basiswerterhebung:
	  Verschlechterung der 6-MWD ≥15 % im Vergleich zum Basiswert (bestätigt durch 2 unabhängige Tests an unterschiedlichen Tagen in einem Zeitraum von 2 Wochen)
	  und
	  Notwendigkeit einer zusätzlicher PAH-spezifischen Therapie
	Sämtliche Ereignisse des primären Endpunkts wurden durch ein verblindetes, unabhängiges Critical Event Committee (CEC) begutachtet. Das CEC bestand aus 3 klinischen Experten mit Erfahrung in der PAH und in der Überwachung klinischer Studien. Als Mitglied des CEC war es nicht möglich, gleichzeitig als Prüfarzt an der Studie beteiligt zu sein. Die Hauptaufgabe des CEC war die Bestätigung des Auftretens eines Morbiditäts- oder Mortalitätsereignisses, gleichzeitig sollten aber auch alle Todesfälle auf einen möglichen Zusammenhang zur PAH überprüft werden.
	Die Hauptanalyse des Endpunkts wurde für den Zeitraum von der Randomisierung bis 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase (*end-of-treatment*, EOT) vorgenommen. Auswertungspopulation für die Hauptanalyse war die FAS-Population.

	Zensierungsregeln/Ersetzungsregeln Es wurde keine Imputationsmethode angewendet. Bei Patienten, bei denen bis zu 7 Tage nach der letzten Einnahme der Studienmedikation kein CEC-bestätigtes M/M-Ereignis eingetreten war, wurde die Zeit bis zum ersten CEC-bestätigten M/M-Ereignis nach folgenden Regeln zensiert: • Randomisierte Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und keine Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302-Behandlungsphase plus 7 Tage, Zeitpunkt der EOS-Untersuchung, Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. • Randomisierte Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und ihre Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302-Behandlungsphase plus 7 Tage, Zeitpunkt des letzten Kontakts, 27.04.2014) minus Zeitpunkt der Randomisierung plus 1 Tag. • Randomisierte Patienten, die keine Studienmedikation erhalten hatten: Minimum (Maximum [Zeitpunkt der EOS-Untersuchung, Zeitpunkt der Randomisierung], Zeitpunkt des letzten Kontakts, 27.04.2014) minus Zeitpunkt der Randomisierung plus 1 Tag. Entsprechend dem 5. Global Amendment wurden CEC-bestätigte M/M-Ereignisse, die (nach Einschätzung des CEC) vor dem 4. Global Amendment vom 16.08.2011 aufgetreten waren, zum Zeitpunkt des Auftretens des Ereignisses zensiert. In solchen Fällen, in denen Patienten mit einem M/M-Ereignis mit Auftreten vor dem 16.08.2011 ein zweites Ereignis hatten, das nach dem 16.08.2011 aufgetreten war, wurde das erste Ereignis verworfen und nur das zweite in die Analyse eingeschlossen. Es handelt sich hierbei um rechts-zensierte Beobachtungen.
[1]) Jeglicher unfreiwillige Krankenhausaufenthalt (≥24 Stunden) aufgrund einer Verschlechterung der PAH; Verschlechterung der PAH schließt Anzeichen und Symptome des Rechtsherzversagens (z. B. (Prä-)Synkope, Zyanose, Verschlechterung der Kurzatmigkeit, klinisch relevante Verschlechterung der körperlichen Leistungsfähigkeit, Abfall der arteriellen Sauerstoffsättigung, verschlimmerte Ödeme, Hepatomegalie und Ascites) mit ein.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-18: Bewertung des Verzerrungspotentials für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-19: Ergebnisse für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis-Ereignis" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**	
Intervention	Selexipag N=574	Placebo N=582
Zeit bis zum ersten Ereignis (Monate)		
1. Quartil [95 % KI]	18,1 [15,9; 21,9]	9,0 [7,3; 12,0]
Median [95 % KI]	- [36,4; -]	29,0 [23,8; 34,8]
3. Quartil [95 % KI]	-	-
Kaplan-Meier-Schätzer ereignisfreier Patienten		
Monat 6		
Patienten unter Risiko	455	433
Zensierte Patienten	73	48
Patienten mit Ereignis	46	101
KM-Schätzer, % [95 % KI]	91,2 [88,5; 93,4]	81,7 [78,2; 84,7]
Monat 12		
Patienten unter Risiko	361	347
Zensierte Patienten	128	77
Patienten mit Ereignis	85	158
KM-Schätzer, % [95 % KI]	83,1 [79,5; 86,1]	70,6 [66,6; 74,3]
Monat 18		
Patienten unter Risiko	246	220
Zensierte Patienten	214	165
Patienten mit Ereignis	114	197
KM-Schätzer, % [95 % KI]	75,5 [71,2; 79,2]	61,3 [56,8; 65,5]
Monat 24		
Patienten unter Risiko	171	149
Zensierte Patienten	267	212
Patienten mit Ereignis	136	221
KM-Schätzer, % [95 % KI]	67,9 [63,0; 72,4]	53,9 [49,0; 58,5]
Monat 30		
Patienten unter Risiko	101	88
Zensierte Patienten	324	262
Patienten mit Ereignis	149	232
KM-Schätzer, % [95 % KI]	61,9 [56,3; 66,9]	49,3 [44,2; 54,3]

Monat 36		
Patienten unter Risiko	40	28
Zensierte Patienten	381	313
Patienten mit Ereignis	153	241
KM-Schätzer, % [95 % KI]	58,2 [51,7; 64,1]	41,7 [35,2; 48,0]
Monat 42		
Patienten unter Risiko	7	4
Zensierte Patienten	412	336
Patienten mit Ereignis	155	242
KM-Schätzer, % [95 % KI]	55,1 [47,5; 62,0]	39,2 [31,6; 46,8]
Monat 48		
Patienten unter Risiko	1	-
Zensierte Patienten	418	340
Patienten mit Ereignis	155	242
KM-Schätzer, % [95 % KI]	55,1 [47,5; 62,0]	39,2 [31,6; 46,8]
Behandlungseffekt (Selexipag gegenüber Placebo)		
HR [99 % KI]		0,60 [0,46; 0,78]
p (Log-Rang-Test)		<0,0001

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen.
CEC – *Critical Event Committee*, KI – Konfidenzintervall, KM-Schätzer – Kaplan-Meier-Schätzer, HR – Hazard Ratio, FAS – *full analysis set*

Hauptanalyse des Endpunkts

Die Analyse der Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis wurde auf Grundlage der FAS-Population durchgeführt.

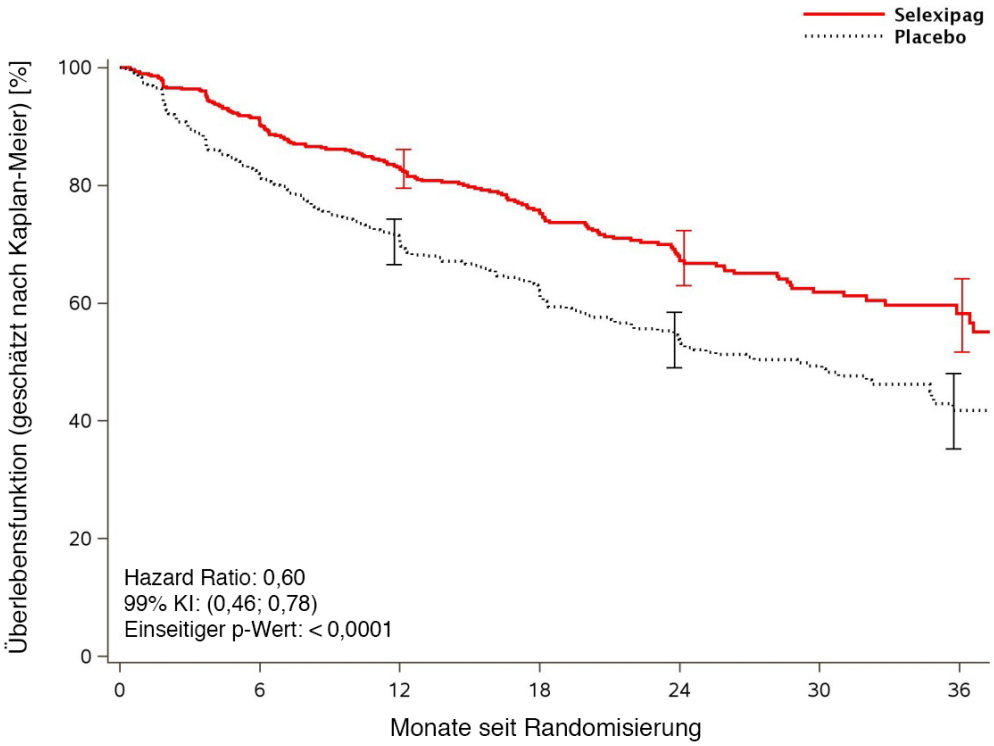
Um Bedenken ausschließen zu können, dass die Protokolländerungen, die mit der 4. globalen Änderung des Studienprotokolls vom 16. August 2011 eingeführt wurden, als informiert angesehen werden könnten, wurde beschlossen, alle vor diesem Datum aufgetretenen M/M-Ereignisse zu zensieren und nicht in die Hauptanalyse des Studienberichts einzuschließen. Der Vergleich der Analysen mit und ohne Zensierung der genannten Ereignisse (vgl. Tabelle 4-19 und Tabelle 4-20) zeigt jedoch nahezu identische Ergebnisse in beiden Analysen. Analog zum Vorgehen im Studienbericht (Berücksichtigung tatsächlich aller Patientendaten, die im Rahmen der Studie erhoben wurden) wird daher auch im Nutzendossier die Analyse des Endpunkts unter Einschluss auch der Ereignisse vor dem 16. August 2011 als Hauptanalyse angesehen. Zudem werden Sensitivitätsanalyse und andere supportive Analysen des primären Endpunkts ebenfalls unter Einschluss dieser Ereignisse durchgeführt.

Insgesamt trat im Beobachtungszeitraum (EOT + 7 Tage) bei 155 Patienten in der Selexipag-Gruppe und bei 242 Patienten in der Placebo-Gruppe ein CEC-bestätigtes kombiniertes M/M-Ereignis auf. Die Hazard Ratio (HR) [99 % Konfidenzintervall (KI)] in der

Ereigniszeitanalyse für das Auftreten eines CEC-bestätigten M/M-Ereignisses betrug 0,60 [0,46; 0,78] und war statistisch signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p<0,0001). Dies entspricht einer Reduktion des relativen Risikos von 40 %.

Die mediane Zeit bis zum ersten Ereignis wurde im Selexipag-Arm im Verlauf der Behandlungsphase noch nicht erreicht, im Placebo-Arm betrug sie 29,0 Wochen (95 % KI: 23,8; 34,8).

Abbildung 4-5 zeigt die Kaplan-Meier-Kurve des Endpunkts in der FAS-Population. Die Trennung zwischen der Selexipag-Gruppe und der Placebo-Gruppe erfolgte früh und ist nach 6 Monaten bereits sehr deutlich (Kaplan-Meier-Schätzer ereignisfreier Patienten: 91,2 % in der Selexipag-Gruppe vs. 81,7 % in der Placebo-Gruppe). Zu allen untersuchten Zeitpunkten war der Kaplan-Meier-Schätzer in der Placebo-Gruppe niedriger als in der Selexipag-Gruppe.



Selexipag-Patienten:							
unter Risiko	574	455	361	246	171	101	40
Ereignisse	0	46	85	114	136	149	153
zensiert	0	73	128	214	267	324	381

Placebo-Patienten:							
unter Risiko	582	433	347	220	149	88	28
Ereignisse	0	101	158	197	221	232	241
zensiert	0	48	77	165	212	262	313

Abbildung 4-5: Kaplan-Meier-Kurve für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Fehlerbalken geben das 95 % KI an. CEC – *Critical Event* Committee, FAS – *full analysis set*, KI – Konfidenzintervall

Nebenanalysen des Endpunkts

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf den Zeitraum von der Randomisierung bis EOT + 7 Tage.

Sensitivitätsanalysen

Um die Robustheit der Ergebnisse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ einschätzen zu können, wurden verschiedene Sensitivitätsanalysen durchgeführt. Diese sind in der folgenden Tabelle 4-20 dargestellt.

Wie von der FDA aufgrund der Protokolländerung im 4. globalen Amendment vom 16. August 2011 gefordert, wurden für die Hauptanalyse des Endpunkts im Studienbericht

sämtliche Ereignisse, die vor dem 16. August 2011 aufgetreten waren, in der Analyse des primären Endpunkts zensiert. Dennoch ergab sich ein sehr ähnliches Ergebnis für den Endpunkt (HR [99 % KI]: 0,61 [0,46; 0,81], p>0,0001), sodass beschlossen wurde, alle weiteren Analysen wie ursprünglich geplant unter Einschluss aller Ereignisse durchzuführen.

In weiteren Sensitivitätsanalysen wurde der Einfluss der Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“ auf die Ergebnisse des primären Endpunkts untersucht. Der Grund hierfür ist, dass beide Komponenten, insbesondere jedoch „Fortschreiten der Erkrankung“, vom G-BA als von unterschiedlicher Schwere im Vergleich zu den anderen Komponenten des primären Endpunkts angesehen werden [3]. Für die Komponente „Fortschreiten der Erkrankung“ bestehen seitens des G-BA zudem Zweifel an der Patientenrelevanz [3]. Es soll darauf hingewiesen werden, dass die genannten Komponenten trotz der vorgenommenen Sensitivitätsanalysen weiterhin von Actelion als patientenrelevant angesehen werden. Analysen unter Ausschluss des „Fortschreitens der Erkrankung“ (HR [95 % KI]: 0,73 [0,58; 0,92]) bzw. unter Ausschluss beider Komponenten (HR [95 % KI]: 0,75 [0,59; 0,96]) zeigten jedoch, dass deren Einfluss auf die Ergebnisse des kombinierten Endpunkts beschränkt ist. Zwar verringert sich die Reduktion des relativen Risikos durch den Ausschluss der Komponente „Fortschreiten der Erkrankung“ von 40 % zu 27 %, es handelt sich jedoch weiterhin um einen deutlichen, signifikanten (p=0,0037) und überdies klinisch relevanten Effekt. Der zusätzliche Ausschluss der Komponente „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“ wiederum hat nur einen sehr geringen Effekt auf das Ergebnis der Ereigniszeitanalyse. Die relative Risikoreduktion in dieser Sensitivitätsanalyse betrug 25 %.

Weiterhin wurde eine Sensitivitätsanalyse in der PPS-Population durchgeführt. Ebenso wie in der FAS-Population war auch bei der Analyse der PPS-Population der Unterschied zwischen den Behandlungsgruppen statistisch signifikant (p<0,0001). Der Effektschätzer des Behandlungseffektes (HR [99 % KI]: 0,58 [0,44; 0,76]) war zudem konsistent zum Effektschätzer der Analyse in der FAS-Population. Die relative Risikoreduktion für das Auftreten eines M/M-Ereignisses in dieser Sensitivitätsanalyse betrug 42 %.

Zum Ausschluss möglicher informativer Zensierungen in GRIPHON wurden verschiedene Sensitivitätsanalysen durchgeführt, die diesen Sachverhalt untersuchen sollten. In allen Sensitivitätsanalysen zur möglichen informativen Zensierung zeigte sich jedoch ein signifikanter Unterschied zwischen den Behandlungsgruppen. Der Effektschätzer für den Behandlungseffekt war überdies konsistent zur Hauptanalyse. Die relative Risikoreduktion für das Auftreten eines M/M-Ereignisses bewegte sich innerhalb dieser Sensitivitätsanalysen zwischen 35 % und 40 %.

In einer weiteren Sensitivitätsanalyse wurde der primäre Endpunkt unter Annahme der ursprünglichen Fallzahlplanung (angepasst durch das 5. Amendment des Studienprotokolls), d. h. mit den ersten 670 randomisierten Patienten und bis zum Auftreten von 202 CEC-bestätigten M/M-Ereignissen ausgewertet. Auch in dieser Analyse war der Effektschätzer des

Behandlungseffekts konsistent zur Hauptanalyse (HR [99 % KI]: 0,65 [0,45; 0,95]) sowie der Unterschied zwischen den Behandlungsgruppen signifikant (p=0,0014). Die relative Risikoreduktion für das Auftreten eines M/M-Ereignisses betrug in der Sensitivitätsanalyse 35 %.

Werden nur Patienten in die Auswertung des primären Endpunkts eingeschlossen, die von den Prüfärzten als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden (entsprechen etwa 93 % der Gesamtstudienpopulation), so ergibt sich eine HR [99 % KI] von 0,62 [0,47; 0,81] und damit eine relative Risikoreduktion für das Auftreten eines M/M-Ereignisses von 38 %. Der Unterschied zwischen den Behandlungsgruppen war auch in dieser Sensitivitätsanalyse statistisch signifikant (p<0,0001).

Eine Sensitivitätsanalyse, in der alle vom Prüfarzt gemeldeten M/M-Ereignisse in die Analyse eingeschlossen wurden, unabhängig davon, ob sie vom CEC bestätigt wurden oder nicht, zeigte ebenfalls zur Hauptanalyse konsistente Ergebnisse. Die HR [99 % KI] in der Sensitivitätsanalyse betrug 0,64 [0,50; 0,82], der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (p<0,0001). Die relative Risikoreduktion für das Auftreten eines M/M-Ereignisses betrug in dieser Sensitivitätsanalyse somit 36 %.

In Bezug auf die Komponente „Fortschreiten der Erkrankung“ kam es vor, dass die Werte für den zweiten, bestätigenden 6-MWT nicht vorlagen. Für diese Fälle war es im Studienprotokoll vorgesehen, dass die Entscheidung, ob tatsächlich ein Fortschreiten der Erkrankung vorliegt, durch das CEC getroffen werden sollte. Dies war bei insgesamt 5 Patienten (0,9 %) im Selexipag-Arm und bei 10 Patienten (1,7 %) im Placebo-Arm der Fall. Eine Sensitivitätsanalyse, in der sämtliche Ereignisse der Komponente „Fortschreiten der Erkrankung“ ausgeschlossen wurden, bei denen kein zweiter 6-MWT zur Bestätigung durchgeführt wurde, zeigte zur Hauptanalyse konsistente Ergebnisse (HR [99 % KI]: 0,60 [0,46; 0,78], p<0,0001). Die relative Risikoreduktion für das Auftreten eines M/M-Ereignisses betrug in dieser Sensitivitätsanalyse somit 40 %.

Insgesamt gesehen bestätigen sämtliche der durchgeführten Sensitivitätsanalysen das Ergebnis der Hauptanalyse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ und unterstreichen damit die Robustheit der Ergebnisse des Endpunkts.

Tabelle 4-20: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“

Sensitivitätsanalyse	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Analyse unter Ausschluss der Ereignisse vor dem 16. August 2011	
Analyse unter Ausschluss der Ereignisse, die vor der Protokolländerung am 16. August 2011 aufgetreten waren (insgesamt 45 Ereignisse ausgeschlossen)	0,61 [0,46; 0,81] p<0,0001
Analyse unter Ausschluss einzelner Komponenten	
Analyse unter Ausschluss der Komponente „Fortschreiten der Erkrankung“	0,73 [0,58; 0,92] p=0,0037
Analyse unter Ausschluss der Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstoff-therapie aufgrund einer Verschlechterung der PAH“	0,75 [0,59; 0,96] p=0,0107
Per-protocol-Analyse	
Analyse in der PPS-Population	0,58 [0,44; 0,76] p<0,0001
Analysen zur Untersuchung möglicher informativer Zensierung	
Analyse unter Einschluss aller CEC-bestätigten Ereignisse und aller unbestätigten Ereignisse, die zum Abbruch der Therapie führten	0,61 [0,47; 0,79] p<0,0001
Analyse unter Einschluss aller CEC-bestätigten Ereignisse und der *Lost-to-follow-up*-Patienten in der Behandlungsphase[1]	0,60 [0,46; 0,79] p<0,0001
Analyse unter Einschluss aller CEC-bestätigten Ereignisse und der Patienten, die kein bestätigtes Ereignis hatten, aber bei denen zum Zeitpunkt des vorzeitigen Therapieabbruchs Anzeichen der Krankheitsverschlechterung beobachtet wurden	0,65 [0,51; 0,83] p<0,0001

Weitere Sensitivitätsanalysen	
Analyse des Endpunkts bei den ersten 670 randomisierten Patienten und bis zum Auftreten von 202 CEC-bestätigten Ereignissen[2]	0,65 [0,45; 0,95] p=0,0014
Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden	0,62 [0,47; 0,81] p<0,0001
Analyse unter Einschluss der vom Prüfarzt gemeldeten Ereignisse unabhängig von ihrer Bestätigung durch das CEC	0,64 [0,50; 0,82] p<0,0001
Analyse unter Ausschluss von Ereignissen der Komponente „Fortschreiten der Erkrankung“, wenn kein zweiter 6-MWT zur Bestätigung durchgeführt wurde	0,60 [0,46; 0,78] p<0,0001

[1]) Für die Analyse wurde bei den Lost-to-follow-up-Patienten davon ausgegangen, dass sie zum Zeitpunkt der letzten Untersuchung ein Ereignis hatten.

[2]) Die Analyse entspricht der ursprünglichen Studienplanung vor dem 5. Amendment des Studienprotokolls.

Weitere Erläuterungen zu den Sensitivitätsanalysen finden sich in Abschnitt 9.8.3.3 des Studienberichts.

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Sofern nicht anders angegeben, war die FAS-Population Grundlage der Analyse.

HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*, PPS – *per protocol set*, 6-MWT – 6-Minuten-Gehtest

Wirksamkeit nach IMD-Kategorien

Derzeit wird davon ausgegangen, dass die Rezeptordichte des IP-Rezeptors zwischen einzelnen Patienten um bis zu dem Faktor 6 variieren kann. Diese Beobachtung wird auch als möglicher Grund für die unterschiedlichen individuellen Erhaltungsdosen (*individual maintenance dose*, IMD) diskutiert, da eine geringere Rezeptordichte eine größere Wirkstoffdosis notwendig macht, um einen ausreichenden therapeutischen Effekt zu erreichen. Das Anwendungskonzept von Selexipag sieht daher vor, die Dosierung so lange zu steigern, bis die patientenindividuell höchste Dosis erreicht wird, bei der spezifische Nebenwirkungen der Selexipag-Therapie gerade noch durch adäquate Gegenmaßnahmen behandelt werden können.

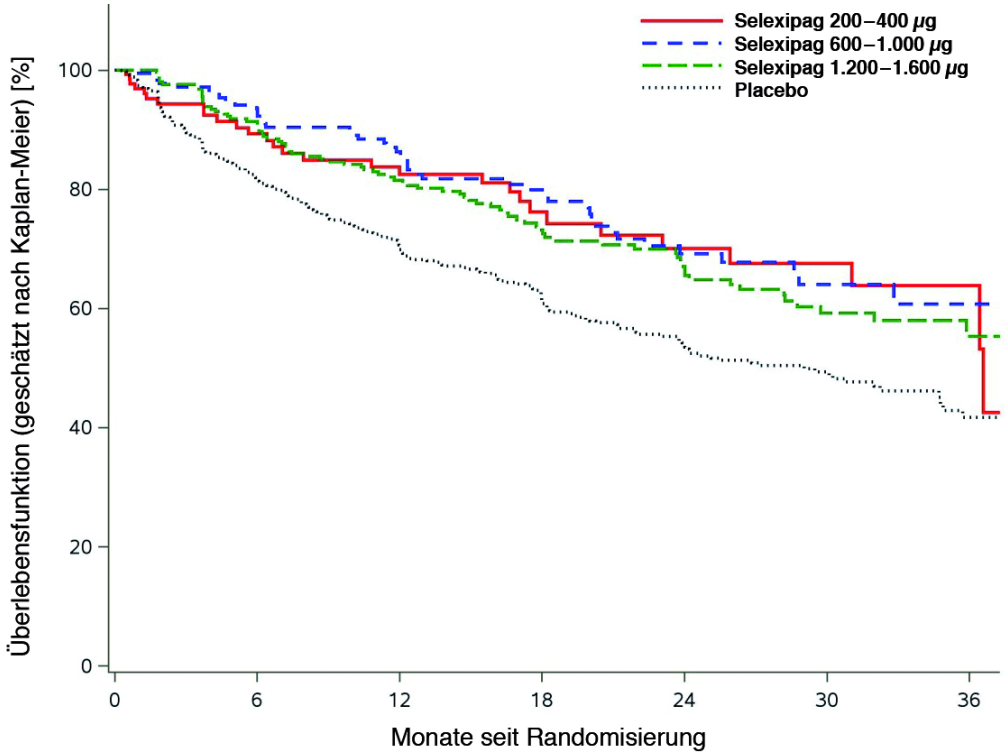
Zur Überprüfung dieses Therapiekonzepts wurde in einer supportiven Analyse die Wirksamkeit von Selexipag in Abhängigkeit von der IMD ausgewertet. Hierzu wurden die

Patienten entsprechend ihrer IMD in insgesamt 5 Kategorien aufgeteilt: 0[30], 200-400 µg Selexipag zweimal täglich, 600-1.000 µg Selexipag zweimal täglich, 1.200-1.600 µg Selexipag zweimal täglich und „andere“[31]. Anschließend wurde für jede der IMD-Kategorien (mit Ausnahme von 0 und „andere“) eine Ereigniszeitanalyse der Zeit bis zum ersten CEC-bestätigten M/M-Ereignis durchgeführt. Die Ergebnisse dieser Ereigniszeitanalyse (siehe Tabelle 4-21 und Abbildung 4-6) zeigen einen konsistenten Therapieeffekt über alle Wirkstärkekategorien und unterstreichen damit die Effektivität des gewählten Therapiekonzepts für Selexipag.

[30] Patienten hatten die initale Selexipag-Dosierung erhalten, hatten aber auf dieser Dosierung die Titrationsphase abgebrochen.
[31] Patienten wurden mit einem anderen Therapieregime als zweimal täglich behandelt.

Tabelle 4-21: Ergebnisse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ nach IMD-Kategorien

IMD-Kategorie	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [95 % KI], p-Wert (Log-Rang)
200-400 µg Selexipag, zweimal täglich	0,60 [0,41; 0,88] p=0,0038
600-1.000 µg Selexipag, zweimal täglich	0,53 [0,38; 0,72] p<0,0001
1.200-1.600 µg Selexipag, zweimal täglich	0,64 [0,49; 0,82] p=0,0002
In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Als Grundlage der Analyse diente die FAS-Population, jedoch wurden Patienten in der IMD-Kategorie „0“, d. h. solche Patienten, die die initale Selexipag-Dosierung erhalten hatten, aber auf dieser Dosierung in der Titrationsphase die Behandlung abbrachen, ausgeschlossen. Ebenso ausgeschlossen wurden Patienten, die mit einem anderen Therapieregime als zweimal täglich behandelt wurden („andere“), da sie durch die definierten IMD-Kategorien nicht erfasst werden konnten. Insgesamt wurden aufgrund dessen 16 Patienten in die Analyse nicht eingeschlossen. IMD – individuelle Erhaltungsdosis (*individual maintenance dose*), HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*	



Selexipag-IMD 200–400 µg-Patienten:							
unter Risiko	133	84	68	41	29	19	9
Ereignisse	0	12	17	22	25	26	27
Selexipag-IMD 600–1.000 µg-Patienten:							
unter Risiko	180	152	116	84	52	28	12
Ereignisse	0	11	22	30	40	43	44
Selexipag-IMD 1.200–1.600 µg-Patienten:							
unter Risiko	246	219	177	121	90	54	19
Ereignisse	0	21	44	60	69	78	80
Placebo-Patienten:							
unter Risiko	582	433	347	220	149	88	28
Ereignisse	0	101	158	197	221	232	241

Abbildung 4-6: Kaplan-Meier-Kurven für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“, aufgeteilt nach IMD-Kategorien
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen; ausgeschlossen wurden Patienten mit IMD=0 und „andere“. Analysierte Population: FAS. CEC – *Critical Event Committee*, IMD – individuelle Erhaltungsdosis (*individual maintenance dose*), FAS – *full analysis set*

Informationen zur Zensierung

Insgesamt wurden in GRIPHON 434 Patienten (75,6 %) im Selexipag-Arm und 370 Patienten (63,6 %) im Placebo-Arm für die Hauptanalyse des Endpunkts (unter Ausschluss der Ereignisse, die vor dem 16. August 2011 begannen) zensiert (vgl. Tabelle 4-22). Der Hauptgrund für die Zensierung war in beiden Studienarmen der Abschluss der Studie, ohne dass ein CEC-bestätigtes M/M-Ereignis aufgetreten war (Selexipag: 50,3 %, Placebo: 43,3 %). Weiterhin wurden 20,9 % der Patienten im Selexipag-Arm und 13,7 % der Patienten im

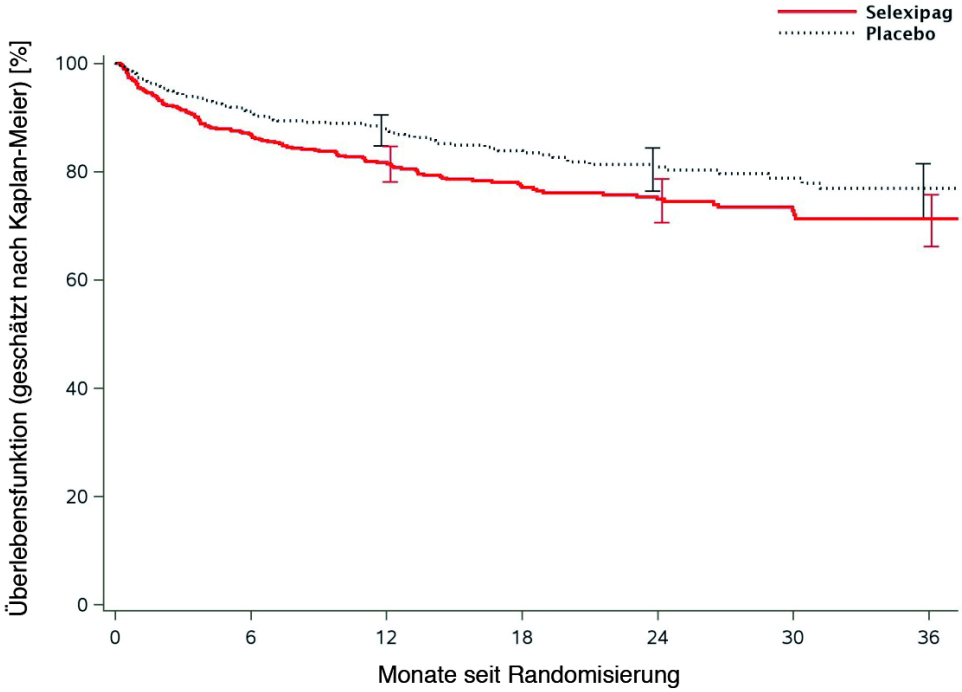
Placebo-Arm zensiert, da sie die Behandlung abgebrochen hatten, ohne dass ein CEC-bestätigtes M/M-Ereignis aufgetreten war. Weitere Gründe für die Zensierung waren der Rückzug der Einverständniserklärung ohne CEC-bestätigtes M/M-Ereignis (1,4 % bzw. 0,9 %) und Patienten, die ohne CEC-bestätigtes M/M-Ereignis *lost-to-follow-up* waren (0,3 % bzw. 0,5 %).

Tabelle 4-22: Gründe für die Zensierung in der Hauptanalyse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ (unter Ausschluss der Ereignisse, die vor dem 16. August 2011 begannen)

Gründe für die Zensierung				
Anzahl zensierte Patienten je Behandlungsgruppe				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Zensierte Beobachtungen, gesamt	434	75,6	370	63,6
Abschluss der Studie ohne CEC-bestätigtes M/M-Ereignis	289	50,3	252	43,3
Behandlung abgebrochen (andere Gründe) ohne CEC-bestätigtes M/M-Ereignis	120	20,9	80	13,7
Ereignis vor dem 16. August 2011	15	2,6	30	5,2
Einverständniserklärung zurückgezogen ohne CEC-bestätigtes M/M-Ereignis	8	1,4	5	0,9
Lost-to-follow-up ohne CEC-bestätigtes M/M-Ereignis	2	0,3	3	0,5
Als Grundlage der Analyse diente die FAS-Population. CEC – *Critical Event Committee*, FAS – *full analysis set*				

Als zusätzliche unterstützende Analyse für den primären Endpunkt wurde eine inverse Zensierung vorgenommen, d. h. Patienten, bei denen ein CEC-bestätigtes M/M-Ereignis aufgetreten war (einschließlich Ereignisse, die vor dem 16. August 2011 aufgetreten waren) und Patienten, die die Studie ohne CEC-bestätigtes M/M-Ereignis abgeschlossen hatten (Selexipag: 443 Patienten, Placebo: 494 Patienten) wurden zensiert, während Patienten, die die Behandlung dauerhaft ohne Ereignis abgebrochen hatten (Selexipag: 130 Patienten, Placebo: 88 Patienten), als Ereignis gewertet wurden.

Die Kaplan-Meier-Kurven dieser Analyse (siehe Abbildung 4-7) deutet auf ein höheres Risiko der Zensierung (d. h. des Drop-outs für den primären Endpunkt) in der Selexipag-Gruppe im Vergleich zur Placebo-Gruppe. Die Tatsache, dass sich die Kaplan-Meier-Kurven insbesondere in den ersten 6 Monaten aufteilen, anschließend aber weitestgehend konstant zueinander und relativ flach verlaufen, deutet darauf hin, dass insbesondere die aufgrund des Therapieschemas zu erwartende höhere Abbruchrate in der Titrationsphase zum höheren Risiko der Zensierung in der Selexipag-Gruppe führen könnte.



Selexipag-Patienten:

unter Risiko	574	455	361	246	171	101	40
Ereignisse	0	73	99	117	123	127	129
zensiert	0	46	114	211	280	346	405

Placebo-Patienten:

unter Risiko	582	433	347	220	149	88	28
Ereignisse	0	48	62	76	83	86	88
zensiert	0	101	173	286	350	408	466

Abbildung 4-7: Kaplan-Meier-Kurve der inversen Zensierung für „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ (d. h. Zeit bis zum Studienabbruch ohne M/M-Ereignis)
Für die inverse Zensierung wurden alle CEC-bestätigten M/M-Ereignisse sowie alle Patienten, die die Studie ohne CEC-bestätigtes M/M-Ereignis abschlossen, zensiert. Analysierte Population: FAS.
CEC – *Critical Event Committee*, FAS – *full analysis set*

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nicht-stratifizierten Analyse des primären Endpunkts mit Analysen verglichen, die nach geographischer Region, PAH-Ätiologie, WHO-/NYHA-Klasse am Basiswert, PAH-spezifischer Begleittherapie am Basiswert bzw. WHO-/NYHA-Klasse und PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. Die Effektschätzer des Behandlungseffekts, die in diesen Analysen bestimmt wurden, waren konsistent über alle stratifizierten Analysen sowie zur nicht-stratifizierten Analyse (siehe Tabelle 4-23). In allen Analysen war der Unterschied zwischen den Behandlungsgruppen signifikant, die Reduktion des relativen Risikos für das Auftreten eines M/M-Ereignis bewegte sich in einem Bereich von 39 bis 41 %.

Tabelle 4-23: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“

Modell	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Nicht-stratifiziertes Modell	0,60 [0,46; 0,78] p<0,0001
Modell stratifiziert nach Region[1]	0,59 [0,46; 0,78] p<0,0001
Modell stratifiziert nach PAH-Ätiologie[2]	0,61 [0,46; 0,79] p<0,0001
Modell stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I-II/III-IV)	0,61 [0,47; 0,80] p<0,0001
Modell stratifiziert nach PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)	0,60 [0,46; 0,79] p<0,0001
Modell stratifiziert nach WHO-/NYHA-Klasse und PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung	0,61 [0,47; 0,80] p<0,0001

[1]) Asien, Osteuropa, Lateinamerika, Nordamerika, Westeuropa, Australien

[2]) (Idiopathisch, angeboren, Drogen/Arzneimittel- oder Toxin-induziert), assoziiert mit HIV-Infektion, assoziiert mit Bindegewebserkrankung, assoziiert mit angeborenem Herzfehler mit korrigiertem Links-Rechts-Shunt.

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Als Grundlage der Analyse diente die FAS-Population.

HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*

Competing-Risk-Analyse und Analyse der Zeit bis zu den Einzelereignissen

Um ein besseres Verständnis über den Einfluss der einzelnen Komponenten auf den Behandlungseffekt im primären Endpunkt zu erreichen, wurde unterstützend eine Competing-Risk-Analyse nach der Methode von Gray mit den vier Komponenten „Tod“, „Hospitalisierung wegen PAH“, „Verschlechterung der PAH“ (enthält „Lungentransplantation oder atriale Septostomie“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie“) und „Fortschreiten der Erkrankung“ durchgeführt. Jede der 4 Komponenten konkurriert in der Behandlungsphase mit allen anderen, da das Auftreten von einer der Komponenten das Auftreten der anderen verhindert. Die Ergebnisse des Gray-Tests hierzu sind in Tabelle 4-24, die kumulative Inzidenzfunktion in Abbildung 4-8

dargestellt. In der Analyse zeigt sich, dass die Ergebnisse des primären Endpunkts vorwiegend durch die Verringerung des Risikos für das Auftreten einer „Hospitalisierung wegen PAH“ (p=0,0402) und des Risikos für das Auftreten eines „Fortschreitens der Erkrankung“ (p<0,0001) getrieben werden. In Bezug auf das Risiko für „Tod“ (p=0,0827) oder „Verschlechterung der PAH“ (p=0,5342) wurde kein signifikanter Unterschied beobachtet.

Tabelle 4-24: Competing-Risk-Analyse für den Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ (Methode nach Gray)

Competing-Risk-Analyse (Gray-Test)			
Intervention	Selexipag N=574		Placebo N=582
Behandlungsdifferenz			
	Chi-Quadrat	Freiheitsgrade	p
Selexipag vs. Placebo:			
Tod	3,0112	1	0,0827
Hospitalisierung wegen PAH	4,2076	1	0,0402
Verschlechterung der PAH	0,3864	1	0,5342
Fortschreiten der Erkrankung	27,8940	1	<0,0001
In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Grundlage der Analyse war die FAS-Population.			

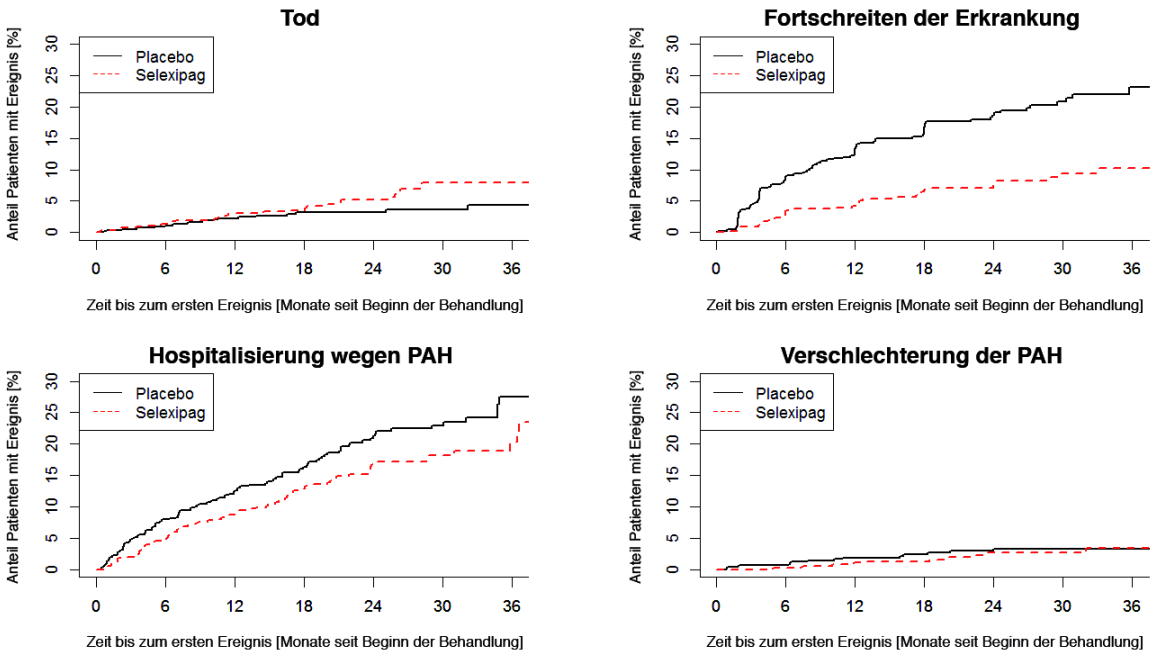


Abbildung 4-8: Kumulative Inzidenzfunktion für das erste Morbiditäts- oder Mortalitätsereignis bis zu EOT + 7 Tage (Competing-Risk-Analyse nach Gray)

CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. CEC – *Critical Event Committee*, FAS – *full analysis set*

Ergänzend zur Competing-Risk-Analyse wurde auf Wunsch der US Food and Drug Administration (FDA) zudem ein Analyse der Zeit bis zu den jeweiligen Einzelereignissen des primären Endpunkts der GRIPHON-Studie durchgeführt (vgl. Tabelle 4-25). Für diese Analyse wurden sämtliche Ereignisse von Studienbeginn bis EOT + 7 Tage berücksichtigt, d. h. unabhängig davon, ob ein anderes Ereignis des primären Endpunkts zuvor aufgetreten war oder ob ein zeitgleich aufgetretenes Ereignis als schwereres Ereignis im primären Endpunkt gezählt wurde.

Mit Ausnahme des Ereignisses „Tod“ war das relative Risiko für das Auftreten eines Ereignisses bei allen Komponenten des primären Endpunkts in der Selexipag-Gruppe zumindest numerisch niedriger als in der Placebo-Gruppe. Konsistent zu den Ergebnissen der Competing-Risk-Analyse zeigt auch die Anzahl der Zeit bis zu den Einzelereignissen, dass der primäre Endpunkt vorwiegend durch die Komponenten „Hospitalisierung wegen PAH“ (HR [99 % KI]: 0,65 [0,46; 0,94], p=0,0011) und „Fortschreiten der Erkrankung“ (HR [99 % KI]: 0,44 [0,30; 0,64], p<0,0001) getrieben wird.

Tabelle 4-25: Analyse der Zeit bis zu den Einzelereignissen des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“

Analyse der Zeit bis zu den Einzelereignissen						
	Anzahl Ereignisse je Behandlungsgruppe für den gesamten kombinierten Endpunkt und seine Einzelkomponenten (CEC-bestätigt, EOT)				Behandlungseffekt (Selexipag gegenüber Placebo)	
Intervention	Selexipag N=574		Placebo N=582		HR [99 % KI]	p
	n	%	n	%		
Morbiditäts-/ Mortalitätsereignis, gesamt	155	27,0	242	41,6	0,60 [0,46; 0,78]	<0,0001
Tod	46	8,0	37	6,4	1,17 [0,66; 2,07]	0,7656
Hospitalisierung wegen PAH	86	15,0	124	21,3	0,65 [0,46; 0,94]	0,0011
Lungentransplantation oder atriale Septo-stomie	2	0,3	4	0,7	0,48 [0,05; 4,49]	0,1943
Parenterale Prostanoid- oder dauerhafte Sauer-stofftherapie	33	5,7	49	8,4	0,63 [0,35; 1,13]	0,0205
Fortschreiten der Erkrankung	67	11,7	147	25,3	0,44 [0,30; 0,64]	<0,0001

In die Analyse wurden sämtliche Ereignisse eingeschlossen, die bis EOT + 7 Tage aufgetreten waren, d. h. unabhängig davon, ob ein anderes Ereignis des primären Endpunkts früher aufgetreten war oder das andere Ereignis zeitgleich aufgetreten war, aber als schweres Ereignis ausschlaggebend für die Wertung im primären Endpunkt war. Aufgrund dessen konnte je Patient mehr als ein Ereignis in der Auswertung berücksichtigt werden, sodass sich die Anzahl der Einzelereignisse nicht zur Gesamtzahl der Ereignisse des primären Endpunkts summiert.

Es wurden sämtliche CEC-bestätigten Ereignisse berücksichtigt, d. h. auch die Ereignisse, die vor dem 16. August 2011 begannen. Grundlage der Analyse war die FAS-Population. Die Analysen wurde *post hoc* durchgeführt.

HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, EOT – *end-of-treatment*, FAS – *full analysis set*

Naive Raten

Zusätzlich zur Ereigniszeitanalyse wurde die Anzahl der Ereignisse im Untersuchungszeitraum (die „Gesamtraten“), wie vom G-BA im Beratungsgespräch vom 13.05.2015 gefordert [3], als naive Raten, d. h. in Form von Risk Ratios (RR), Odds Ratios (OR) und absoluter Risikodifferenz (RD) samt zugehörigen 95 %-Konfidenzintervallen ausgewertet (die Ergebnisse dieser Analyse sind in Tabelle 4-26 dargestellt). Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt eine solche Auswertung jedoch voraus, dass alle Patienten zu einem fixen Endpunkt ausgewertet werden [83]. Hier findet jedoch eine Auswertung des Zeitraumes bis zum Auftreten eines Ereignisses des primären Endpunktes, d. h. zum Zeitpunkt des patientenindividuellen Endes des Behandlungszeitraums (EOT) statt

und damit eben nicht zu einem festgelegten und für alle Patienten gleichen Zeitpunkt. Für die Post-hoc-Analyse der naiven Raten ist daher abweichend von der Hauptanalyse von einem hohen Verzerrungspotential auszugehen.

Wie bereits zuvor beschrieben, traten im Zeitraum von Randomisierung der Patienten bis EOT + 7 Tage insgesamt 155 Ereignisse im Selexipag-Arm auf und damit bedeutend weniger als im Placebo-Arm (242 Ereignisse). Dies entspricht einer signifikanten Reduktion des relativen Risikos von 35 % (RR [95 % KI]: 0,65 [0,55; 0,77]) sowie einer signifikanten Reduktion des absoluten Risikos von 14,6 Prozentpunkten (RD [95 % KI]: −14,6 %-P. [−20,0 %-P.; −9,2 %-P.]). Die Auswertung der naiven Raten stimmt damit vollständig mit der Aussage der Hauptanalyse überein.

Das in beiden Gruppen am häufigsten aufgetretene erste Ereignis in der Studie war die „Hospitalisierung wegen PAH“, die bei 13,6 % der Patienten im Selexipag-Arm und bei 18,7 % der Patienten im Placebo-Arm aufgetreten war. Der Unterschied zwischen den Behandlungsgruppen war sowohl für die relative Risikoreduktion (RR [95 % KI]: 0,73 [0,56; 0,95]) als auch für die absolute Risikoreduktion (RD [95 % KI]: −5,1 %-P. [−9,4 %-P.; −0,9 %-P.]) statistisch signifikant.

Ebenso wie für „Hospitalisierung wegen PAH“ war der Unterschied zwischen den Behandlungsgruppen auch für Patienten mit „Fortschreiten der Erkrankung“ als erstem Ereignis statistisch signifikant (RR [95 % KI]: 0,39 [0,27; 0,55]; RD [95 % KI]: −10,6 %-P. [−14,2 %-P.; −6,9 %-P.]).

Das Ereignis „Tod“ trat in beiden Behandlungsarmen nur bei einer vergleichsweise geringen Anzahl von Patienten als erstes Ereignis auf (Selexipag: 4,9 %, Placebo: 3,1 %), die Häufigkeiten sind zudem statistisch nicht signifikant verschieden. Die Ereignisse „Lungentransplantation oder atriale Septostomie“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie“ traten jeweils nur sehr selten auf und lassen allein stehend nur eine sehr limitierte Aussage zum Behandlungseffekt zu.

Tabelle 4-26: Analyse der naiven Raten für den Endpunkt „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ (einschließlich Einzelkomponenten)

Naive Raten				
Anzahl Ereignisse je Behandlungsgruppe für den gesamten kombinierten Endpunkt und seine Einzelkomponenten (Naive Raten, CEC-bestätigt, EOT)				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Morbiditäts-/ Mortalitätsereignis, gesamt	155	27,0	242	41,6
Tod	28	4,9	18	3,1
Hospitalisierung wegen PAH	78	13,6	109	18,7
Lungentransplantation oder atriale Septostomie	1	0,2	2	0,3
Parenterale Prostanoid- oder dauerhafte Sauer-stofftherapie	10	1,7	13	2,2
Fortschreiten der Erkrankung[1]	38	6,6	100	17,2
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Morbiditäts-/ Mortalitätsereignis, gesamt	0,65 [0,55; 0,77]	0,52 [0,41; 0,67]	-14,6 [-20,0; -9,2]	<0,0001
Tod	1,58 [0,88; 2,82]	1,61 [0,88; 2,94]	1,8 [-0,5; 4,0]	0,1289
Hospitalisierung wegen PAH	0,73 [0,56; 0,95]	0,68 [0,50; 0,94]	-5,1 [-9,4; -0,9]	0,0179
Lungentransplantation oder atriale Septostomie	0,51 [0,05; 5,58]	0,51 [0,05; 5,60]	-0,2 [-0,8; 0,4]	0,6831
Parenterale Prostanoid- oder dauerhafte Sauer-stofftherapie	0,78 [0,35; 1,76]	0,78 [0,34; 1,78]	-0,5 [-2,1; 1,1]	0,5806
Fortschreiten der Erkrankung[1]	0,39 [0,27; 0,55]	0,34 [0,23; 0,51]	-10,6 [-14,2; -6,9]	<0,0001

[1]) Die weitere Aufteilung dieser Einzelkomponente ist in der folgenden Tabelle 4-27 aufgeführt.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Wenn mehr als ein Ereignis mit dem gleichen Anfangsdatum aufgetreten war, dann wurde das schwerere Ereignis als erstes Ereignis ausgewertet. Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, CEC – *Critical Event Committee*, EOT – *end-of-treatment*, FAS – *full analysis set*

Tabelle 4-27 zeigt die weitere Aufteilung der Komponente „Fortschreiten der Erkrankung“ in die beiden Kriterien, die zum Eintritt der Komponente und damit zur Zählung im primären Endpunkt führen konnten. Die Verschlechterung der 6-MWD um mindestens 15 % im Vergleich zum Basiswert war unabhängig von der WHO-/NYHA-Klasse bei beiden Kriterien Voraussetzung für die Wertung im primären Endpunkt. Während bei Patienten in WHO-/NYHA-Klasse II zusätzlich zur Verschlechterung der 6-MWD auch noch eine Verschlechterung ihrer Klasse beobachtet werden musste, musste bei Patienten in WHO-/NYHA-Klasse IV für die Wertung im primären Endpunkt die Notwendigkeit für zusätzliche PAH-Therapie bestehen. Bei Patienten in WHO-/NYHA-Klasse III konnten beide genannten zusätzlichen Ereignisse zu einer Zählung im primären Endpunkt führen. Da alle drei Ereignisse auch gleichzeitig stattfinden konnten, war bei Patienten in der WHO-/NYHA-Klasse III eine Wertung im primären Endpunkt nach beiden Kriterien der Komponente möglich. Im Rahmen der primären Auswertung des Endpunkts spielt diese Tatsache keine Rolle, aber bei einer getrennten Auswertung der beiden Kriterien, wie in Tabelle 4-27 dargestellt, hat dies eine Doppelzählung von Patienten zur Folge.

In beiden Kriterien, die zur Zählung der Komponente „Fortschreiten der Erkrankung“ führen konnten, zeigt sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von Selexipag. Dies ist sowohl bei der Auswertung des relativen Risikos (RR [95 % KI]: 0,37 [0,23; 0,60] bzw. 0,51 [0,33; 0,80]) als auch bei der Auswertung der absoluten Risikodifferenz der Fall (RD [95 % KI]: −6,8 %-P. [−9,9 %-P.; −3,8 %-P.]). Da jeweils beide Einzelereignisse (Verschlechterung der 6-MWD ≥15 % und Verschlechterung der WHO-/NYHA-Klasse bzw. Notwendigkeit neuer PAH-Therapie) gleichzeitig eintreten mussten, damit eines der beiden Kriterien als „Fortschreiten der Erkrankung“ gezählt wurde, unterscheiden sich die Raten der Einzelereignisse nicht von den Raten der beiden Kriterien.

Tabelle 4-27: Analyse der naiven Raten der Einzelbestandteile der Komponente „Fortschreiten der Erkrankung“ des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“

Naive Raten				
Anzahl Ereignisse je Behandlungsgruppe für die Einzelbestandteile der Komponente „Fortschreiten der Erkrankung“ (Naive Raten, CEC-bestätigt, EOT)				
Intervention	Selexipag N=567		Placebo N=569	
	n	%	n	%
Patienten mit Ereignis in WHO-/NYHA-Klasse II/III zum Basiswert	23	4,1	62	10,9
Verschlechterung der 6-MWD ≥15 %[1]	23	4,1	62	10,9
Verschlechterung der WHO-/NYHA-Klasse	23	4,1	62	10,9
Intervention	Selexipag N=296		Placebo N=322	
	n	%	n	%
Patienten mit Ereignis in WHO-/NYHA-Klasse III/IV zum Basiswert	25	8,4	53	16,5
Verschlechterung der 6-MWD ≥15 %[1]	25	84	53	16,5
Notwendigkeit zusätzlicher PAH-Therapie	25	8,4	53	16,5
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit Ereignis in WHO-/NYHA-Klasse II/III zum Basiswert	0,37 [0,23; 0,60]	0,35 [0,21; 0,57]	-6,8 [-9,9; -3,8]	<0,0001
Verschlechterung der 6-MWD ≥15 %[1]	0,37 [0,23; 0,60]	0,35 [0,21; 0,57]	-6,8 [-9,9; -3,8]	<0,0001
Verschlechterung der WHO-/NYHA-Klasse	0,37 [0,23; 0,60]	0,35 [0,21; 0,57]	-6,8 [-9,9; -3,8]	<0,0001

Patienten mit Ereignis in WHO-/NYHA-Klasse III/IV zum Basiswert	0,51 [0,33; 0,80]	0,47 [0,28; 0,78]	-8,0 [-13,2; -2,9]	0,0028
Verschlechterung der 6-MWD ≥15 %[1]	0,51 [0,33; 0,80]	0,47 [0,28; 0,78]	-8,0 [-13,2; -2,9]	0,0028
Notwendigkeit zusätzlicher PAH-Therapie	0,51 [0,33; 0,80]	0,47 [0,28; 0,78]	-8,0 [-13,2; -2,9]	0,0028

[1]) Es wird davon ausgegangen, dass die Verschlechterung vom CEC bestätigt wurde.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, CEC – *Critical Event Committee*, EOT – *end-of-treatment*, FAS – *full analysis set*

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Zusätzlich zur primären Analyse des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ bis EOT wurde als explorative Analyse der Endpunkt mit den gleichen Komponenten, jedoch bis Studienschluss ausgewertet. Beide Analysen unterscheiden sich durch den Umgang mit Patienten, die ohne CEC-bestätigtes M/M-Ereignis die Behandlungsphase beendet hatten. In der Analyse der Zeit bis zum ersten CEC-bestätigten M/M-Ereignis bis EOT wurden diese Patienten zum individuellen Ende ihrer Behandlungsphase zensiert, während sie in der Analyse bis Studienschluss auch über diesen Zeitpunkt hinaus weiterbeobachtet wurden und somit weiterhin ein M/M-Ereignis bei ihnen auftreten konnte. Die Ergebnisse der Analyse sind in Tabelle 4-28 dargestellt.

In der Auswertung der „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ bis Studienschluss zeigte sich ein deutlicher und statistisch signifikanter Unterschied zwischen Selexipag und Placebo (HR [95 % KI]: 0,65 [0,54; 0,79], p<0,0001). Der Unterschied zwischen den Behandlungsgruppen entspricht einer relativen Risikoreduktion für das Auftreten eines CEC-bestätigten M/M-Ereignisses von 35 %.

Neben der Analyse aller Ereignisse des primären Endpunkts bis Studienschluss wurden zwei weitere Analysen durchgeführt, in denen die Ereignisse der Komponente „Fortschreiten der Erkrankung“ bzw. der Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“ aus der Analyse ausgeschlossen wurden. Die Ergebnisse dieser Analysen sind ebenfalls in Tabelle 4-28 dargestellt.

Ähnlich wie auch bereits in der Analyse des Endpunkts bis EOT verringert sich die relative Risikoreduktion durch den Ausschluss der Komponente „Fortschreiten der Erkrankung“ leicht (von 35 % auf 19 %), der Unterschied zwischen den Behandlungsgruppen ist jedoch weiterhin statistisch signifikant (p=0,0189). Der zusätzliche Ausschluss der Komponente „Initiierung

einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie“ hat nur einen geringfügigen Einfluss auf das Ergebnis der Analyse. Auch hier ist jedoch der Unterschied zwischen den Behandlungsgruppen weiterhin statistisch signifikant.

Tabelle 4-28: Analyse des Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

Alternative Analyse	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [95 % KI], p-Wert (Log-Rang)
Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis bis zum Studienschluss	
Unter Einschluss aller Komponenten.	0,65 [0,54; 0,79] p<0,0001
Unter Ausschluss der Komponente „Fortschreiten der Erkrankung“.	0,81 [0,66; 0,99] p=0,0189
Unter Ausschluss der Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstoff-therapie aufgrund einer Verschlechterung der PAH“.	0,82 [0,66; 1,01] p=0,0322
In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Grundlage der Analyse war die FAS-Population. EOT – *end-of-treatment*, HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*, 6-MWT – 6-Minuten-Gehtest	

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-20). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.2 Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-29: Operationalisierung von „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH. Es handelt sich um einen kombinierten Endpunkt, als Morbiditäts- oder Mortalitätsereignis wurden die folgenden Ereignisse gezählt: • Tod wegen PAH • Hospitalisierung aufgrund einer Verschlechterung der PAH[1] Beide Ereignisse dieses kombinierten Endpunkts wurden durch ein verblindetes, unabhängiges Critical Event Committee (CEC) begutachtet. Das CEC bestand aus 3 klinischen Experten mit Erfahrung in der PAH und in der Überwachung klinischer Studien. Als Mitglied des CEC war es nicht möglich, gleichzeitig als Prüfarzt an der Studie beteiligt zu sein. Die Hauptaufgabe des CEC war die Bestätigung des Auftretens eines Morbiditäts- oder Mortalitätsereignisses, gleichzeitig sollten aber auch alle Todesfälle auf einen möglichen Zusammenhang zur PAH überprüft werden. Die Hauptanalyse des Endpunkts wurde für den Zeitraum von der Randomisierung bis 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase (*end-of-treatment*, EOT) vorgenommen. Auswertungspopulation für die Hauptanalyse war die FAS-Population. *Zensierungsregeln/Ersetzungsregeln* Es wurde keine Imputationsmethode angewendet. Bei Patienten, bei denen bis zu 7 Tage nach der letzten Einnahme der Studienmedikation kein CEC-bestätigter Todesfall und keine CEC-bestätigte Hospitalisierung aufgrund einer Verschlechterung der PAH eingetreten war, wurde die Zeit bis zum Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH nach folgenden Regeln zensiert: • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und innerhalb der Behandlungsphase + 7 Tage aus einem anderen Grund als der PAH verstorben waren: Zeitpunkt des Todes • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und keine Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302/GRIPHON-Behandlungsphase plus 7 Tage, Zeitpunkt der EOS-Untersuchung, Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und ihre Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302/GRIPHON-Behandlungsphase plus 7 Tage, Zeitpunkt des letzten Kontakts, 27.04.2014) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die keine Studienmedikation erhalten hatten: Minimum (Maximum [Zeitpunkt der EOS-Untersuchung, Zeitpunkt der Randomisierung], Zeitpunkt des letzten Kontakts, 27.04.2014) minus Zeitpunkt der Randomisierung plus 1 Tag. Es handelt sich hierbei um (potentiell informative) rechts-zensierte Beobachtungen.
[1]) Jeglicher unfreiwillige Krankenhausaufenthalt (≥24 Stunden) aufgrund einer Verschlechterung der PAH; Verschlechterung der PAH schließt Anzeichen und Symptome des Rechtsherzversagens (z. B. (Prä-)Synkope, Zyanose, Verschlechterung der Kurzatmigkeit, klinisch relevante Verschlechterung der körperlichen Leistungsfähigkeit, Abfall der arteriellen Sauerstoffsättigung, verschlimmerte Ödeme, Hepatomegalie und Ascites) mit ein.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-30: Bewertung des Verzerrungspotentials für „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunktherheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-31: Ergebnisse für „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**	
Intervention	Selexipag N=574	Placebo N=582
Zeit bis zum ersten Ereignis (Monate)		
1. Quartil [95 % KI]	28,6 [21,9; 36,6]	19,4 [16,2; 24,0]
Median [95 % KI]	- 	- [38,5; -]
3. Quartil [95 % KI]	-	-
Kaplan-Meier-Schätzer ereignisfreier Patienten (%)		
Monat 6		
Patienten unter Risiko	457	437
Zensierte Patienten	85	89
Patienten mit Ereignis	32	56
KM-Schätzer, % [95 % KI]	93,9 [91,4; 95,6]	89,6 [86,6; 91,9]
Monat 12		
Patienten unter Risiko	364	351
Zensierte Patienten	152	145
Patienten mit Ereignis	58	86
KM-Schätzer, % [95 % KI]	88,3 [85,1; 90,8]	83,0 [79,4; 86,0]
Monat 18		
Patienten unter Risiko	250	227
Zensierte Patienten	248	249
Patienten mit Ereignis	76	106
KM-Schätzer, % [95 % KI]	83,2 [79,3; 86,4]	77,2 [73,0; 80,9]
Monat 24		
Patienten unter Risiko	172	152
Zensierte Patienten	311	307
Patienten mit Ereignis	91	123
KM-Schätzer, % [95 % KI]	77,4 [72,7; 81,4]	70,7 [65,7; 75,1]
Monat 30		
Patienten unter Risiko	102	89
Zensierte Patienten	374	363
Patienten mit Ereignis	98	130
KM-Schätzer, % [95 % KI]	73,6 [68,2; 78,2]	67,0 [61,4; 71,9]

Monat 36		
Patienten unter Risiko	40	28
Zensierte Patienten	434	418
Patienten mit Ereignis	100	136
KM-Schätzer, % [95 % KI]	71,1 [64,6; 76,6]	58,7 [50,4; 66,1]
Monat 42		
Patienten unter Risiko	7	4
Zensierte Patienten	465	441
Patienten mit Ereignis	102	137
KM-Schätzer, % [95 % KI]	67,2 [58,9; 74,3]	55,3 [44,8; 64,5]
Monat 48		
Patienten unter Risiko	1	-
Zensierte Patienten	471	445
Patienten mit Ereignis	102	137
KM-Schätzer, % [95 % KI]	67,2 [58,9; 74,3]	55,3 [44,8; 64,5]
Behandlungseffekt (Selexipag gegenüber Placebo)		
HR [99 % KI]		0,70 [0,50; 0,98]
p (Log-Rang-Test)		0,0031

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen.
KI – Konfidenzintervall, KM-Schätzer – Kaplan-Meier-Schätzer, HR – Hazard Ratio, FAS – *full analysis set*

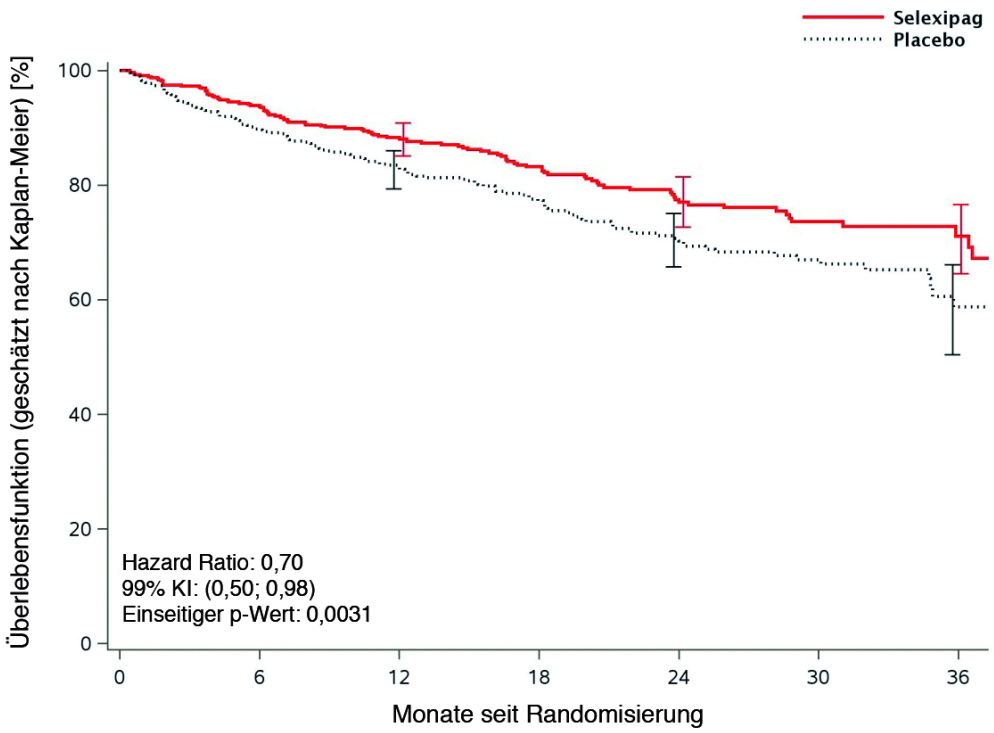
Hauptanalyse des Endpunkts

Die Analyse der Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH wurde auf Grundlage der FAS-Population durchgeführt.

Insgesamt trat im Beobachtungszeitraum (EOT + 7 Tage) bei 102 Patienten in der Selexipag-Gruppe und bei 137 Patienten in der Placebo-Gruppe ein Ereignis auf, d. h. entweder sie verstarben wegen PAH oder wurden aufgrund einer Verschlechterung der PAH stationär aufgenommen. Die HR [99 % KI] in der Ereigniszeitanalyse eines Todesfalles wegen PAH oder einer Hospitalisierung aufgrund einer Verschlechterung der PAH betrug 0,70 [0,50; 0,98] und war statistisch signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p<0,0031). Dies entspricht einer Reduktion des relativen Risikos von 30 %.

Die mediane Zeit bis zum ersten Ereignis wurde in beiden Armen im Behandlungszeitraum nicht erreicht, das erste Quartil [95 % KI] war im Selexipag-Arm nach 28,6 Wochen [21,9 Wochen; 36,6 Wochen], im Placebo-Arm hingegen bereits nach 19,4 Wochen [16,2 Wochen; 24,0 Wochen] erreicht.

Abbildung 4-9 zeigt die Kaplan-Meier-Kurve des Endpunkts in der FAS-Population.



Selexipag-Patienten:							
unter Risiko	574	457	364	250	172	102	40
Ereignisse	0	32	58	76	91	98	100
zensiert	0	85	152	248	311	374	434

Placebo-Patienten:							
unter Risiko	582	437	351	227	152	89	28
Ereignisse	0	56	86	106	123	130	136
zensiert	0	89	145	249	307	363	418

Abbildung 4-9: Kaplan-Meier-Kurve für „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Fehlerbalken geben das 95 % KI an. CEC – *Critical Event* Committee, FAS – *full analysis set*, KI – Konfidenzintervall

Nebenanalysen des Endpunkts

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf den Zeitraum von der Randomisierung bis EOT + 7 Tage.

Sensitivitätsanalysen

Um die Robustheit der Ergebnisse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung

der PAH“ einschätzen zu können, wurden verschiedene Sensitivitätsanalysen durchgeführt. Diese sind in der folgenden Tabelle 4-32 dargestellt.

Wie auch in der FAS-Population war der Unterschied zwischen den Behandlungsgruppen in der PPS-Population statistisch signifikant (p=0,0016), der Effektschätzer des Behandlungseffekts in der PPS-Population (HR [99 % KI]: 0,68 [0,48; 0,95]) war konsistent zum Effektschätzer in der FAS-Population (HR [99 % KI]: 0,70 [0,50; 0,98]). Die relative Risikoreduktion für das Auftreten eines Todesfalles wegen PAH oder einer Hospitalisierung aufgrund einer Verschlechterung der PAH betrug in dieser Sensitivitätsanalyse 32 %.

Wurden nur Patienten in die Auswertung des Endpunkts eingeschlossen, die von den Prüfärzten als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden, so ergab sich eine HR [99 % KI] von 0,70 [0,50; 0,99] und damit eine relative Risikoreduktion für das Auftreten eines Todesfalles wegen PAH oder einer Hospitalisierung aufgrund einer Verschlechterung der PAH von 30 %. Der Unterschied zwischen den Behandlungsgruppen war auch in dieser Sensitivitätsanalyse statistisch signifikant (p<0,0036).

Insgesamt gesehen bestätigen sämtliche der durchgeführten Sensitivitätsanalysen das Ergebnis der Hauptanalyse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ und unterstreichen damit die Robustheit der Ergebnisse dieses Endpunkts.

Tabelle 4-32: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“

Sensitivitätsanalyse	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Per-protocol-Analyse	
Analyse in der PPS-Population	0,68 [0,48; 0,95] p=0,0016
Weitere Sensitivitätsanalysen	
Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden.	0,70 [0,50; 0,99] p=0,0036
Weitere Erläuterungen zu den Sensitivitätsanalysen finden sich in Abschnitt 9.8.4.1.3 des Studienberichts. In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Sofern nicht anders angegeben, war die FAS-Population Grundlage der Analyse. HR – Hazard Ratio, KI – Konfidenzintervall, FAS – *full analysis set*, PPS – *per protocol set*	

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nicht-stratifizierten Analyse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ mit Analysen verglichen, die nach geographischer Region, PAH-Ätiologie, WHO-/NYHA-Klasse am Basiswert, PAH-spezifischer Begleittherapie am Basiswert bzw. WHO-/NYHA-Klasse und PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. Die Effektschätzer des Behandlungseffekts, die in diesen Analysen bestimmt wurden, waren konsistent über alle stratifizierten Analysen sowie zur nicht-stratifizierten Analyse (siehe Tabelle 4-33). In allen Analysen war der Unterschied zwischen den Behandlungsgruppen signifikant, die Reduktion des relativen Risikos für das Auftreten eines M/M-Ereignis bewegte sich in einem Bereich von 29 bis 31 %.

Tabelle 4-33: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“

Modell	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Nicht-stratifiziertes Modell	0,70 [0,50; 0,98] p=0,0031
Modell stratifiziert nach Region[1]	0,69 [0,49; 0,97] p=0,0024
Modell stratifiziert nach PAH-Ätiologie[2]	0,70 [0,50; 0,99] p=0,0035
Modell stratifiziert nach WHO-/ NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I-II/III-IV)	0,71 [0,51; 1,00] p=0,0043
Modell stratifiziert nach PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)	0,70 [0,50; 0,98] p=0,0033
Modell stratifiziert nach WHO-/ NYHA-Klasse und PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung	0,71 [0,51; 1,00] p=0,0048

[1]) Asien, Osteuropa, Lateinamerika, Nordamerika, Westeuropa, Australien

[2]) (Idiopathisch, angeboren, Drogen/Arzneimittel- oder Toxin-induziert), assoziiert mit HIV-Infektion, assoziiert mit Bindegewebserkrankung, assoziiert mit angeborenem Herzfehler mit korrigiertem Links-Rechts-Shunt.

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Als Grundlage der Analyse diente die FAS-Population.

HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*

Naive Raten

Zusätzlich zur Ereigniszeitanalyse wurde die Anzahl der Ereignisse im Untersuchungszeitraum (die „Gesamtraten“), wie vom G-BA im Beratungsgespräch vom 13.05.2015 gefordert [3], als naive Raten, d. h. in Form von Risk Ratios (RR), Odds Ratios (OR) und absoluter Risikodifferenz (RD) samt zugehörigen 95 %-Konfidenzintervallen ausgewertet (die Ergebnisse dieser Analyse sind in Tabelle 4-34 dargestellt). Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt eine solche Auswertung jedoch voraus, dass alle Patienten zu einem fixen Endpunkt ausgewertet werden [83]. Hier findet jedoch eine Auswertung der Zeit bis zum Auftreten des ersten Ereignisses innerhalb

eines Zeitraums von Randomisierung bis zum Zeitpunkt des patientenindividuellen Endes des Behandlungszeitraums (EOT) statt und damit eben nicht zu einem festgelegten und für alle Patienten gleichen Zeitpunkt. Für die Post-hoc-Analyse der naiven Raten ist daher abweichend von der Hauptanalyse von einem hohen Verzerrungspotential auszugehen.

Wie bereits zuvor beschrieben, traten im Zeitraum von Randomisierung der Patienten bis EOT + 7 Tage insgesamt 102 Ereignisse im Selexipag-Arm auf und damit weniger als im Placebo-Arm (137 Ereignisse). Dies entspricht einer signifikanten Reduktion des relativen Risikos von 34 % (RR [95 % KI]: 0,76 [0,60; 0,95]) sowie einer signifikanten Reduktion des absoluten Risikos um 5,8 Prozentpunkte (RD [95 % KI]: −5,8 %-P. [−10,4 %-P.; −1,1 %-P.]). Die Auswertung der naiven Raten stimmt damit vollständig mit der Aussage der Hauptanalyse überein.

Das in beiden Gruppen am häufigsten aufgetretene erste Ereignis dieses kombinierten Endpunkts war die „Hospitalisierung wegen PAH“, die bei 15,0 % der Patienten im Selexipag-Arm und bei 21,1 % der Patienten im Placebo-Arm aufgetreten war. Der Unterschied zwischen den Behandlungsgruppen war sowohl für die relative Risikoreduktion (RR [95 % KI]: 0,71 [0,55; 0,91]) als auch für die absolute Risikoreduktion (RD [95 % KI]: −6,2 %-P. [−10,6 %-P.; −1,7 %-P.]) statistisch signifikant.

Das Ereignis „Tod“ trat in beiden Behandlungsarmen nur bei einer geringen Anzahl von Patienten als erstes Ereignis auf (Selexipag: 2,8 %, Placebo: 2,4 %), die Häufigkeiten sind zudem statistisch nicht signifikant verschieden.

Tabelle 4-34: Analyse der naiven Raten für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ (einschließlich Einzelkomponenten)

Naive Raten				
Anzahl Ereignisse je Behandlungsgruppe für den gesamten kombinierten Endpunkt und seine Einzelkomponenten (Naive Raten, CEC-bestätigt, EOT)				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Gesamt	102	17,8	137	23,5
Todesfall wegen PAH	16	2,8	14	2,4
Hospitalisierung wegen PAH	86	15,0	123	21,1
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Gesamt	0,76 [0,60; 0,95]	0,70 [0,53; 0,94]	-5,8 [-10,4; -1,1]	0,0156
Todesfall wegen PAH	1,16 [0,57; 2,35]	1,16 [0,56; 2,41]	0,4 [-1,5; 2,2]	0,7534
Hospitalisierung wegen PAH	0,71 [0,55; 0,91]	0,66 [0,49; 0,89]	-6,2 [-10,6; -1,7]	0,0067

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, CEC – *Critical Event Committee*, EOT – *end-of-treatment*, FAS – *full analysis set*

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Neben der Analyse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ bis zum Ende der Behandlungsphase (EOT) wurde auch eine Analyse der Ereignisse bis zum Studienschluss durchgeführt. Die Ergebnisse dieser Analyse sind in Tabelle 4-35 dargestellt.

In der Auswertung des Endpunkts bis Studienschluss zeigte sich ein deutlicher und statistisch signifikanter Unterschied zwischen Selexipag und Placebo (HR [95 % KI]: 0,76 [0,60; 0,95], p<0,0081). Der Unterschied zwischen den Behandlungsgruppen entspricht einer relativen Risikoreduktion für das Auftreten eines CEC-bestätigten Todesfalles wegen PAH oder einer Hospitalisierung aufgrund einer Verschlechterung der PAH von 24 %.

Tabelle 4-35: Analyse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

Alternative Analyse	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [95 % KI], p-Wert (Log-Rang)
Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH bis zum Studienschluss	0,76 [0,60; 0,95] p=0,0081
In die Analyse wurden auch die CEC-bestätigten Ereignisse miteingeschlossen, die vor dem 16. August 2011 begannen. Grundlage der Analyse war die FAS-Population. HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*	

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-21). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.3 Zeit bis zum Tod (Studienschluss)

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-36: Operationalisierung von „Zeit bis zum Tod (Studienschluss)“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Zeit bis zum Tod jeglicher Ursache bis zum Studienschluss. Die Analyse des Endpunkts wurde für den Zeitraum von der Randomisierung bis zum Studienschluss vorgenommen. Auswertungspopulation für die Hauptanalyse war die FAS-Population.
	Zensierungsregeln/Ersetzungsregeln Es wurde keine Imputationsmethode angewendet. Patienten, bei denen nicht bekannt war, ob sie verstorben waren, wurden anhand der folgenden Regeln in der Ereigniszeitanalyse zensiert: • Patienten, die keine Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase oder zur Nachverfolgung der Sterblichkeit am Studienende gegeben hatten: Minimum (Zeitpunkt der EOS-Untersuchung, Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die keine Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten, aber für die Nachverfolgung der Sterblichkeit am Studienende: Minimum (Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die ihre Zustimmung zur Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Maximum [Zeitpunkt der Abschlussuntersuchung der PTOP, Zeitpunkt der EOS-Untersuchung], Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. Es handelt sich hierbei um rechts-zensierte Beobachtungen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-37: Bewertung des Verzerrungspotentials für „Zeit bis zum Tod (Studienschluss)“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotentiala uf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes konnten nicht beobachtet werden, jedoch besteht ein bedeutendes Verzerrungspotential durch den erlaubten Therapiewechsel am Ende der verblindeten Behandlungsphase (EOT). Patienten im Placebo-Arm, bei denen ein nicht-tödliches M/M-Ereignis aufgetreten war, konnten anschließend die verblindete Behandlung beenden und stattdessen mit Selexipag oder einem anderen PAH-spezifischen Arzneimittel weiterbehandelt werden. Insgesamt 155 Patienten aus dem Placebo-Arm wechselten nach Ende der Behandlungsphase die Therapie und wurden in offener Form mit Selexipag behandelt. Tendenziell ist aufgrund dessen jedoch nur mit einer Verzerrung des Therapieeffekts zuungunsten von Selexipag zu rechnen.

Das Verzerrungspotential für diesen Endpunkt ist insgesamt somit als hoch anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zum Tod (Studienschluss)“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-38: Ergebnisse für „Zeit bis zum Tod (Studienschluss)“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**	
Intervention	Selexipag N=574	Placebo N=582
Zeit bis zum ersten Ereignis (Monate)		
1. Quartil [95 % KI]	38,2 [28,7; -]	- [28,1; -]
Median [95 % KI]	-	-
3. Quartil [95 % KI]	-	-
Kaplan-Meier-Schätzer ereignisfreier Patienten (%)		
Monat 6		
Patienten unter Risiko	543	546
Zensierte Patienten	10	14
Patienten mit Ereignis	21	22
KM-Schätzer, % [95 % KI]	96,3 [94,4; 97,6]	96,2 [94,2; 97,5]
Monat 12		
Patienten unter Risiko	473	491
Zensierte Patienten	56	40
Patienten mit Ereignis	45	51
KM-Schätzer, % [95 % KI]	91,9 [89,4; 93,9]	91,0 [88,3; 93,1]
Monat 18		
Patienten unter Risiko	350	356
Zensierte Patienten	163	155
Patienten mit Ereignis	61	71
KM-Schätzer, % [95 % KI]	88,5 [85,5; 91,0]	86,8 [83,6; 89,4]
Monat 24		
Patienten unter Risiko	257	273
Zensierte Patienten	239	222
Patienten mit Ereignis	78	87
KM-Schätzer, % [95 % KI]	83,8 [80,1; 86,9]	82,5 [78,8; 85,7]
Monat 30		
Patienten unter Risiko	161	168
Zensierte Patienten	322	314
Patienten mit Ereignis	91	100
KM-Schätzer, % [95 % KI]	78,9 [74,4; 82,7]	78,0 [73,7; 81,7]

Monat 36		
Patienten unter Risiko	64	70
Zensierte Patienten	413	407
Patienten mit Ereignis	97	105
KM-Schätzer, % [95 % KI]	75,2 [69,9; 79,7]	75,2 [70,3; 79,5]
Monat 42		
Patienten unter Risiko	10	12
Zensierte Patienten	464	465
Patienten mit Ereignis	100	105
KM-Schätzer, % [95 % KI]	69,5 [61,1; 76,5]	75,2 [70,3; 79,5]
Monat 48		
Patienten unter Risiko	1	-
Zensierte Patienten	473	477
Patienten mit Ereignis	100	105
KM-Schätzer, % [95 % KI]	69,5 [61,1; 76,5]	75,2 [70,3; 79,5]
Behandlungseffekt (Selexipag gegenüber Placebo)		
HR [99 % KI]		0,97 [0,68; 1,39]
p (Log-Rang-Test)		0,4214

In die Analyse wurden auch die Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten.
KI – Konfidenzintervall, KM-Schätzer – Kaplan-Meier-Schätzer, HR – Hazard Ratio, FAS – *full analysis set*

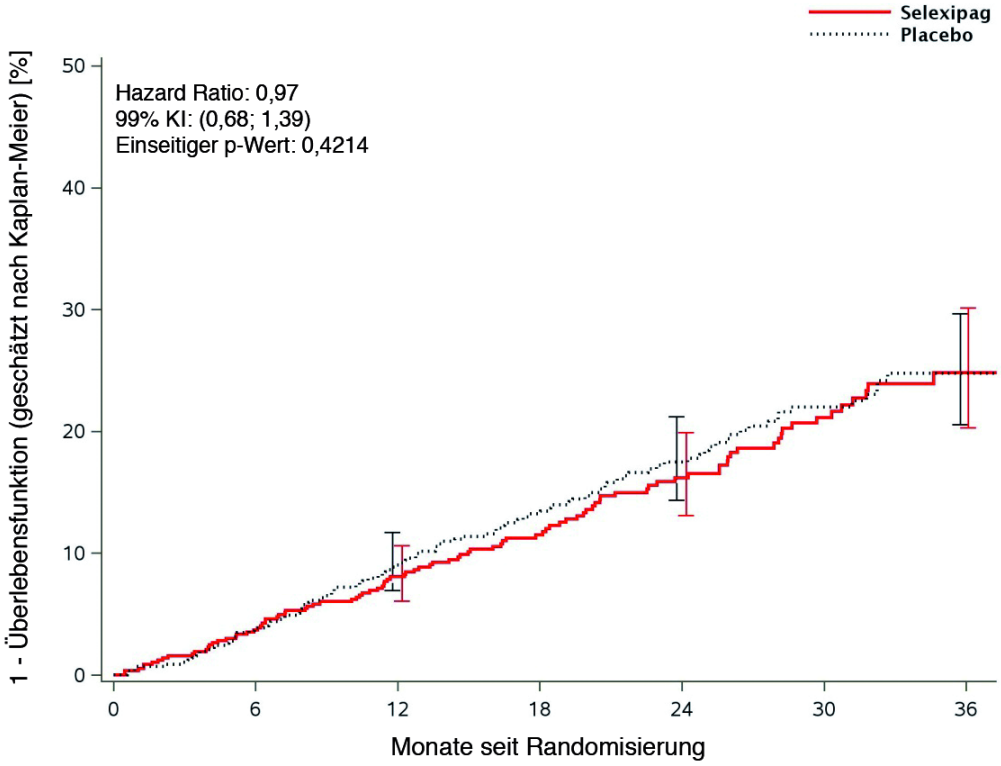
Hauptanalyse des Endpunkts

Die Analyse der Zeit bis zum Tod bis Studienschluss wurde auf Grundlage der FAS-Population durchgeführt.

Insgesamt verstarben bis zum Studienschluss unabhängig von der Ursache 100 Patienten in der Selexipag-Gruppe und 105 Patienten in der Placebo-Gruppe. Die HR [99 % KI] in der Ereigniszeitanalyse der Todesfälle bis Studienschluss betrug 0,97 [0,68; 1,39] und war statistisch nicht signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p=0,4214).

Die mediane Überlebenszeit wurde in beiden Armen bis zum Ende der Studie nicht erreicht, und auch für das erste Quartil lassen die vorliegenden Daten keine exakte Abschätzung der Werte zu.

Abbildung 4-10 zeigt die Kaplan-Meier-Kurve des Endpunkts in der FAS-Population.



Selexipag-Patienten:							
unter Risiko	574	543	473	350	257	161	64
Ereignisse	0	21	45	61	78	91	97
zensiert	0	10	56	163	239	322	413

Placebo-Patienten:							
unter Risiko	582	546	491	356	273	168	70
Ereignisse	0	22	51	71	87	100	105
zensiert	0	14	40	155	222	314	407

Abbildung 4-10: Kaplan-Meier-Kurve für „Zeit bis zum Tod (Studienschluss)“
Todesfälle, die vor dem 16. August 2011 auftraten, sind als Todesfälle in die Analyse eingeschlossen. Analysierte Population FAS. Fehlerbalken geben das 95 % KI an. FAS – *full analysis set*, KI – Konfidenzintervall

Nebenanalysen des Endpunkts

Analysen zum Einfluss des erlaubten Therapiewechsels („cross-over“)

Mithilfe des *Rank Preserving Structural Accelerated Failure Time*(RPSFT)-Modells sowie des *Marginal Structural Cox Proportional Hazards*-Modells (MSM) mit zeitabhängiger Gewichtung anhand des *Inverse Probability of Censoring Weighting*(IPCW)-Schemas sollte der Einfluss des erlaubten Therapiewechsels („cross-over“) vom Placebo-Arm in den Selexipag-Arm untersucht und damit die Ergebnisse der Zeit bis zum Tod bis Studienschluss adjustiert werden. Die Analysen wurden *post hoc* durchgeführt, eine detaillierte Beschreibung der jeweils angewendeten Methodik findet sich im Anhang [84, 85].

RPSFT-Modell

Das RPSFT-Modell ist ein nicht-parametrisches statistisches Modell, das auf Basis der randomisierten Gruppenzuteilung einen Schätzwert für den Therapieeffekt liefert, dabei aber die Verzerrung durch den erlaubten Therapiewechsel im Zeitraum nach dem Ende der doppelblinden Untersuchungsphase korrigiert (Nachbeobachtungsphase). Im Speziellen wurden die folgenden Änderungen der Behandlung im Modell untersucht:

1. Therapiewechsel von Placebo zu Selexipag
2. Abbruch der Behandlung mit Selexipag

Die durchschnittliche Dauer, mit der die Patienten, die ursprünglich auf Placebo randomisiert worden waren, mit Selexipag behandelt wurden, betrug 12,2 % der Gesamtzeit in der Nachbeobachtungsphase. Im Selexipag-Arm betrug die durchschnittliche Dauer der Selexipag-Behandlung in der Nachbeobachtungsphase hingegen 82,8 %.

Aus den vorliegenden Daten ergibt sich die Interpretation, dass Patienten, die durchgehend mit Selexipag behandelt worden wären, gegenüber Patienten, die niemals Selexipag erhalten hätten, eine relative Verlängerung des Überlebens um das 1,16-Fache erreichen (siehe Tabelle 4-39).

Tabelle 4-39: RPSFT-Modell: Schätzer für Ψ und für das relative Überleben

	Schätzer [95 % KI]
Schätzer für Ψ	-0,15 [-0,71; 0,54]
Schätzer des relativen Überlebens	1,16 [0,58; 2,03]
Werte >1 für den Schätzer des relativen Überlebens entsprechen einem Überlebensvorteil durch Selexipag, Werte <1 entsprechen einem Überlebensvorteil durch Placebo. Als Grundlage der Analyse diente die FAS-Population. KI – Konfidenzintervall	

Zur Durchführung des RPSFT-Modells mussten verschiedene Annahmen getroffen werden, die im Detail im Anhang beschrieben sind [85]. Eine dieser Annahmen – Annahme „iii“ – beschreibt, dass davon ausgegangen wird, dass neben Selexipag keine anderen Faktoren einen Einfluss auf das Überleben eines individuellen Patienten je nach Gruppenzuteilung haben. Diese Annahme ist jedoch voraussichtlich verletzt, da Patienten sowohl im Selexipag-Arm als auch im Placebo-Arm nach Ende der verblindeten Behandlung mit anderen, auf den Prostacyclin-Signalweg gerichteten PAH-spezifischen Arzneimitteln (d.h. Epoprostenol, Iloprost, Treprostinil oder Beraprost) behandelt werden konnten. In einem früheren Modell, das auf Wunsch der EMA in der vorgelegten Form angepasst wurde, wurde diesem Umstand insofern Rechnung getragen, als nicht nur die Tage auf Selexipag als aktive Behandlung gewertet wurden, sondern auch die Tage, in denen der Patient nur mit einem Prostanoid

behandelt wurde.[32] Mit einem Schätzer des relativen Überlebens [95 % KI] von 1,19 [0,58; 2,03] war das Ergebnis dieser früheren Analyse konsistent zur hier vorgelegten Analyse.

MSM mit zeitabhängiger Gewichtung anhand des IPCW-Schemas

Im Vergleich zum RPSFT-Modell wird beim *Marginal Structural Cox Proportional Hazards*-Modell das Risiko zu versterben mithilfe eines konventionellen Cox-Proportional-Hazards-Modells modelliert, das jedoch zeitabhängige Fallgewichte zur Bestimmung der Wahrscheinlichkeit enthält, dass ein Patient nicht aus dem Placebo-Arm in den Selexipag-Arm wechselt oder dass er die Selexipag-Behandlung im Selexipag-Arm nicht beendet.

Zur Schätzung der stabilisierten Fallgewichte und des MSM muss die Nachbeobachtungszeit in diskrete Zeitabschnitte unterteilt werden. Dies kann entweder durch die Verwendung einer glatten Zeitfunktion wie z. B. einem kubischen Spline geschehen oder durch die Kombination von Zeitintervallen. Im vorliegenden Modell wurden kubische Splines verwendet. Die verwendeten Zeitabschnitte waren 1 Monat, 3 Monate und 6 Monate, wobei 1 Monat 30,4 Tagen entsprach.

Wie auch bereits im RPSFT-Modell wurden im MSM die folgenden Änderungen der Behandlung untersucht:

1. Therapiewechsel von Placebo zu Selexipag
2. Abbruch der Behandlung mit Selexipag

Die resultierenden Effektschätzer der MSM-Analyse (siehe Tabelle 4-40) können als *Hazard Ratio* der Gesamtmortalität angesehen werden für den Vergleich von Patienten, wenn sie Selexipag durchgehend erhalten hätten, gegenüber Patienten, die niemals Selexipag erhalten hätten.

Die Effektschätzer unterscheiden sich je nach Unterteilung der Zeitabschnitte nur vergleichweise geringfügig, die HR bewegt sich in einer Spanne von 0,60 bis 0,65. Dies entspricht einer Verringerung des relativen Sterberisikos im Bereich von 25 bis 30 %. Der Unterschied zwischen den Behandlungsgruppen war unabhängig von der gewählten Unterteilung der Zeitabschnitte statistisch signifikant.

Die Ergebnisse der MSM-Analyse bestätigen damit die Annahme, dass der erlaubte Therapiewechsel aus dem Placebo-Arm in den Selexipag-Arm zu einer Verzerrung der Ergebnisse des Endpunkts zu Ungunsten von Selexipag führt. Ein negativer Effekt von Selexipag auf die Gesamtmortalität kann damit ausgeschlossen werden.

[32] Tage, an denen Patienten mit Selexipag und einem Prostanoid behandelt wurden, wurden nur einmal gezählt.

Tabelle 4-40: MSM – Schätzer des Behandlungseffekts (HR) mit 95 % KI, robustem Standardfehler und p-Wert

Unterteilung der Zeitabschnitte	**HR [95 % KI]**	**SE**	**Alpha**	**Pr > ChiSq**
1 Monat	0,65 [0,42; 0,99]	0,14	0,05	0,0462
3 Monate	0,61 [0,39; 0,95]	0,14	0,05	0,0295
6 Monate	0,60 [0,39; 0,93]	0,13	0,05	0,0221
HR – Hazard Ratio, KI – Konfidenzintervall, SE – Standardfehler (*standard error*)				

Ähnlich wie auch beim RPSFT-Modell wurden in einer früheren Ausführung des MSM auch Wechsel aus dem Selexipag- oder Placebo-Arm auf ein anderes, auf den Prostacyclin-Signalweg gerichtetes PAH-spezifisches Arzneimittel berücksichtigt. Auf Wunsch der EMA wurde dieses Modell in der vorgelegten Form angepasst und berücksichtigt nun ausschließlich Wechsel von Placebo zu Selexipag und Therapieabbrüche im Selexipag-Arm.

Die Ergebnisse der früheren Version der Analyse sind aber grundsätzlich konsistent zu der hier vorgelegten Analyse bzw. zeigen sogar stärkere und durchgehend statistisch signifikante Effekte.

Sensitivitätsanalysen

Um die Robustheit der Ergebnisse des Endpunkts „Zeit bis zum Tod (Studienschluss)“ einschätzen zu können, wurde eine Sensitivitätsanalyse durchgeführt, in der nur Patienten in die Auswertung eingeschlossen wurden, die von den Prüfärzten als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden (siehe Tabelle 4-41). Die HR [99 % KI] in dieser Analyse betrug 0,94 [0,65; 1,37], der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant.

Tabelle 4-41: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum Tod (Studienschluss)“

Sensitivitätsanalyse	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden.	0,94 [0,65; 1,37] p=0,3408
Weitere Erläuterungen zu den Sensitivitätsanalysen finden sich in Abschnitt 9.8.4.1.3 des Studienberichts. In die Analyse wurden auch die Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten. Als Grundlage der Analyse diente die FAS-Population. HR – Hazard Ratio, KI – Konfidenzintervall, FAS – *full analysis set*	

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nicht-stratifizierten Analyse des Endpunkts „Zeit bis zum Tod (Studienschluss)“ mit Analysen verglichen, die nach geographischer Region, PAH-Ätiologie, WHO-/NYHA-Klasse am Basiswert, PAH-spezifischer Begleittherapie am Basiswert bzw. WHO-/NYHA-Klasse und PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. Die Effektschätzer des Behandlungseffekts, die in diesen Analysen bestimmt wurden, waren konsistent über alle stratifizierten Analysen sowie zur nicht-stratifizierten Analyse (siehe Tabelle 4-42). In keiner der Analysen war der Unterschied zwischen den Behandlungsgruppen statistisch signifikant.

Tabelle 4-42: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Zeit bis zum Tod (Studienschluss)“

Modell	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Nicht-stratifiziertes Modell	0,97 [0,68; 1,39] p=0,4214
Modell stratifiziert nach Region[1]	0,97 [0,67; 1,39] p=0,4021
Modell stratifiziert nach PAH-Ätiologie[2]	0,98 [0,69; 1,41] p=0,4539
Modell stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I-II/III-IV)	0,99 [0,69; 1,42] p=0,4752
Modell stratifiziert nach PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)	0,97 [0,68; 1,39] p=0,4148
Modell stratifiziert nach WHO-/NYHA-Klasse und PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung	1,00 [0,70; 1,43] p=0,4884

[1]) Asien, Osteuropa, Lateinamerika, Nordamerika, Westeuropa, Australien

[2]) (Idiopathisch, angeboren, Drogen/Arzneimittel- oder Toxin-induziert), assoziiert mit HIV-Infektion, assoziiert mit Bindegewebserkrankung, assoziiert mit angeborenem Herzfehler mit korrigiertem Links-Rechts-Shunt.

In die Analyse wurden auch die Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten. Als Grundlage der Analyse diente die FAS-Population.

HR – Hazard Ratio, KI – Konfidenzintervall, CEC – *Critical Event Committee*, FAS – *full analysis set*

Naive Raten

Zusätzlich zur Ereigniszeitanalyse wurde die Anzahl der Ereignisse im Untersuchungszeitraum (die „Gesamtraten“), wie vom G-BA im Beratungsgespräch vom 13.05.2015 gefordert [3], als naive Raten, d. h. in Form von Risk Ratios (RR), Odds Ratios (OR) und absoluter Risikodifferenz (RD) samt zugehörigen 95 %-Konfidenzintervalle ausgewertet (die Ergebnisse dieser Analyse sind in Tabelle 4-43 dargestellt). Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt eine solche Auswertung jedoch voraus, dass alle Patienten zu einem fixen Endpunkt ausgewertet werden [83]. Hier findet jedoch eine Auswertung der Zeit bis zum Auftreten des ersten Ereignisses innerhalb eines Zeitraums von Randomisierung bis zum Ende der Studie (Studienschluss) statt und damit eben nicht zu einem festgelegten und für alle Patienten gleichen Zeitpunkt. Für die Post-hoc-Analyse der naiven Raten ist daher abweichend von der Hauptanalyse von einem hohen Verzerrungspotential auszugehen.

Wie bereits zuvor beschrieben, traten im Zeitraum von Randomisierung der Patienten bis Studienschluss insgesamt 100 Ereignisse im Selexipag-Arm auf und damit etwas weniger als im Placebo-Arm (105 Ereignisse). Dies entspricht einer numerischen Reduktion des relativen Risikos von 3 % (RR [95 % KI]: 0,97 [0,75; 1,24]) sowie einer numerischen Reduktion des absoluten Risikos um 0,6 Prozentpunkte (RD [95 % KI]: −0,6 %-P. [−5,0 %-P.; 3,8 %-P.]. Der Unterschied zwischen den Behandlungsgruppen ist für beide Effektmaße als statistisch nicht signifikant anzusehen. Die Auswertung der naiven Raten stimmt damit mit der Aussage der Hauptanalyse überein.

Tabelle 4-43: Analyse der naiven Raten für den Endpunkt „Zeit bis zum Tod (Studienschluss)“

Naive Raten				
Anzahl Todesfälle je Behandlungsgruppe (Naive Raten, Studienschluss)				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Gesamt	100	17,4	105	18,0
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Gesamt	0,97 [0,75; 1,24]	0,96 [0,71; 1,30]	-0,6 [-5,0; 3,8]	0,8480

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

In die Analyse wurden auch die Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten. Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-22). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa ähnlich ausgeprägt wie in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation nicht von einem unterschiedlichen Therapieffekt für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.4 Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-44: Operationalisierung von „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Zeit bis zum CEC-bestätigten Tod wegen PAH bis EOT. Die Ereignisse dieses Endpunkts wurden durch ein verblindetes, unabhängiges Critical Event Committee (CEC) begutachtet. Das CEC bestand aus 3 klinischen Experten mit Erfahrung in der PAH und in der Überwachung klinischer Studien. Als Mitglied des CEC war es nicht möglich, gleichzeitig als Prüfarzt an der Studie beteiligt zu sein. Die Hauptaufgabe des CEC war die Bestätigung des Auftretens eines Morbiditäts- oder Mortalitätsereignisses, gleichzeitig sollten aber auch alle Todesfälle auf einen möglichen Zusammenhang zur PAH überprüft werden. Die Hauptanalyse des Endpunkts wurde für den Zeitraum von der Randomisierung bis 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase (*end-of-treatment*, EOT) vorgenommen. Auswertungspopulation für die Hauptanalyse war die FAS-Population. *Zensierungsregeln/Ersetzungsregeln* Es wurde keine Imputationsmethode angewendet. Bei Patienten, bei denen bis zu 7 Tage nach der letzten Einnahme der Studienmedikation kein CEC-bestätigter Todesfall wegen PAH eingetreten war, wurde die Zeit bis zum Tod wegen PAH bis EOT nach folgenden Regeln zensiert: • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und innerhalb der Behandlungsphase + 7 Tage aus einem anderen Grund als der PAH verstorben waren: Zeitpunkt des Todes • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und keine Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302-Behandlungsphase plus 7 Tage, Zeitpunkt der EOS-Untersuchung, Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten und ihre Zustimmung zur Teilnahme an der Beobachtungsphase nach der Behandlungsphase gegeben hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302-Behandlungsphase plus 7 Tage, Zeitpunkt des letzten Kontakts, 27.04.2014) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die keine Studienmedikation erhalten hatten: Minimum (Maximum [Zeitpunkt der EOS-Untersuchung, Zeitpunkt der Randomisierung], Zeitpunkt des letzten Kontakts, 27.04.2014) minus Zeitpunkt der Randomisierung plus 1 Tag. Es handelt sich hierbei um (potentiell informative) rechts-zensierte Beobachtungen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-45: Bewertung des Verzerrungspotentials für „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es existieren jedoch sonstige das Verzerrungspotential beeinflussende Aspekte.

Bei PAH-Patienten tritt in aller Regel erst ein Morbiditätsereignis auf, bevor sie im weiteren Verlauf der Krankheit versterben. In der Studiensituation von GRIPHON bedeutet dies, dass Todesfälle, die sich nach Auftreten eines Morbiditätsereignisses ergeben und dadurch die randomisierte Behandlungsphase für den Patienten beenden, in Auswertungen bis EOT nicht erfasst werden. Diese informative Zensierung führt letztlich zu einer Unterschätzung des Sterberisikos in beiden Behandlungsgruppen und kann damit zu einer substantiellen Verzerrung der Ergebnisse führen.

Das Verzerrungspotential ist für diesen Endpunkt daher als hoch anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-46: Ergebnisse für „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**	
Intervention	Selexipag N=574	Placebo N=582
Zeit bis zum ersten Ereignis (Monate)		
1. Quartil [95 % KI]	-	-
Median [95 % KI]	-	-
3. Quartil [95 % KI]	-	-
Kaplan-Meier-Schätzer ereignisfreier Patienten (%)		
Monat 6		
Patienten unter Risiko	459	446
Zensierte Patienten	105	124
Patienten mit Ereignis	10	12
KM-Schätzer, % [95 % KI]	98,1 [96,5; 99,0]	97,6 [95,9; 98,7]
Monat 12		
Patienten unter Risiko	366	360
Zensierte Patienten	189	203
Patienten mit Ereignis	19	19
KM-Schätzer, % [95 % KI]	96,0 [93,8; 97,4]	95,9 [93,7; 97,4]
Monat 18		
Patienten unter Risiko	255	233
Zensierte Patienten	295	325
Patienten mit Ereignis	24	24
KM-Schätzer, % [95 % KI]	94,4 [91,7; 96,2]	94,2 [91,3; 96,1]
Monat 24		
Patienten unter Risiko	177	156
Zensierte Patienten	368	401
Patienten mit Ereignis	29	25
KM-Schätzer, % [95 % KI]	92,3 [89,0; 94,7]	93,8 [90,8; 95,8]
Monat 30		
Patienten unter Risiko	103	92
Zensierte Patienten	438	463
Patienten mit Ereignis	33	27
KM-Schätzer, % [95 % KI]	89,7 [85,3; 92,9]	92,3 [88,3; 94,9]

Monat 36		
Patienten unter Risiko	41	30
Zensierte Patienten	500	525
Patienten mit Ereignis	33	27
KM-Schätzer, % [95 % KI]	89,7 [85,3; 92,9]	92,3 [88,3; 94,9]
Monat 42		
Patienten unter Risiko	7	4
Zensierte Patienten	534	551
Patienten mit Ereignis	33	27
KM-Schätzer, % [95 % KI]	89,7 [85,3; 92,9]	92,3 [88,3; 94,9]
Monat 48		
Patienten unter Risiko	1	-
Zensierte Patienten	540	555
Patienten mit Ereignis	33	27
KM-Schätzer, % [95 % KI]	89,7 [85,3; 92,9]	92,3 [88,3; 94,9]
Behandlungseffekt (Selexipag gegenüber Placebo)		
HR [99 % KI]		1,16 [0,70; 1,93]
p (Log-Rang-Test)		0,7153

In die Analyse wurden auch die CEC-bestätigten Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten.
KI – Konfidenzintervall, KM-Schätzer – Kaplan-Meier-Schätzer, HR – Hazard Ratio, FAS – *full analysis set*

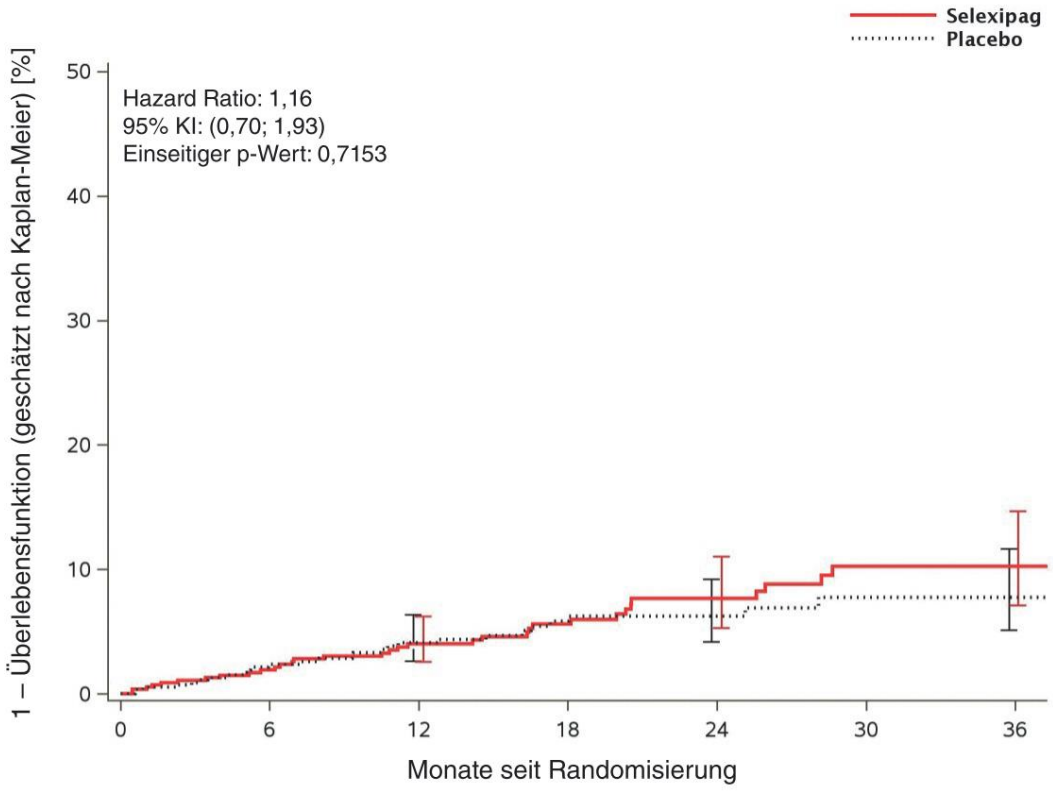
Hauptanalyse

Die Analyse der Zeit bis zum CEC-bestätigten Tod wegen PAH bis EOT wurde auf Grundlage der FAS-Population durchgeführt.

Insgesamt verstarben wegen PAH in der Behandlungsphase (EOT + 7 Tage) 33 Patienten in der Selexipag-Gruppe und 27 Patienten in der Placebo-Gruppe. Die HR [99 % KI] in der Ereigniszeitanalyse des CEC-bestätigten Todes wegen PAH bis EOT betrug 1,16 [0,70; 1,93] und war statistisch nicht signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p=0,7153).

Die mediane Überlebenszeit wurde in beiden Armen bis zum Ende der Studie nicht erreicht.

Abbildung 4-11 zeigt die Kaplan-Meier-Kurve des Endpunkts in der FAS-Population.



Selexipag-Patienten:							
unter Risiko	574	459	366	255	177	103	41
Ereignisse	0	10	19	24	29	33	33
zensiert	0	105	189	295	368	438	500

Placebo-Patienten:							
unter Risiko	582	446	360	233	156	92	30
Ereignisse	0	12	19	24	25	27	27
zensiert	0	124	203	325	401	463	525

Abbildung 4-11: Kaplan-Meier-Kurve für „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ Todesfälle, die vor dem 16. August 2011 auftraten, sind als Todesfälle in die Analyse eingeschlossen. Analysierte Population FAS. Fehlerbalken geben das 95 % KI an. FAS – *full analysis set*, KI – Konfidenzintervall

Nebenanalysen

Sensitivitätsanalysen

Um die Robustheit der Ergebnisse des Endpunkts „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“ einschätzen zu können, wurde eine Sensitivitätsanalyse durchgeführt, in der nur Patienten in die Auswertung eingeschlossen wurden, die von den Prüfärzten als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden (siehe Tabelle 4-47). Die HR [99 % KI] in dieser Analyse betrug 1,19 [0,70; 2,02], der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant.

Tabelle 4-47: Sensitivitätsanalysen für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“

Sensitivitätsanalyse	**Behandlungseffekt von Selexipag gegenüber Placebo** HR [99 % KI], p-Wert (Log-Rang)
Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden.	1,19 [0,70; 2,02] p=0,7440
Weitere Erläuterungen zu den Sensitivitätsanalysen finden sich in Abschnitt 9.8.4.1.3 des Studienberichts. In die Analyse wurden auch die CEC-bestätigten Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten. Als Grundlage der Analyse diente die FAS-Population. HR – Hazard Ratio, KI – Konfidenzintervall, FAS – *full analysis set*	

Naive Raten

Zusätzlich zur Ereigniszeitanalyse wurde die Anzahl der Ereignisse im Untersuchungszeitraum (die „Gesamtraten“), wie vom G-BA im Beratungsgespräch vom 13.05.2015 gefordert [3], als naive Raten, d. h. in Form von Risk Ratios (RR), Odds Ratios (OR) und absoluter Risikodifferenz (RD) samt zugehörigen 95 %-Konfidenzintervallen ausgewertet (die Ergebnisse dieser Analyse sind in Tabelle 4-48 dargestellt). Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt eine solche Auswertung jedoch voraus, dass alle Patienten zu einem fixen Endpunkt ausgewertet werden [83]. Hier findet jedoch eine Auswertung der Zeit bis zum Auftreten des ersten Ereignisses innerhalb eines Zeitraums von Randomisierung bis zum Zeitpunkt des patientenindividuellen Endes des Behandlungszeitraums (EOT) statt und damit eben nicht zu einem festgelegten und für alle Patienten gleichen Zeitpunkt. Für die Post-hoc-Analyse der naiven Raten ist daher abweichend von der Hauptanalyse von einem hohen Verzerrungspotential auszugehen.

Wie bereits zuvor beschrieben, traten im Zeitraum von Randomisierung der Patienten bis EOT + 7 Tage insgesamt 33 Ereignisse im Selexipag-Arm auf und damit etwas häufiger als im Placebo-Arm (27 Ereignisse). Dies entspricht einer numerischen Erhöhung des relativen Risikos von 24 % (RR [95 % KI]: 1,24 [0,76; 2,03]) sowie einer numerischen Erhöhung des absoluten Risikos um 1,1 Prozentpunkte (RD [95 % KI]: 1,1 %-P. [−1,4 %-P.; 3,7 %-P.]. Der Unterschied zwischen den Behandlungsgruppen ist für beide Effektmaße als statistisch nicht signifikant anzusehen. Die Auswertung der naiven Raten stimmt damit mit der Aussage der Hauptanalyse überein.

Tabelle 4-48: Analyse der naiven Raten für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)“

Naive Raten				
Anzahl Todesfälle je Behandlungsgruppe (Naive Raten, EOT)				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Gesamt	33	5,7	27	4,6
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Gesamt	1,24 [0,76; 2,03]	1,25 [0,74; 2,11]	1,1 [-1,4; 3,7]	0,5297

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

In die Analyse wurden auch die CEC-bestätigten Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten. Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, EOT – *end-of-treatment*, FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-23). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa ähnlich ausgeprägt wie in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation nicht von einem unterschiedlichen Therapieffekt für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.5 Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-49: Operationalisierung von „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Zeit bis zum CEC-bestätigten Tod wegen PAH bis Studienschluss. Die Ereignisse dieses Endpunkts wurden durch ein verblindetes, unabhängiges Critical Event Committee (CEC) begutachtet. Das CEC bestand aus 3 klinischen Experten mit Erfahrung in der PAH und in der Überwachung klinischer Studien. Als Mitglied des CEC war es nicht möglich, gleichzeitig als Prüfarzt an der Studie beteiligt zu sein. Die Hauptaufgabe des CEC war die Bestätigung des Auftretens eines Morbiditäts- oder Mortalitätsereignisses, gleichzeitig sollten aber auch alle Todesfälle auf einen möglichen Zusammenhang zur PAH überprüft werden. Die Analyse des Endpunkts wurde für den Zeitraum von der Randomisierung bis zum Studienschluss vorgenommen. Auswertungspopulation für die Hauptanalyse war die FAS-Population.
	Zensierungsregeln/Ersetzungsregeln Es wurde keine Imputationsmethode angewendet. Bei Patienten, bei denen bis Studienende kein CEC-bestätigter Todesfall wegen PAH eingetreten war, wurde die Zeit bis zum Tod wegen PAH bis zum Studienschluss nach folgenden Regeln zensiert: • Patienten, die aufgrund einer anderen Ursache als der PAH verstorben waren: Zeitpunkt des Todes • Alle übrigen Patienten: Minimum (Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014] minus Zeitpunkt der Randomisierung plus 1 Tag Es handelt sich hierbei um rechts-zensierte Beobachtungen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-50: Bewertung des Verzerrungspotentials für „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes konnten nicht beobachtet werden, jedoch besteht ein bedeutendes Verzerrungspotential durch den erlaubten Therapiewechsel am Ende der verblindeten Behandlungsphase (EOT). Patienten im Placebo-Arm, bei denen ein nicht-tödliches M/M-Ereignis aufgetreten war, konnten anschließend die verblindete Behandlung beenden und stattdessen mit Selexipag oder einem anderen PAH-spezifischen Arzneimittel weiterbehandelt werden. Insgesamt 155 Patienten aus dem Placebo-Arm wechselten nach Ende der Behandlungsphase die Therapie und wurden in offener Form mit Selexipag behandelt. Tendenziell ist aufgrund dessen jedoch nur mit einer Verzerrung des Therapieeffektes zuungunsten von Selexipag zu rechnen.

Das Verzerrungspotential für diesen Endpunkt ist insgesamt somit als hoch anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-51: Ergebnisse für „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**	
Intervention	Selexipag N=574	Placebo N=582
Zeit bis zum ersten Ereignis (Monate)		
1. Quartil [95 % KI]	- [38,2; -]	- [32,7; -]
Median [95 % KI]	-	-
3. Quartil [95 % KI]	-	-
Kaplan-Meier-Schätzer ereignisfreier Patienten (%)		
Monat 6		
Patienten unter Risiko	543	546
Zensierte Patienten	14	20
Patienten mit Ereignis	17	16
KM-Schätzer, % [95 % KI]	97,0 [95,2; 98,1]	97,2 [95,5; 98,3]
Monat 12		
Patienten unter Risiko	473	491
Zensierte Patienten	68	50
Patienten mit Ereignis	33	41
KM-Schätzer, % [95 % KI]	94,1 [91,7; 95,7]	92,7 [90,2; 94,5]
Monat 18		
Patienten unter Risiko	350	356
Zensierte Patienten	178	171
Patienten mit Ereignis	46	55
KM-Schätzer, % [95 % KI]	91,1 [88,3; 93,3]	89,6 [86,6; 91,9]
Monat 24		
Patienten unter Risiko	257	273
Zensierte Patienten	258	242
Patienten mit Ereignis	59	67
KM-Schätzer, % [95 % KI]	87,4 [83,9; 90,2]	86,2 [82,8; 89,1]
Monat 30		
Patienten unter Risiko	161	168
Zensierte Patienten	349	335
Patienten mit Ereignis	64	79
KM-Schätzer, % [95 % KI]	85,2 [81,3; 88,4]	81,9 [77,7; 85,4]

Monat 36		
Patienten unter Risiko	64	70
Zensierte Patienten	441	429
Patienten mit Ereignis	69	83
KM-Schätzer, % [95 % KI]	81,9 [76,9; 86,0]	79,6 [74,8; 83,5]
Monat 42		
Patienten unter Risiko	10	12
Zensierte Patienten	494	487
Patienten mit Ereignis	70	83
KM-Schätzer, % [95 % KI]	80,0 [73,4; 85,1]	79,6 [74,8; 83,5]
Monat 48		
Patienten unter Risiko	1	-
Zensierte Patienten	503	499
Patienten mit Ereignis	70	83
KM-Schätzer, % [95 % KI]	80,0 [73,4; 85,1]	79,6 [74,8; 83,5]
Behandlungseffekt (Selexipag gegenüber Placebo)		
HR [99 % KI]		0,86 [0,63; 1,18]
p (Log-Rang-Test)		0,1763

In die Analyse wurden auch die CEC-bestätigten Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten.
KI – Konfidenzintervall, KM-Schätzer – Kaplan-Meier-Schätzer, HR – Hazard Ratio, FAS – *full analysis set*

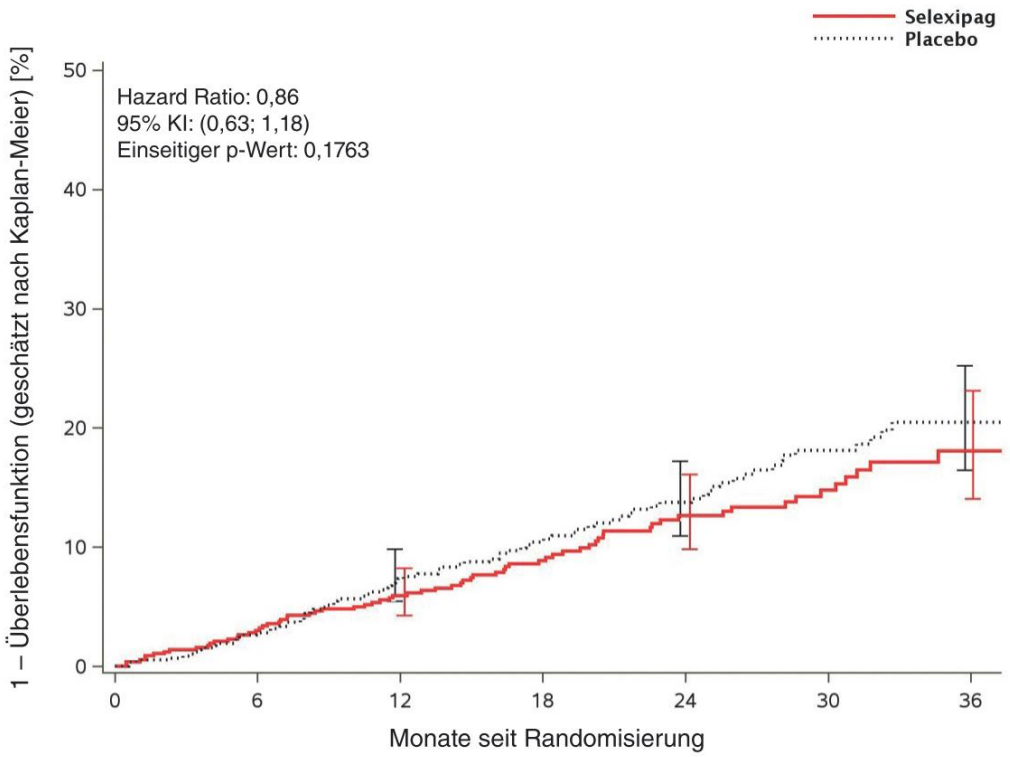
Hauptanalyse

Die Analyse der Zeit bis zum CEC-bestätigten Tod bis Studienschluss wurde auf Grundlage der FAS-Population durchgeführt.

Insgesamt verstarben wegen PAH bis zum Studienschluss 70 Patienten in der Selexipag-Gruppe und 83 Patienten in der Placebo-Gruppe. Die HR [99 % KI] in der Ereigniszeitanalyse des Todes bis Studienschluss betrug 0,86 [0,63; 1,18] und war statistisch nicht signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p=0,1763).

Die mediane Überlebenszeit wurde in beiden Armen bis zum Ende der Studie nicht erreicht.

Abbildung 4-12 zeigt die Kaplan-Meier-Kurve des Endpunkts in der FAS-Population.



Selexipag-Patienten:							
unter Risiko	574	543	473	350	257	161	64
Ereignisse	0	17	33	46	59	64	69
zensiert	0	14	68	178	258	349	441

Placebo-Patienten:							
unter Risiko	582	546	491	356	273	168	70
Ereignisse	0	16	41	55	67	79	83
zensiert	0	20	50	171	242	335	429

Abbildung 4-12: Kaplan-Meier-Kurve für „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“
Todesfälle, die vor dem 16. August 2011 auftraten, sind als Todesfälle in die Analyse eingeschlossen. Analysierte Population FAS. Fehlerbalken geben das 95 % KI an. FAS – *full analysis set*, KI – Konfidenzintervall

Nebenanalysen

Naive Raten

Zusätzlich zur Ereigniszeitanalyse wurde die Anzahl der Ereignisse im Untersuchungszeitraum (die „Gesamtraten“), wie vom G-BA im Beratungsgespräch vom 13.05.2015 gefordert [3], als naive Raten, d. h. in Form von Risk Ratios (RR), Odds Ratios (OR) und absoluter Risikodifferenz (RD) samt zugehörigen 95 %-Konfidenzintervallen ausgewertet (die Ergebnisse dieser Analyse sind in Tabelle 4-48 dargestellt). Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt eine solche Auswertung jedoch voraus, dass

alle Patienten zu einem fixen Endpunkt ausgewertet werden [83]. Hier findet jedoch eine Auswertung der Zeit bis zum Auftreten des ersten Ereignisses innerhalb eines Zeitraums von Randomisierung bis zum Ende der Studie (Studienschluss) und damit eben nicht zu einem festgelegten und für alle Patienten gleichen Zeitpunkt. Für die Post-hoc-Analyse der naiven Raten ist daher abweichend von der Hauptanalyse von einem hohen Verzerrungspotential auszugehen.

Wie bereits zuvor beschrieben, traten im Zeitraum von Randomisierung der Patienten bis Studienschluss insgesamt 70 Ereignisse im Selexipag-Arm auf und damit weniger als im Placebo-Arm (83 Ereignisse). Dies entspricht einer numerischen Reduktion des relativen Risikos von 14% (RR [95% KI]: 0,86 [0,64; 1,15]) sowie einer numerischen Reduktion des absoluten Risikos um 2,1 Prozentpunkte (RD [95%]: 2,1%-P. [−6,0%-P.; 1,8%-P.]. Der Unterschied zwischen den Behandlungsgruppen ist für beide Effektmaße als statistisch nicht signifikant anzusehen. Die Auswertung der naiven Raten stimmt damit mit der Aussage der Hauptanalyse überein.

Tabelle 4-52: Analyse der naiven Raten für den Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)“

Naive Raten				
Anzahl Todesfälle je Behandlungsgruppe (Naive Raten, Studienschluss)				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Gesamt	70	12,2	83	14,3
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95% KI]	OR [95% KI]	RD [95% KI] (Prozentpunkte)	p
Gesamt	0,86 [0,64; 1,15]	0,84 [0,59; 1,18]	-2,1 [-6,0; 1,8]	0,3057

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

In die Analyse wurden auch die CEC-bestätigten Todesfälle miteingeschlossen, die sich vor dem 16. August 2011 ereigneten. Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt

wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-24). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.6 Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-53: Operationalisierung von „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH. Die Erhebung der Hospitalisierungen erfolgte über die UE-Seite des Prüfbogens (*case report form*, CRF), in der die stationäre Aufnahme des Patienten dokumentiert wurde. PAH-bezogene Hospitalisierungen wurden definiert als Hospitalisierungen, die mit einem UE der folgenden bevorzugten Bezeichnungen (*preferred terms*, PT) der MedDRA-Kodierung (Version 16) assoziiert waren: „Akutes Rechtsherzversagen“, „Rechtsherzversagen“, „Herzversagen“, „Pulmonal arterielle Hypertonie“, „Pulmonale Hypertonie“, „Chronisches Herzversagen“, „Akutes Cor pulmonale“, „kongestive Herzinsuffizienz“, „rechtsventrikuläre Dysfunktion“, „Synkope“, „Bewusslosigkeit“, „Dyspnoe“, „Dyspnoe im Ruhezustand“, „Hypoxie“, „Pulmonale Ödeme“, „Periphere Ödeme“, „Vorhofflimmern“, „Vorhofflattern“, „Herz- und Atemstillstand“, „Kardio-pulmonales Versagen“ Die Ereignisse dieses Endpunkts wurden durch ein verblindetes, unabhängiges Critical Event Committee (CEC) begutachtet. Das CEC bestand aus 3 klinischen Experten mit Erfahrung in der PAH und in der Überwachung klinischer Studien. Als Mitglied des CEC war es nicht möglich, gleichzeitig als Prüfarzt an der Studie beteiligt zu sein. Die Hauptaufgabe des CEC war die Bestätigung des Auftretens eines Morbiditäts- oder Mortalitätsereignisses, gleichzeitig sollten aber auch alle Todesfälle auf einen möglichen Zusammenhang zur PAH überprüft werden. Die Hauptanalyse des Endpunkts wurde für den Zeitraum von der Randomisierung bis 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase (*end-of-treatment*, EOT) vorgenommen. Auswertungspopulation für die Hauptanalyse war die FAS-Population. *Zensierungsregeln/Ersetzungsregeln* Es wurde keine Imputationsmethode angewendet. Zensierungsregeln Bei Patienten ohne Hospitalisierung wegen PAH wurde die Zeit bis zur ersten CEC-bestätigten Hospitalisierung bis EOT nach den folgenden Regeln zensiert: • Patienten, die die Studienmedikation mindestens einmal eingenommen hatten: Minimum (Zeitpunkt der letzten Einnahme der Studienmedikation in der AC-065A302-Behandlungsphase plus 7 Tage, Zeitpunkt des letzten Kontakts, Cut-off-Zeitpunkt der Analyse [27.04.2014]) minus Zeitpunkt der Randomisierung plus 1 Tag. • Patienten, die keine Studienmedikation erhalten hatten: Zeitpunkt der Randomisierung plus 1 Tag.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-54: Bewertung des Verzerrungspotentials für „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-55: Ergebnisse für „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**	
Intervention	Selexipag N=574	Placebo N=582
Zeit bis zum ersten Ereignis (Monate)		
1. Quartil [95 % KI]	36,4 [24,7; -]	21,5 [18,4; 29,4]
Median [95 % KI]	- 	- [39,0; -]
3. Quartil [95 % KI]	-	-
Kaplan-Meier-Schätzer ereignisfreier Patienten		
Monat 6		
Patienten unter Risiko	458	438
Zensierte Patienten	89	95
Patienten mit Ereignis	27	49
KM-Schätzer, % [95 % KI]	94,8 [92,5; 96,4]	90,8 [88,0; 93,0]
Monat 12		
Patienten unter Risiko	371	356
Zensierte Patienten	154	150
Patienten mit Ereignis	49	76
KM-Schätzer, % [95 % KI]	90,0 [87,0; 92,4]	84,9 [81,4; 87,8]
Monat 18		
Patienten unter Risiko	259	233
Zensierte Patienten	250	255
Patienten mit Ereignis	65	94
KM-Schätzer, % [95 % KI]	85,4 [81,6; 88,4]	79,8 [75,7; 83,2]
Monat 24		
Patienten unter Risiko	179	158
Zensierte Patienten	321	313
Patienten mit Ereignis	74	111
KM-Schätzer, % [95 % KI]	81,9 [77,6; 85,5]	73,2 [68,3; 77,5]
Monat 30		
Patienten unter Risiko	111	93
Zensierte Patienten	381	372
Patienten mit Ereignis	82	117
KM-Schätzer, % [95 % KI]	77,6 [72,4; 81,9]	70,0 [64,6; 74,7]

Monat 36		
Patienten unter Risiko	47	37
Zensierte Patienten	444	423
Patienten mit Ereignis	83	122
KM-Schätzer, % [95 % KI]	76,8 [71,3; 81,3]	63,3 [55,4; 70,2]
Monat 42		
Patienten unter Risiko	9	6
Zensierte Patienten	479	452
Patienten mit Ereignis	86	124
KM-Schätzer, % [95 % KI]	70,9 [62,2; 78,0]	57,8 [46,9; 67,2]
Monat 48		
Patienten unter Risiko	1	0
Zensierte Patienten	487	458
Patienten mit Ereignis	86	124
KM-Schätzer, % [95 % KI]	70,9 [62,2; 78,0]	57,8 [46,9; 67,2]
Behandlungseffekt (Selexipag gegenüber Placebo)		
HR [95 % KI]		0,65 [0,50; 0,86]
p (Log-Rang-Test)		0,0022

Bei der Ereigniszeitanalyse handelt sich um die Post-hoc-Analyse eines prospektiv geplanten Endpunkts.
KI – Konfidenzintervall, KM-Schätzer – Kaplan-Meier-Schätzer, HR – Hazard Ratio, FAS – *full analysis set*

Hauptanalyse

PAH-bedingte Hospitalisierungen sind ein objektives Kriterium für die Verschlechterung der PAH sowie ein Indiz für die Schwere der Erkrankung und die Belastung für den Patienten. In GRIPHON wurde die Anzahl an Hospitalisierungen wegen PAH und die Anzahl an Tagen mit stationärem Aufenthalt als vordefinierte pharmakoökonomische Endpunkte analysiert.

Hierbei ist jedoch zu beachten, dass aufgrund des Studiendesigns in GRIPHON bei manchen Patienten lediglich ein Hospitalisierungsereignis aufgezeichnet wurde, da in vielen Fällen die betreffende Hospitalisierung mit einem Ereignis des primären Endpunktes und damit mit dem Behandlungsende (EOT) assoziiert war. *Post hoc* wurde daher stattdessen die Zeit bis zum ersten Hospitalisierungsereignis wegen PAH analysiert, was die genannte Problematik berücksichtigt und somit als die geeignetere Analysemethode angesehen wird.

Die Analyse der Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH wurde auf Grundlage der FAS-Population durchgeführt.

Insgesamt wurden im Beobachtungszeitraum (EOT + 7 Tage) 86 Patienten in der Selexipag-Gruppe und 124 Patienten in der Placebo-Gruppe stationär wegen PAH aufgenommen. Die Hazard Ratio (HR) [95 % KI] in der Ereigniszeitanalyse für das Auftreten einer PAH-bedingten Hospitalisierung betrug 0,65 [0,50; 0,86] und war statistisch signifikant (einseitiger

nicht-stratifizierter Log-Rang-Test: p=0,0022). Dies entspricht einer Reduktion des relativen Risikos von 35 %.

Die mediane Zeit bis zum ersten Ereignis wurde in beiden Armen im Behandlungszeitraum nicht erreicht, das erste Quartil [95 % KI] war im Selexipag-Arm nach 36,4 Wochen [24,7 Wochen; –], im Placebo-Arm hingegen bereits nach 21,5 Wochen [18,4 Wochen; 29,4 Wochen] erreicht.

Abbildung 4-13 zeigt die Kaplan-Meier-Kurve des Endpunkts in der FAS-Population.

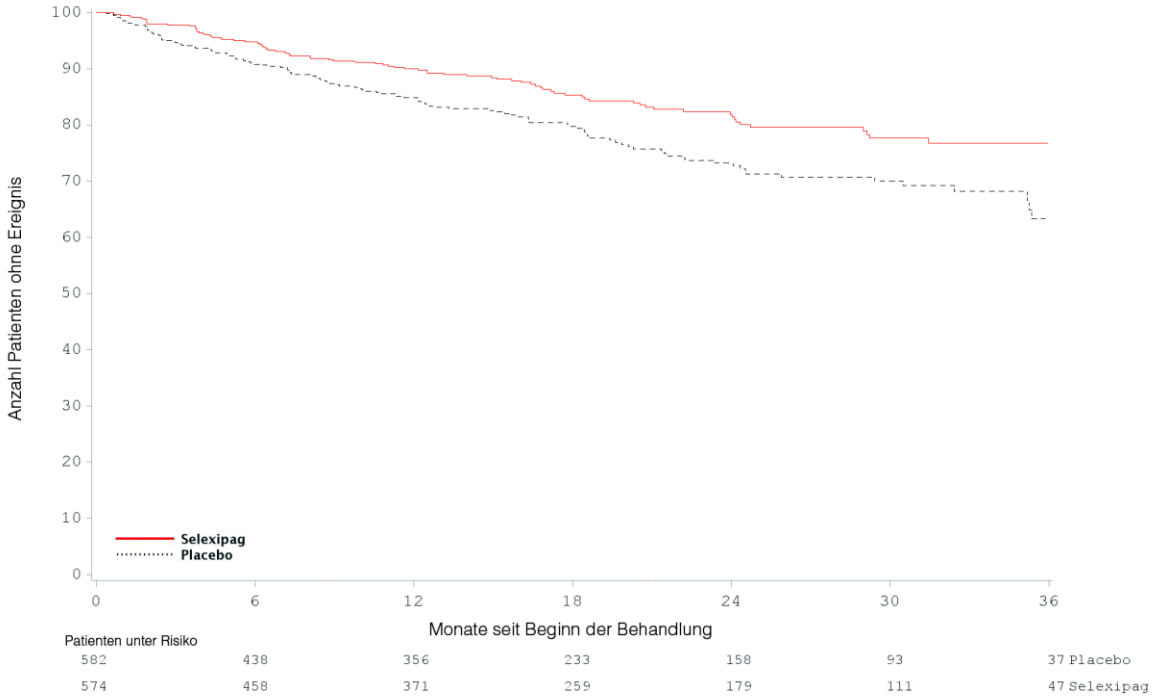


Abbildung 4-13: Kaplan-Meier-Kurve für „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. CEC – *Critical Event* Committee, FAS – *full analysis set*

Nebenanalysen

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf den Zeitraum von der Randomisierung bis EOT + 7 Tage.

Naive Raten

Zusätzlich zur Ereigniszeitanalyse wurde die Anzahl der Ereignisse im Untersuchungszeitraum (die „Gesamtraten“), wie vom G-BA im Beratungsgespräch vom 13.05.2015 gefordert [3], als naive Raten, d. h. in Form von Risk Ratios (RR), Odds Ratios (OR) und absoluter Risikodifferenz (RD) samt zugehörigen 95 %-Konfidenzintervallen ausgewertet (die Ergebnisse dieser Analyse sind in Tabelle 4-56 dargestellt). Gemäß dem Cochrane Handbook

for Systematic Reviews of Interventions setzt eine solche Auswertung jedoch voraus, dass alle Patienten zu einem fixen Endpunkt ausgewertet werden [83]. Hier findet jedoch eine Auswertung des Zeitraumes bis zum Auftreten eines Ereignisses des primären Endpunktes, d. h. zum Zeitpunkt des patientenindividuellen Endes des Behandlungszeitraums (EOT) statt und damit eben nicht zu einem festgelegten und für alle Patienten gleichen Zeitpunkt. Für die Post-hoc-Analyse der naiven Raten ist daher abweichend von der Hauptanalyse von einem hohen Verzerrungspotential auszugehen.

Wie bereits zuvor beschrieben, traten im Zeitraum von Randomisierung der Patienten bis EOT + 7 Tage insgesamt 86 Ereignisse im Selexipag-Arm auf und damit bedeutend weniger als im Placebo-Arm (124 Ereignisse). Dies entspricht einer signifikanten Reduktion des relativen Risikos von 30 % (RR [95 % KI]: 0,70 [0,55; 0,90]) sowie einer signifikanten Reduktion des absoluten Risikos von 6,3 Prozentpunkten (RD [95 % KI]: −6,3 %-P. [−10,7 %-P.; −1,9 %-P.]. Die Auswertung der naiven Raten stimmt damit vollständig mit der Aussage der Hauptanalyse überein.

Tabelle 4-56: Analyse der naiven Raten für den Endpunkt „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“

Naive Raten				
Anzahl Ereignisse je Behandlungsgruppe (Naive Raten, EOT)				
Intervention	Selexipag N=574		Placebo N=582	
	n	%	n	%
Gesamt	86	15,0	124	21,3
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Gesamt	0,70 [0,55; 0,90]	0,65 [0,48; 0,88]	-6,3 [-10,7; -1,9]	0,0054

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt und besitzen ein hohes Verzerrungspotential.

Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, EOT – *end-of-treatment*, FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt

wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-25). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.7 Symptome der PAH

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-57: Operationalisierung von „Symptome der PAH“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Anteil der Patienten mit mindestens einem Symptom der PAH nach Definition des G-BA. Als Symptom der PAH nach Definition des G-BA werden UE definiert, die einer der folgenden bevorzugten Bezeichnungen (PT) der MedDRA-Kodierung (Version 16) zuzuordnen sind: • Schwindel • Schwindel unter Belastung • Dyspnoe • Dyspnoe in Ruhe • Paroxysmale nächtliche Dyspnoe • Erschöpfung • Prä-Synkope • Synkope • Überschüssige Flüssigkeit • Wasseransammlung • Generalisiertes Ödem • Lokalisiertes Ödem • Ödem • Peripheres Ödem • Aszites • Hypervolämie • Hydrothorax Der Endpunkt „Symptome der PAH“ war in GRIPHON nicht prospektiv definiert, sondern wurde *post hoc* als Ergebnis der Beratungsgespräche zu Macitentan [3] bzw. Selexipag [14] definiert (vgl. Abschnitt 4.2.5.2). Die Zusammenstellung der PT richtet sich nach den vom G-BA als wichtig erachteten Symptomen der PAH (Dyspnoe, Synkope, Ödeme, Schwindel, Müdigkeit). Zur Operationalisierung der UE siehe Tabelle 4-86.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-58: Bewertung des Verzerrungspotentials für „Symptome der PAH“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es liegen jedoch sonstige das Verzerrungspotential beeinflussende Aspekte vor. So setzt die Auswertung der Ereignisse des Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen gemäß dem Cochrane Handbook for Systematic Reviews of Interventions jedoch voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden [83]. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer bzw. nicht mehr handhabbarer prostacyclinspezifischer Nebenwirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben.

Insgesamt ist das Verzerrungspotential für diesen Endpunkt als hoch anzusehen. Aufgrund der vergleichsweise längeren medianen Behandlungsdauer in der Selexipag-Gruppe (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchstens von einer Verzerrung der Ergebnisse zuungunsten von Selexipag auszugehen.

Stellen Sie die Ergebnisse für den Endpunkt „Symptome der PAH" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-59: Ergebnisse für „Symptome der PAH" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: SAF**			
Intervention	Selexipag N=575		Placebo N=577	
Anzahl Patienten mit mindestens einem vom G-BA definierten Symptom der PAH				
	n	%	n	%
Patienten mit mindestens einem vom G-BA definierten Symptom der PAH	251	43,7	297	51,5
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem vom G-BA definierten Symptom der PAH	0,85 [0,75; 0,96]	0,73 [0,58; 0,92]	-7,8 [-13,6; -2,1]	0,0081

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Hauptanalyse

Die Analyse der Symptome der PAH (nach Definition des G-BA) wurde auf Grundlage der SAF-Population durchgeführt.

Insgesamt trat im Verlauf der Behandlungsphase von GRIPHON bei 43,7 % der Patienten im Selexipag-Arm und damit bei deutlich weniger Patienten als im Placebo-Arm (51,5 %) mindestens ein Symptom der PAH auf. Dies entspricht einer Reduktion des relativen Risikos von 15 % (RR [95 % KI]: 0,85 [0,75; 0,96]) bzw. einer Reduktion des absoluten Risikos von 7,8 Prozentpunkten (RD [95 % KI]: -7,8 %-P. [−13,6 %-P.; −2,1 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (Barnards exakter Test: p=0,0081).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden.

Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-26). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.8 Absolute Veränderung der 6-MWD

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-60: Operationalisierung von „Absolute Veränderung der 6-MWD“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Absolute Veränderung der 6-MWD vom Basiswert bis zu Woche 26, jeweils gemessen zum Talspiegel der Selexipag-Konzentration. Die Erhebung der 6-MWD erfolgte nach einem von Actelion vorgegebenen standardisierten Protokoll (es handelt sich dabei um ein modifiziertes Protokoll der American Thoracic Society). Die absolute Veränderung wurde bestimmt als Differenz des 6-MWD-Werts in Woche 26 minus dem 6-MWD-Wert zum Zeitpunkt der Basiswerterhebung. Für den Woche-26-Wert wurde der 6-MWD-Wert herangezogen, der innerhalb des Woche-26-Zeitfensters (Studientage 148-271) bestimmt wurde. In den Fällen, in denen mehr als ein 6-MWD-Wert innerhalb des Woche-26-Zeitfensters vorlag, wurde derjenige Wert ausgewählt, der näher am nominalen Wert für Woche 26 (Studientag 182) lag. Als „zum Talspiegel gemessen“ wurde der 6-MWT dann gezählt, wenn die letzte Selexipag-Anwendung entweder am Vortag der Testdurchführung stattgefunden hatte oder am Testtag selbst, aber mindestens 12 Stunden zurücklag. Wenn der 6-MWT im Zusammenhang mit der Untersuchung nach dem Auftreten eines M/M-Ereignis gemessen wurde, wurde er standardmäßig als „zum Talspiegel gemessen“ betrachtet. Auswertungspopulation für die Hauptanalyse war die FAS-Population.

	Ersetzungregeln
	Für die Hauptanalyse wurden sämtliche Ergebnisse des 6-MWT in Woche 26 verwendet, unabhängig davon, ob der Test zum Talspiegel durchgeführt wurde oder nicht.
	Bei Patienten, für die in Woche 26 kein Wert für die 6-MWD vorlag, wurde der Wert nach den folgenden Regeln ersetzt:
	• Regel 1: Für Patienten, die in Woche 26 nicht gehen konnten, wurde der fehlende Wert durch 0 m ersetzt. Hierzu gehörten:
	  ○ Patienten, die vor Tag 271 (Obergrenze des Zeitfensters der Woche-26-Untersuchung) verstarben, ohne dass eine Woche-26-Untersuchung stattgefunden hatte.
	  ○ Patienten, bei denen die Untersuchung in Woche 26 der Untersuchung aufgrund eines Morbiditäts- oder Mortalitätsereignisses entsprach und die aus PAH-bezogenen Gründen nicht gehen konnten, d. h. als Grund war „Dyspnoe/Schwäche“ oder „im Zusammenhang mit der PAH“ kodiert.
	Anmerkung: Patienten, die zu den geplanten Untersuchungen aus PAH-bezogenen Gründen nicht gehen konnten, wurde 0 m als Wert in das CRF eingetragen.
	• Regel 2 (wenn Regel 1 nicht anwendbar war): Der fehlende Wert wurde durch den zweitniedrigsten 6-MWD-Wert ersetzt, der in Woche 26 in der gleichen Auswertungspopulation beobachtet wurde (unabhängig davon, ob er zum Zeitpunkt des Talspiegels gemessen wurde, und unabhängig von der Behandlungsgruppe). In der FAS-Population betrug dieser Wert 10 m.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-61: Bewertung des Verzerrungspotentials für „Absolute Veränderung der 6-MWD“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Absolute Veränderung der 6-MWD“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-62: Ergebnisse für „Absolute Veränderung der 6-MWD“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**					
Intervention	Selexipag (N=574)			Placebo (N=582)		
Veränderung der 6-MWD vom Basiswert bis zu Woche 26 (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	574	358,5	76,3	582	348,0	83,2
Woche 26	574	306,5	170,0	582	281,7	173,8
Veränderung Woche 26 gegenüber Basiswert	**574**	**-52,0**	**150,2**	**582**	**-66,3**	**148,2**
Veränderung der 6-MWD vom Basiswert bis zu Woche 26 (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	574	376	90; 482	582	369	50; 515
Woche 26	574	370	0; 617	582	346	0; 650
Veränderung Woche 26 gegenüber Basiswert	**574**	**4,0**	**-448; 260**	**582**	**-9,0**	**-438; 262**
Imputation fehlender Werte für Woche 26						
	n	%		n	%	
Insgesamt ersetzt	114	19,9		136	23,4	
Grund: Tod[1]	25	4,4		26	4,5	
Grund: nicht Tod	89	15,5		110	18,9	
Woche-26-Untersuchung durchgeführt	11	1,9		13	2,2	
6-MWT nicht durchgeführt, Grund: nicht PAH-bezogen[2]	11	1,9		13	2,2	
6-MWT durchgeführt, Grund: PAH-bezogen[1]	-	-		-	-	
Woche-26-Untersuchung nicht durchgeführt[2]	78	13,6		97	16,7	

Behandlungseffekt (Veränderung der 6-MWD vom Basiswert bis Woche 26)		
Punktschätzer [99 % KI][3]		12,0 [1; 24]
p-Wert (Wilcoxon-Mann-Whitney)[4]		0,0027

[1]) Der fehlende Wert wurde durch 0 m ersetzt (Regel 1)
[2]) Der fehlende Wert wurde durch den zweitschwersten Rangwert in Woche 26 ersetzt (Regel 2; in der FAS-Population: 10 m)
[3]) Nach der Hodges-Lehmann-Methode
[4]) Aus nicht-parametrischer ANCOVA mit 6-MWD-Basiswert als Kovariate
Die 6-MWD ist in Metern (m) angegeben.
6-MWD – 6-Minuten-Gehstrecke, MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max – Maximalwert, KI – Konfidenzintervall, ANCOVA – Kovarianzanalyse, FAS – *full analysis set*

Hauptanalyse

Die Analyse der absoluten Veränderung der 6-MWD wurde auf Grundlage der FAS-Population durchgeführt.

Die mediane absolute Veränderung der 6-MWD vom Basiswert bis Woche 26 betrug 4,0 m bei Patienten in der Selexipag-Gruppe, während sich die Patienten in der Placebo-Gruppe im gleichen Zeitraum im Median um 9,0 m verschlechterten. Der Behandlungseffekt [99 % KI] von Selexipag im Vergleich zu Placebo betrug 12,0 m [1 m; 24 m], der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (Wilcoxon-Mann-Whitney-Test aus nicht-parametrischer ANCOVA mit 6-MWD-Basiswert als Kovariate: p=0,0027).

Im Vergleich zu den 6-MWD-Werten, die in vergangenen RCTs gemessen wurden, ist der in GRIPHON gemessene Wert vergleichsweise niedrig. Dies ist in Anbetracht der Patientencharakteristika der in GRIPHON eingeschlossenen Patienten jedoch nicht überraschend. Insbesondere hervorzuheben wäre in diesem Zusammenhang, dass der Großteil (ca. 80 %) der Patienten in GRIPHON zum Zeitpunkt der Basiswerterhebung schon mit einem oder sogar zwei PAH-spezifischen Arzneimitteln behandelt wurde (vgl. Tabelle 4-13). Dies könnte jedoch zu einer Limitierung der Verbesserung des einzelnen Patienten führen (*Ceiling*-Effekt), wie sie in der Vergangenheit bei Patienten mit bestehender PAH-spezifischer Therapie auch schon beobachtet wurde [86], und damit auch die Unterschiede zwischen den Behandlungsgruppen nivellieren. In früheren Studien war der Therapieeffekt auf die 6-MWD dann am größten, wenn das untersuchte Arzneimittel als erste PAH-spezifische Therapie eingesetzt wurde, während die zusätzliche Anwendung eines zweiten oder sogar eines dritten PAH-spezifischen Arzneimittels nicht zwangsläufig zu einer weiteren Verbesserung symptomatischer Endpunkte führen musste [86]. Konsistent zu dieser früheren Beobachtung war auch in GRIPHON der Behandlungseffekt in der Subgruppe der Patienten ohne PAH-spezifische Begleittherapie sehr viel höher als in der Gesamtstudienpopulation (Behandlungseffekt [95 % KI]: 34,0 m [15,0 m, 55,0 m], vgl. auch Abbildung 4-27). Dieser Wert entspricht wiederum den Beobachtungen mit therapienaiven Patienten in früheren Studien (z. B. BREATHE-1: 35 m) [87].

Der Therapieeffekt könnte zudem auch durch den hohen Anteil an Patienten in der vergleichsweise weniger schweren WHO-/NYHA-Klasse II (45,8 %, vgl. Tabelle 4-13) limitiert sein. Für diese Patientengruppe wurde in der Vergangenheit ebenfalls ein *Ceiling*-Effekt beobachtet [86, 88], sodass diese Patienten ihre vergleichsweise wenig eingeschränkte 6-MWD zum Zeitpunkt der Basiswerterhebung nicht in gleichem Maße steigern konnten wie Patienten mit schwerer Symptomatik.

Wie bereits in Abschnitt 4.2.5.2 beschrieben, stellen schließlich auch die äußerst konservativ gewählte Ersetzungsstrategie und der ebenfalls konservative Zeitpunkt der Messung in Bezug auf die letzte Einnahme der Studienmedikation (Talspiegel vs. Spitzenspiegel) deutliche Limitationen im Nachweis eines Therapieeffekts dar.

Nebenanalysen

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf die absolute Veränderung der 6-MWD in Woche 26 im Vergleich zum Basiswert.

Sensitivitätsanalysen

Um die Robustheit der Ergebnisse des Endpunkts „Absolute Veränderung der 6-MWD“ einschätzen zu können, wurden verschiedene Sensitivitätsanalysen durchgeführt. Diese sind in der folgenden Tabelle 4-63 dargestellt.

In einer ersten Sensitivitätsanalyse wurden diejenigen Patienten (5,1 % der FAS-Population) aus der Analyse ausgeschlossen, bei denen der 6-MWT in Woche 26 nicht am Talspiegel der Wirkstoffkonzentration durchgeführt wurde. Der hierbei zwischen den Behandlungsgruppen ermittelte Unterschied in der medianen Veränderung zum Basiswert von 11,0 m [99 % KI: 0 m; 22 m] war konsistent zum Ergebnis der Hauptanalyse.

In einer weiteren Sensitivitätsanalyse wurde die absolute Veränderung der 6-MWD in der PPS-Population ermittelt. Diese Sensitivitätsanalyse ergab einen Behandlungseffekt von Selexipag im Vergleich zu Placebo von 10,0 m [99 % KI: −2 m; 21 m] und war ebenfalls konsistent zur Hauptanalyse.

Die Robustheit der Ersetzungsstrategie wurde durch weitere Sensitivitätsanalysen überprüft, in denen der Imputationsalgorithmus variiert wurde. Der beobachtete Therapieeffekt war hierbei über sämtliche Imputationsalgorithmen hinweg sowie auch zur Hauptanalyse konsistent.

Wurden nur Patienten in die Auswertung des Endpunkts eingeschlossen, die von den Prüfärzten als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden, so ergab sich ein Behandlungseffekt von Selexipag im Vergleich zu Placebo von 12,0 m [99 % KI: 1 m; 23 m].

Insgesamt gesehen bestätigen sämtliche der durchgeführten Sensitivitätsanalysen das Ergebnis der Hauptanalyse des Endpunkts „Absolute Veränderung der 6-MWD“ und unterstreichen damit die Robustheit der Ergebnisse dieses Endpunkts.

Tabelle 4-63: Sensitivitätsanalysen für den Endpunkt „Absolute Veränderung der 6-MWD“

Sensitivitätsanalyse	**Behandlungseffekt von Selexipag gegenüber Placebo** Median [99 % KI][1], p-Wert (Wilcoxon-Mann-Whitney)[2]
Analyse unter Ausschluss der Patienten mit nicht am Talspiegel durchgeführtem 6-MWT	
Analyse unter Ausschluss der Patienten, bei denen der 6-MWT nicht am Talspiegel der Wirkstoffkonzentration durchgeführt wurde.	11,0 [0; 22] p=0,0086
Per-protocol-Analyse	
Analyse in der PPS-Population und unter Ausschluss der Patienten, bei denen der 6-MWT nicht am Tiefpunkt der Wirkstoff-konzentration durchgeführt wurde.	10,0 [-2; 21] p=0,0207
Analyse mit alternativen Imputationsalgorithmen	
Regel 1 wurde wie definiert angewandt, Regel 2 wurde nicht angewandt. Fehlende Werte wurden nicht ersetzt.	6,0 [-3; 15] p=0,0279
Regel 1 wurde wie definiert angewandt, Regel 2 wurde durch die folgende Regel ersetzt: Fehlende Werte wurden durch das 25. Perzentil der beobachteten 6-MWD-Werte (unabhängig von der Behandlungsgruppe) ersetzt.	9,0 [1; 18] p=0,0035
Regel 1 wurde wie definiert angewandt, Regel 2 wurde durch die folgende Regel ersetzt: Fehlende Werte wurden durch die letzte verfügbare 6-MWD (ggf. Basiswert) ersetzt.	12,0 [4; 21] p=0,0001
Weitere Sensitivitätsanalysen	
Analyse nur der Patienten, die als Patienten mit „konsistenter“ oder „typischer“ PAH identifiziert wurden.	12,0 [1; 23] p=0,0041

[1]) Nach der Hodges-Lehmann-Methode

[2]) Aus nicht-parametrischer ANCOVA mit 6-MWD-Basiswert als Kovariate

Sofern nicht anderweitig angegeben, wurden fehlende Werte nach den gleichen Regeln ersetzt wie in der Hauptanalyse:

- Regel 1: Für Patienten, die den 6-MWT aus PAH-bezogenen Gründen nicht durchführen konnten oder verstorben waren, wurden fehlende Werte durch 0 m ersetzt.
- Regel 2: Anderweitig fehlende Werte wurde durch den zweitschwersten Rangwert in Woche 26 ersetzt (in der FAS-Population: 10 m).

Weitere Erläuterungen zu den Sensitivitätsanalysen finden sich in Abschnitt 9.8.4.1.3 des Studienberichts. Sofern nicht anderweitig angegeben, ist die FAS-Population Grundlage der Analyse.

Die 6-MWD ist in Metern (m) angegeben.

KI – Konfidenzintervall, PPS – *per protocol set*, FAS – *full analysis set*, 6-MWT – 6-Minuten-Gehtest, ANCOVA – Kovarianzanalyse

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nicht-stratifizierten Analyse des Endpunkts „Absolute Veränderung der 6-MWD“ mit Analysen verglichen, die nach geographischer Region, PAH-Ätiologie, WHO-/NYHA-Klasse am Basiswert bzw. PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. Das Ergebnis der Analysen war konsistent über alle stratifizierten Analysen sowie zur nicht-stratifizierten Analyse (siehe Tabelle 4-64). In allen Analysen war der Unterschied zwischen den Behandlungsgruppen signifikant.

Tabelle 4-64: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Absolute Veränderung der 6-MWD“

Modell	Behandlungseffekt von Selexipag gegenüber Placebo
Nicht-stratifiziertes Modell mit 6-MWD-Basiswert als Kovariate[3]	p=0,0027
Modell mit 6-MWD-Basiswert als Kovariate, stratifiziert nach Region[1,4]	p=0,0045
Modell mit 6-MWD-Basiswert als Kovariate, stratifiziert nach PAH-Ätiologie[2,4]	p=0,0086
Modell mit 6-MWD-Basiswert als Kovariate, stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I-II/III-IV)[4]	p=0,0108
Modell mit 6-MWD-Basiswert als Kovariate, stratifiziert nach PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)[4]	p=0,0051

[1]) Asien, Osteuropa, Lateinamerika, Nordamerika, Westeuropa, Australien
[2]) (Idiopathisch, angeboren, Drogen/Arzneimittel- oder Toxin-induziert), assoziiert mit HIV-Infektion, assoziiert mit Bindegewebserkrankung, assoziiert mit angeborenem Herzfehler mit korrigiertem Links-Rechts-Shunt.
[3]) p-Wert aus einseitigem Wilcoxon-Mann-Whitney-Test
[4]) p-Wert aus zweiseitigem van Elteren-Test
Grundlage der Analyse war die FAS-Population.
6-MWD – 6-Minuten-Gehstrecke, FAS – *full analysis set*

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Tabelle 4-65 zeigt in einer Übersicht die absolute Veränderung der 6-MWD zu verschiedenen Studienzeitpunkten (einschließlich der EOS-Untersuchung) im Vergleich zum Basiswert. Mit Ausnahme von Monat 36, in dem nur die Werte von vergleichsweise wenigen Patienten in die Analyse eingingen, war die mediane absolute Veränderung der 6-MWD (gemessen zum

Talspiegel) zu allen untersuchten Zeitpunkten im Selexipag-Arm höher als im Placebo-Arm (vgl. auch Abbildung 4-27).

Tabelle 4-65: Analyse des Endpunkts „Absolute Veränderung der 6-MWD“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

Intervention	**Selexipag N=574**			**Placebo N=582**		
Veränderung der 6-MWD im Vergleich zum Basiswert (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	574	358,5	76,3	582	348,0	83,2
Veränderung gegenüber Basiswert						
Woche 8	529	8,5	49,6	547	1,5	57,5
Woche 16	495	10,6	53,7	507	0,0	68,3
Woche 26	460	15,3	59,3	446	4,9	65,5
Monat 12	397	17,4	67,2	387	-0,6	77,5
Monat 18	297	17,5	72,9	279	1,3	74,2
Monat 24	204	10,4	77,0	183	8,2	80,7
Monat 30	137	17,2	78,0	122	18,6	77,4
Monat 36	75	7,8	72,1	61	11,3	83,0
EOS-Untersuchung	474	-11,1	95,1	499	-34,0	101,6
Veränderung der 6-MWD im Vergleich zum Basiswert (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	574	376	90; 482	582	369	50; 515
Veränderung gegenüber Basiswert						
Woche 8	529	8,0	-300; 195	547	7,0	-405; 301
Woche 16	495	10,0	-237; 190	507	4,0	-400; 280
Woche 26	460	16,0	-313; 260	446	6,0	-372; 262
Monat 12	397	16,0	-384; 235	387	5,0	-421; 274
Monat 18	297	18,0	-417; 200	279	3,0	-393; 242
Monat 24	204	18,5	-297; 198	183	5,0	-440; 294
Monat 30	137	26,0	-413; 212	122	13,5	-326; 276
Monat 36	75	6,0	-180; 200	61	15,0	-333; 262
EOS-Untersuchung	474	3,0	-417; 222	499	-12,0	-440; 262

Bei Patienten, bei denen der 6-MWT nicht am Tiefpunkt der Wirkstoffkonzentration durchgeführt wurde, wurde dennoch dieser Wert für die Analyse verwendet. Bei Patienten, die den 6-MWT an geplanten Untersuchungsterminen aufgrund PAH-bezogener Gründe nicht durchführen konnten, wurden fehlende Werte mit 0 ersetzt (gilt nicht für EOS-Untersuchung).

Die 6-MWD ist in Metern (m) angegeben. Grundlage der Analyse war die FAS-Population.

6-MWD – 6-Minuten-Gehstrecke, 6-MWT – 6-Minuten-Gehtest, MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max – Maximalwert, EOS – *end-of-study*, FAS – *full analysis set*

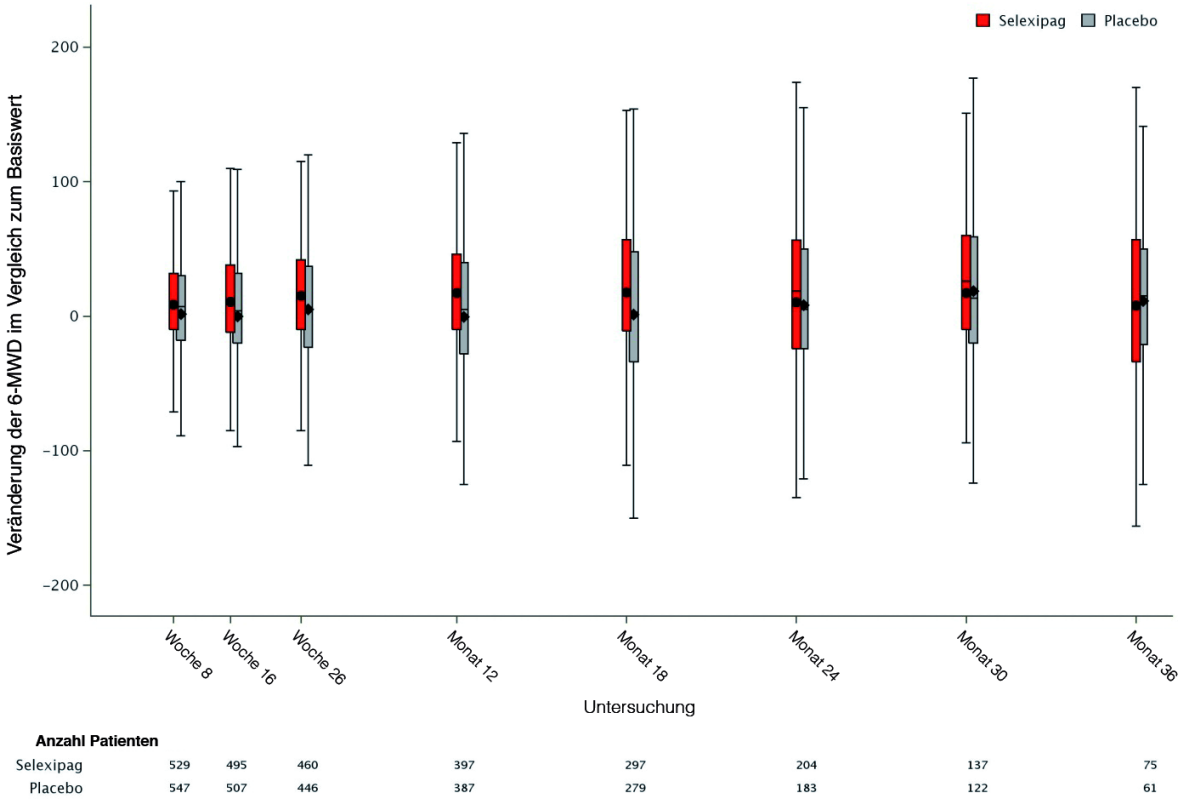


Abbildung 4-14: Absolute Veränderung der 6-MWD (reguläre Untersuchungstermine im Vergleich zum Basiswert)
Die Erhebung der 6-MWD erfolgte jeweils zum Talspiegel der Selexipag-Konzentration. Auswertungspopulation war die FAS-Population. FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen

Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-27). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.9 Ausbleibende Verschlechterung der WHO-/NYHA-Klasse

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-66: Operationalisierung von „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Anteil der Patienten mit ausbleibender Verschlechterung der WHO-/NYHA-Klasse in Woche 26 im Vergleich zum Basiswert. Als Verschlechterung der WHO-/NYHA-Klasse wurde jegliche Veränderung in eine höhere WHO-/NYHA-Klasse im Vergleich zum Zeitpunkt der Basiswertuntersuchung gezählt. Patienten, die zu diesem Zeitpunkt in der WHO-/NYHA-Klasse IV waren, wurden vollständig aus der Analyse ausgeschlossen, da sie sich per Definition nicht mehr verschlechtern konnten. *Ersetzungsregeln* Fehlende Werte in Woche 26 wurden als Verschlechterung interpretiert und als solche in in der Hauptanalyse berücksichtigt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-67: Bewertung des Verzerrungspotentials für „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotential auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotential Endpunkt
GRIPHON	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-68: Ergebnisse für „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON							**Ausgewertete Population: FAS[1]**					
Veränderung der WHO-/NYHA-Klasse												
Intervention	Basis-wert	n	Woche 26									
			I		II		III		IV		Fehlend[2]	
			n	%	n	%	N	%	n	%	n	%
Selexipag (N=571)	I	4	4	100	-	-	-	-	-	-	-	-
	II	274	7	2,6	207	75,5	21	7,7	-	-	39	14,2
	III	293	1	0,3	67	22,9	158	53,9	10	3,4	57	19,5
	gesamt	571	12	2,1	274	48,0	179	31,3	10	1,8	96	16,8
Placebo (N=574)	I	5	4	80,0	-	-	-	-	-	-	1	20,0
	II	255	7	2,7	197	77,3	16	6,3	-	-	35	13,7
	III	314	4	1,3	37	11,8	181	57,6	14	4,5	78	24,8
	gesamt	574	15	2,6	234	40,8	197	34,3	14	2,4	114	19,9

Anteil Patienten mit Verschlechterung/ohne Verschlechterung der WHO-/NYHA-Klasse					
Intervention	WHO-/NYHA-Klasse	Woche 26			
		Ausbleibende Verschlechterung		Verschlechterung	
		n	%	n	%
Selexipag (N=571)	I	4	100,0	-	-
	II	214	78,1	60	21,9
	III	226	77,1	67	22,9
	gesamt	**444**	**77,8**	**127**	**22,2**
Placebo (N=574)	I	4	80,0	1	20,0
	II	204	80,0	51	20,0
	III	222	70,7	92	29,3
	gesamt	**430**	**74,9**	**144**	**25,1**

Behandlungseffekt (Selexipag gegenüber Placebo)		
p-Wert (Cochran-Mantel-Haenszel)[3]		0,2843
OR [99 % KI]		1,16 [0,81; 1,66]
p-Wert (Breslow-Day)		0,1916
RR [95 % KI][4,5]		1,04 [0,97; 1,11]
RD [95 % KI][4,6] (Prozentpunkte)		2,8 [-2,1; 7,8]

[1]) Patienten in WHO-/NYHA-Klasse IV wurden aus der Analyse ausgeschlossen, da diese sich per Definition nicht weiter verschlechtern konnten.
[2]) Fehlende Werte wurden als Verschlechterung interpretiert und in die Analyse eingeschlossen
[3]) Stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung
[4]) Nach den Vorgaben des Nutzendossiers *post hoc* berechnet.
[5]) Dargestellt wird das nicht-adjustierte, übliche RR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte mithilfe der Farrington-Manning-Methode berechnet.
[6]) Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode.
OR – Odds Ratio, RR – Risk Ratio, RD – absolute Risikodifferenz, KI – Konfidenzintervall, FAS – *full analysis set*

Hauptanalyse

Die Analyse der ausbleibenden Verschlechterung der WHO-/NYHA-Klasse wurde auf Grundlage der FAS-Population durchgeführt. Patienten, die zum Zeitpunkt der Basiswerterhebung bereits in WHO-/NYHA-Klasse IV waren, wurden aus der Analyse ausgeschlossen, da sie sich in der WHO-/NYHA-Klasse nicht weiter verschlechtern konnten.

Der Anteil der Patienten, die in Woche 26 ohne Verschlechterung der WHO-/NYHA-Klasse im Vergleich zum Basiswert waren, betrug 77,8 % im Selexipag-Arm und 74,9 % im Placebo-Arm. Dies entspricht einer numerischen Verbesserung des absoluten Risikos um 2,8 Prozentpunkte (RD [95 % KI]: 2,8 %-P. [−2,1 %-P.; 7,8 %-P.] und einer OR [99 % KI] von 1,16 [0,81; 1,66]. Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (zweiseitiger Cochran-Mantel-Haenszel-Test, p=0,2843).

Die Tatsache, dass kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen vorlag, ist möglicherweise durch die Tatsache erklärbar, dass der Großteil der Patienten zum Zeitpunkt der Basiswerterhebung bereits mit einem oder zwei PAH-spezifischen Arzneimitteln vorbehandelt wurde (vgl. Tabelle 4-13) und daher voraussichtlich mit hoher Wahrscheinlichkeit zu diesem Zeitpunkt klinisch stabil war. Gestützt wird diese Hypothese durch die Ergebnisse der Subgruppenanalyse der Patienten mit oder ohne PAH-spezifische Begleittherapie, in der bei Patienten ohne PAH-spezifische Begleittherapie ein sehr viel stärker ausgeprägter Behandlungseffekt beobachtet wurde als bei Patienten mit Begleittherapie (siehe Abbildung 4-28).

Nebenanalysen

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf die ausbleibende Verschlechterung der WHO-/NYHA-Klasse in Woche 26 im Vergleich zum Basiswert.

Sensitivitätsanalysen

Um die Robustheit der Ergebnisse des Endpunkts „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ einschätzen zu können, wurden verschiedene Sensitivitätsanalysen durchgeführt. Diese sind in der folgenden Tabelle 4-69 dargestellt.

In einer ersten Sensitivitätsanalyse wurde der Einfluss der Strategie zur Ersetzung fehlender Werte überprüft. Hierzu wurde die Analyse der ausbleibenden Verschlechterung der WHO-/NYHA-Klasse ohne Ersetzung fehlender Werte durchgeführt. Diese Analyse ergab ein OR [99% KI] von 1,01 [0,51; 2,00], der Unterschied zwischen den Behandlungsgruppen war wie bereits in der Hauptanalyse statistisch nicht signifikant (p=0,9767).

In einer weiteren Sensitivitätsanalyse wurde die ausbleibende Verschlechterung der WHO-/NYHA-Klasse in der PPS-Population ermittelt. Diese Sensitivitätsanalyse ergab eine OR [99% KI] von 1,11 [0,76; 1,60] und damit einen zur Hauptanalyse konsistenten leichten numerischen Vorteil von Selexipag gegenüber Placebo. Der Unterschied zwischen den Behandlungsgruppen war jedoch auch in der PPS-Population statistisch nicht signifikant.

Tabelle 4-69: Sensitivitätsanalysen für den Endpunkt „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“

Sensitivitätsanalyse	**Behandlungseffekt von Selexipag gegenüber Placebo** p-Wert (Cochran-Mantel-Haenszel)[1] OR [99% KI] p-Wert (Breslow-Day)
Analyse ohne Ersetzung fehlender Werte	
Analyse ohne Ersetzung fehlender Werte (in der FAS-Population)	p=0,9767 1,01 [0,51; 2,00] p=0,2894
Per-protocol-Analyse	
Analyse in der PPS-Population	p=0,4473 1,11 [0,76; 1,60] p=0,0692

[1]) Stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswert-erhebung

Weitere Erläuterungen zu den Sensitivitätsanalysen finden sich in Abschnitt 9.8.4.1.3 des Studienberichts.

OR – Odds Ratio, KI – Konfidenzintervall, PPS – *per protocol set*, FAS – *full analysis set*

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nach WHO-/NYHA-Klasse am Basiswert stratifizierten Analyse des Endpunkts „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ mit Analysen verglichen, die zusätzlich nach geographischer Region, PAH-Ätiologie bzw. PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. Die Effektschätzer des Behandlungseffekts, die in diesen Analysen bestimmt wurden, waren konsistent über alle zusätzlich nach weiteren Faktoren als der WHO-/NYHA-Klasse stratifizierten Analysen sowie zur Analyse, die ausschließlich nach WHO-/NYHA-Klasse stratifiziert worden war (siehe Tabelle 4-70). Der Effektschätzer (OR) des Therapieeffekts von Selexipag gegenüber Placebo bewegte sich hierbei konsistent in einem Bereich von 1,14 bis 1,17. Ein statistisch signifikanter Unterschied wurde in keinem der stratifizierten Modelle beobachtet.

Tabelle 4-70: Stratifizierte Analysen des Endpunkts „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“

Modell	**Behandlungseffekt von Selexipag gegenüber Placebo** p-Wert (Cochran-Mantel-Haenszel)[1] OR [99 % KI]
Modell stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung	p=0,2843 1,16 [0,81; 1,66]
Modell stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung und nach Region[2]	p=0,2875 1,16 [0,81; 1,67]
Modell stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung und nach PAH-Ätiologie[3]	p=0,3373 1,14 [0,80; 1,64]
Modell stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung und PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)[4]	p=0,2751 1,17 [0,81; 1,67]

[1]) Stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung

[2]) Asien, Osteuropa, Lateinamerika, Nordamerika, Westeuropa, Australien

[3]) (Idiopathisch, angeboren, Drogen/Arzneimittel- oder Toxin-induziert), assoziiert mit HIV-Infektion, assoziiert mit Bindegewebserkrankung, assoziiert mit angeborenem Herzfehler mit korrigiertem Links-Rechts-Shunt.

Grundlage der Analyse war die FAS-Population.

OR – Odds Ratio, KI – Konfidenzintervall, FAS – *full analysis set*

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Tabelle 4-71 zeigt in einer Übersicht den Anteil der Patienten mit Verbesserung, Verschlechterung oder keiner Veränderung der WHO-/NYHA-Klasse zu verschiedenen Studienzeitpunkten (einschließlich der EOS-Untersuchung) im Vergleich zum Basiswert. Von Woche 8 an war der Anteil der Patienten, die sich in ihrer WHO-/NYHA-Klasse im Vergleich zum Basiswert nicht verschlechtert hatten, zu allen Zeitpunkten im Selexipag-Arm höher als im Placebo-Arm (vgl. auch Abbildung 4-15 und Abbildung 4-16). Lediglich in Woche 26 und in Monat 24 war der Anteil in beiden Gruppen ausgeglichen. Konsistent hierzu war auch der Anteil der Patienten mit einer Verbesserung der WHO-/NYHA-Klasse zu allen untersuchten Zeitpunkten im Selexipag-Arm höher als im Placebo-Arm (siehe Abbildung 4-16). Am ausgeprägtesten war der Unterschied zwischen den Behandlungsgruppen zum Zeitpunkt der EOS-Untersuchung. Während sich zu diesem Zeitpunkt 64,3 % der Patienten im Placebo-Arm nicht verschlechtert hatten, waren im Selexipag-Arm deutlich mehr Patienten (73,9 %) ohne eine Verschlechterung der WHO-/NYHA-Klasse (fehlende Werte wurde in der Analyse dieses Zeitpunkt ersetzt; lag kein Wert vor, wurde davon ausgegangen, dass sich der Patient verschlechtert hatte). Ebenso war auch der Anteil der Patienten, die sich im Vergleich zum Basiswert in ihrer WHO-/NYHA-Klasse verbessert hatten, im Selexipag-Arm höher als im Placebo-Arm (9,6 % unter Selexipag gegenüber 8,7 % unter Placebo).

Tabelle 4-71: Analyse des Endpunkts „Ausbleibende Verschlechterung der WHO-/NYHA-Klasse“ zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

	Selexipag (N=574)		**Placebo (N=582)**	
	n	% [95 % KI][1]	n	% [95 % KI][1]
Woche 4				
Nicht fehlende Werte	562		566	
Fehlende Werte	12		16	
Verbesserung	33	5,9 [4,1; 8,1]	27	4,8 [3,2; 6,9]
Verschlechterung	15	2,7 [1,5; 4,4]	12	2,1 [1,1; 3,7]
Keine Veränderung	514	91,5 [88,8; 93,6]	527	93,1 [90,7; 95,1]
Woche 8				
Nicht fehlende Werte	539		554	
Fehlende Werte	35		28	
Verbesserung	61	11,3 [8,8; 14,3]	43	7,8 [5,7; 10,3]
Verschlechterung	20	3,7 [2,3; 5,7]	28	5,1 [3,4; 7,2]
Keine Veränderung	458	85,0 [81,7; 87,9]	483	87,2 [84,1; 89,9]
Woche 16				
Nicht fehlende Werte	504		512	
Fehlende Werte	70		70	
Verbesserung	70	13,9 [11,0; 17,2]	50	9,8 [7,3; 12,7]
Verschlechterung	25	5,0 [3,2; 7,2]	37	7,2 [5,1; 9,8]
Keine Veränderung	409	81,2 [77,5; 84,5]	425	83,0 [79,5; 86,2]
Woche 26				
Nicht fehlende Werte	478		464	
Fehlende Werte	96		118	
Verbesserung	77	16,1 [12,9; 19,7]	50	10,8 [8,1; 14,0]
Verschlechterung	31	6,5 [4,4; 9,1]	30	6,5 [4,4; 9,1]
Keine Veränderung	370	77,4 [73,4; 81,1]	384	82,8 [79,0; 86,1]
Monat 12				
Nicht fehlende Werte	435		416	
Fehlende Werte	139		166	
Verbesserung	79	18,2 [14,7; 22,1]	58	13,9 [10,8; 17,6]
Verschlechterung	27	6,2 [4,1; 8,9]	53	12,7 [9,7; 16,3]
Keine Veränderung	329	75,6 [71,3; 79,6]	305	73,3 [68,8; 77,5]

Monat 18				
Nicht fehlende Werte	334		311	
Fehlende Werte	240		271	
Verbesserung	67	20,1 [15,9; 24,8]	43	13,8 [10,2; 18,2]
Verschlechterung	27	8,1 [5,4; 11,5]	34	10,9 [7,7; 14,9]
Keine Veränderung	240	71,9 [66,7; 76,6]	234	75,2 [70,1; 79,9]
Monat 24				
Nicht fehlende Werte	241		215	
Fehlende Werte	333		367	
Verbesserung	45	18,7 [14,0; 24,2]	32	14,9 [10,4; 20,4]
Verschlechterung	29	12,0 [8,2; 16,8]	27	12,6 [8,4; 17,7]
Keine Veränderung	167	69,3 [63,1; 75,1]	156	72,6 [66,1; 78,4]
Monat 30				
Nicht fehlende Werte	170		163	
Fehlende Werte	404		419	
Verbesserung	32	18,8 [13,2; 25,5]	29	17,8 [12,3; 24,5]
Verschlechterung	11	6,5 [3,3; 11,3]	21	12,9 [8,2; 19,0]
Keine Veränderung	127	74,7 [67,5; 81,0]	113	69,3 [61,6; 76,3]
Monat 36				
Nicht fehlende Werte	94		101	
Fehlende Werte	480		481	
Verbesserung	25	26,6 [18,0; 36,7]	13	12,9 [7,0; 21,0]
Verschlechterung	9	9,6 [4,5; 17,4]	14	13,9 [7,8; 22,2]
Keine Veränderung	60	63,8 [53,3; 73,5]	74	73,3 [63,5; 81,6]
EOS-Untersuchung[2]				
Nicht fehlende Werte	571[3]		574[3]	
Fehlende Werte	-		-	
Verbesserung	55	9,6 [n. b.]	50	8,7 [n. b.]
Verschlechterung	149	26,1 [n. b.]	205	35,7 [n. b.]
Keine Veränderung	367	64,3 [n. b.]	319	55,6 [n. b.]

[1]) Konfidenzintervall nach Clopper-Pearson

[2]) Die Berechnung der Werte zum Zeitpunkt der EOS-Untersuchung wurde *post hoc* durchgeführt.

[3]) Bei fehlenden Werten zur WHO-/NYHA-Klasse wurde davon ausgegangen, dass sich der Patient in seiner Klasse verschlechtert hatte.

Grundlage der Analyse war die FAS-Population.

KI – Konfidenzintervall, EOS – *end-of-study*, FAS – *full analysis set*, n. b. – nicht berechnet

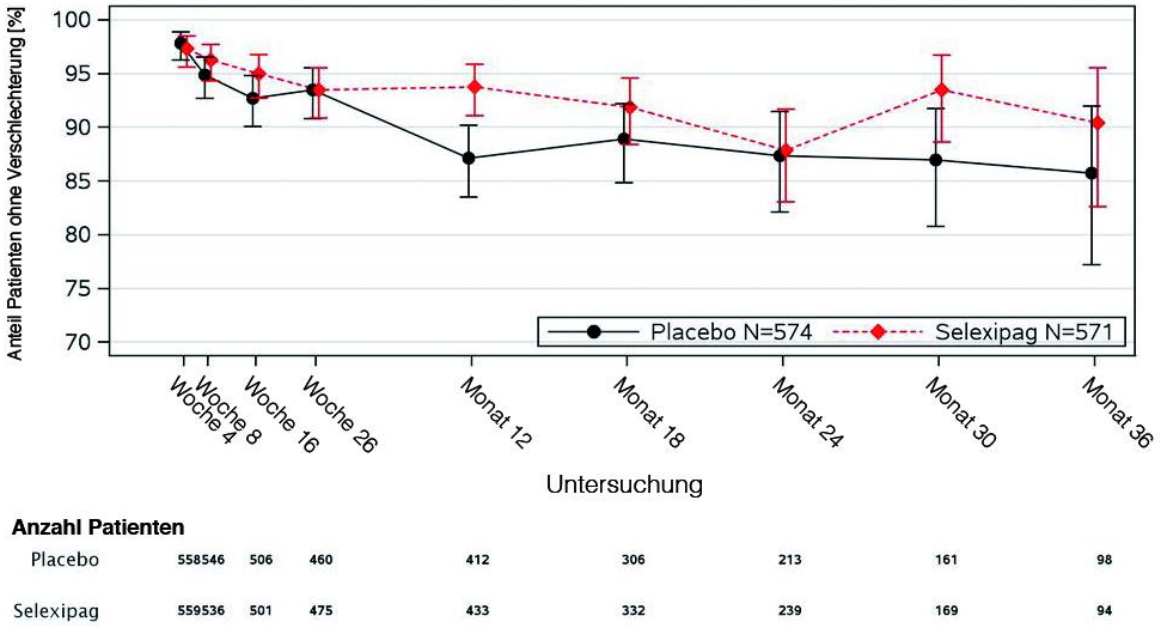


Abbildung 4-15: Ausbleibende Verschlechterung der WHO-/NYHA-Klasse (reguläre Untersuchungstermine im Vergleich zum Basiswert)
Fehlerbalken geben das 95 % KI (nach Clopper-Pearson) an. Analysierte Population: FAS (ohne Patienten in WHO-/NYHA-Klasse IV). FAS – *full analysis set*, KI – Konfidenzintervall

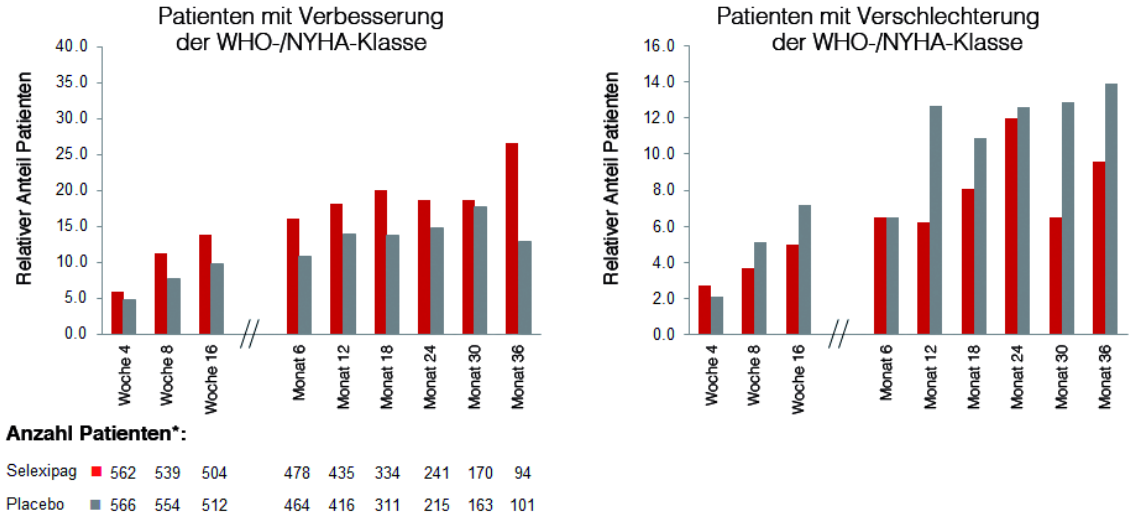


Abbildung 4-16: Patienten mit Verbesserung bzw. Verschlechterung der WHO-/NYHA-Klasse zu regulären Untersuchungsterminen
*Patienten mit fehlenden Werten. Auswertungspopulation war die FAS-Population (ohne Patienten in WHO-/NYHA-Klasse IV). FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-28). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa ähnlich ausgeprägt wie in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation nicht von einem unterschiedlichen Therapieffekt für Deutschland auszugehen ist.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.10 Absolute Veränderung des Borg-Dyspnoe-Index

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-72: Operationalisierung von „Absolute Veränderung des Borg-Dyspnoe-Index“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Absolute Veränderung des Borg-Dyspnoe-Index vom Basiswert bis zu Woche 26. Die Erhebung des Borg-Dyspnoe-Index erfolgte nach einem von Actelion vorgegebenen standardisierten Protokoll und im unmittelbaren Anschluss an die Erhebung der 6-MWD und damit ebenfalls zum Zeitpunkt des Talspiegels der Selexipag-Konzentration (siehe Tabelle 4-60). Zur Erhebung des Borg-Dyspnoe-Index wurden die Patienten im Anschluss an die Durchführung des 6-MWT gebeten, den maximalen Grad ihrer Atemnot während des 6-MWT anhand der Borg-Skala einzuschätzen. Die Borg-Skala ist aus den folgenden zwölf Atemnot-Kategorien zusammengesetzt: 0 – keine 0,5 – sehr, sehr leicht (gerade bemerkbar) 1 – sehr leicht 2 – leicht 3 – moderat 4 – etwas schwer 5 – schwer 6 7 – sehr schwer 8 9 10 – sehr, sehr schwer (maximal) Die absolute Veränderung wurde bestimmt als Differenz des Borg-Dyspnoe-Index-Werts in Woche 26 minus dem Borg-Dyspnoe-Index-Wert zum Zeitpunkt der Basiswerterhebung. Für den Woche-26-Wert wurde der Borg-Dyspnoe-Index-Wert herangezogen, der innerhalb des Woche-26-Zeitfensters (Studientage 148-271) bestimmt wurde. In den Fällen, in denen mehr als ein Borg-Dyspnoe-Index-Wert innerhalb des Woche-26-Zeitfensters vorlag, wurde derjenige Wert ausgewählt, der näher am nominalen Wert für Woche 26 (Studientag 182) lag.
	Ersetzungsregeln Es wurde keine Ersetzung fehlender Werte vorgenommen, da diese *a priori* nicht definiert war. Dargestellt sind die Ergebnisse derjenigen Patienten, für die Werte des Borg-Dyspnoe-Index zum Zeitpunkt der Basiswerterhebung und in Woche 26 vorlagen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-73: Bewertung des Verzerrungspotentials für „Absolute Veränderung des Borg-Dyspnoe-Index“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet.

Da es sich beim Borg-Dyspnoe-Index um einen explorativen Endpunkt handelte, war eine Ersetzung fehlender Werte nicht *a priori* geplant und wurde daher auch nicht durchgeführt. Es wurden ausschließlich diejenigen Patienten in die Analyse miteinbezogen, für die Werte zum Zeitpunkt der Basiswerterhebung und in Woche 26 vorlagen. Das ITT-Prinzip ist damit nicht vollständig umgesetzt.

Ansonsten konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Absolute Veränderung des Borg-Dyspnoe-Index" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-74: Ergebnisse für „Absolute Veränderung des Borg-Dyspnoe-Index" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: FAS**					
Intervention	Selexipag (N=574)			Placebo (N=582)		
Veränderung des Borg-Dyspnoe-Index vom Basiswert bis zu Woche 26 (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	457	3,53	2,11	443	3,48	1,98
Woche 26	457	3,33	1,92	443	3,61	2,13
Veränderung Woche 26 gegenüber Basiswert	**457**	**-0,21**	**1,66**	**443**	**0,13**	**1,64**
Veränderung des Borg-Dyspnoe-Index vom Basiswert bis zu Woche 26 (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	457	3,0	0,0; 10,0	443	3,0	0,0; 10,0
Woche 26	457	3,0	0,0; 10,0	443	3,0	0,0; 10,0
Veränderung Woche 26 gegenüber Basiswert	**457**	**0,0**	**-7,0; 6,0**	**443**	**0,0**	**-6,0; 7,0**
Behandlungseffekt (Veränderung des Borg-Dyspnoe-Index vom Basiswert bis Woche 26)						
Punktschätzer [95 % KI][1]				0,0 [0,0; 0,0]		
p-Wert (Wilcoxon-Mann-Whitney)[2]				0,0061		

[1]) Nach der Hodges-Lehmann-Methode
[2]) Aus nicht-parametrischer ANCOVA mit Borg-Dyspnoe-Index-Basiswert als Kovariate
MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max –Maximalwert, KI – Konfidenzintervall, ANCOVA – Kovarianzanalyse, FAS – *full analysis set*

Hauptanalyse

Die Analyse der absoluten Veränderung des Borg-Dyspnoe-Index wurde auf Grundlage der FAS-Population durchgeführt. Es wurde keine Ersetzung fehlender Werte vorgenommen. Ausgewertet wurden ausschließlich die Daten derjenigen Patienten, für die Werte zum Zeitpunkt der Basiswerterhebung und in Woche 26 vorlagen.

Die mediane absolute Veränderung des Borg-Dyspnoe-Index vom Basiswert bis Woche 26 betrug 0,0 bei Patienten in der Selexipag-Gruppe und 0,0 bei Patienten in der Placebo-Gruppe. Der Behandlungseffekt [95 % KI] von Selexipag im Vergleich zu Placebo betrug 0,0 [0,0; 0,0], als p-Wert für den Unterschied zwischen den Behandlungsgruppen ergab sich ein Wert von p=0,0061 (Wilcoxon-Mann-Whitney-Test aus nicht-parametrischer ANCOVA mit Borg-Dyspnoe-Index-Basiswert als Kovariate).

Nebenanalysen

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Tabelle 4-75 zeigt in einer Übersicht die absolute Veränderung des Borg-Dyspnoe-Index zu verschiedenen Studienzeitpunkten (einschließlich der EOS-Untersuchung) im Vergleich zum Basiswert. Der mittlere absolute Wert des Borg-Dyspnoe-Index war zu allen Zeitpunkten bei den Patienten im Selexipag-Arm numerisch niedriger als bei den Patienten im Placebo-Arm (vgl. auch Abbildung 4-17).

Tabelle 4-75: Analyse des Borg-Dyspnoe-Index zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

Intervention	**Selexipag N=574**			**Placebo N=582**		
Borg-Dyspnoe-Index zu verschiedenen Zeitpunkten (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	574	3,62	2,12	582	3,74	2,11
Wert an regulären Untersuchungsterminen						
Woche 8	528	3,45	2,09	541	3,66	2,17
Woche 16	492	3,36	2,08	504	3,70	2,18
Woche 26	457	3,33	1,92	443	3,61	2,13
Monat 12	395	3,39	2,10	383	3,54	2,13
Monat 18	295	3,32	2,01	274	3,54	2,16
Monat 24	201	3,33	2,03	181	3,65	2,20
Monat 30	136	3,34	1,93	122	3,38	2,00
Monat 36	72	3,63	1,85	60	3,83	2,15
EOS-Untersuchung	451	3,72	2,19	473	4,13	2,31
Borg-Dyspnoe-Index zu verschiedenen Zeitpunkten (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	574	3,0	0,0; 10,0	582	3,0	0,0; 10,0
Wert an regulären Untersuchungsterminen						
Woche 8	528	3,0	0,0; 10,0	541	3,0	0,0; 10,0
Woche 16	492	3,0	0,0; 10,0	504	3,0	0,0; 10,0
Woche 26	457	3,0	0,0; 10,0	443	3,0	0,0; 10,0
Monat 12	395	3,0	0,0; 10,0	383	3,0	0,0; 10,0
Monat 18	295	3,0	0,0; 10,0	274	3,0	0,0; 10,0
Monat 24	201	3,0	0,0; 10,0	181	3,0	0,0; 10,0
Monat 30	136	3,0	0,0; 10,0	122	3,0	0,0; 10,0
Monat 36	72	3,0	0,0; 8,0	60	3,0	0,5; 9,0
EOS-Untersuchung	451	3,0	0,0; 10,0	473	4,0	0,0; 10,0

Grundlage der Analyse war die FAS-Population.

MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max – Maximalwert, EOS – *end-of-study*, FAS – *full analysis set*

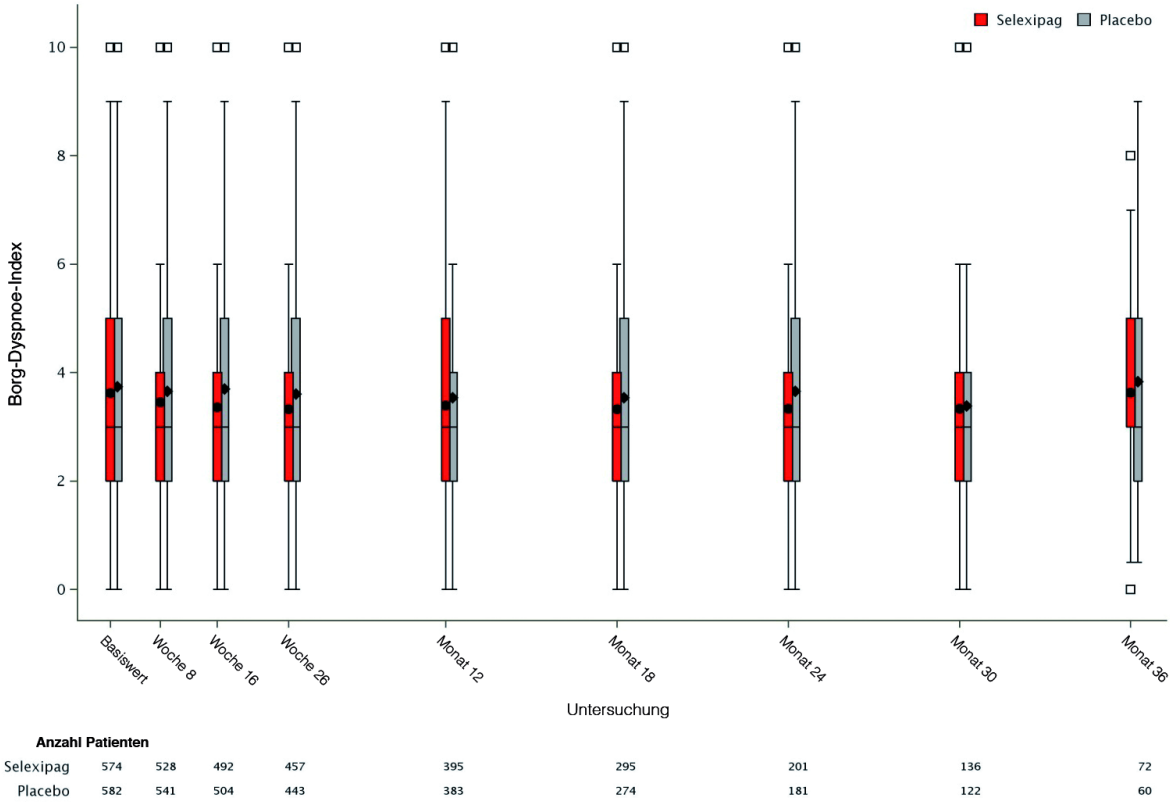


Abbildung 4-17: Borg-Dyspnoe-Index an regulären Untersuchungsterminen
Die Erhebung des Borg-Dyspnoe-Index erfolgte jeweils zum Talspiegel der Selexipag-Konzentration. Auswertungspopulaiton war die FAS-Population. FAS – *full analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen

Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-29). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effekt-modifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.11 Absolute Veränderung des CAMPHOR „Symptome“-Scores

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-76: Operationalisierung von „Absolute Veränderung des CAMPHOR „Symptome“-Scores“

Studie	Operationalisierung

GRIPHON	*Definition des Endpunkts*
	Absolute Veränderung im „Symptome“-Score des CAMPHOR-Fragebogens vom Basiswert bis zu Woche 26.
	Der CAMPHOR „Symptome“-Score ist definiert als die Summe der Items 1-25 des CAMPHOR-Fragebogens. Die Spannweite des „Symptome“-Scores reicht von 0 (gut) bis 25 (schlecht). Die Erhebung des CAMPHOR-Fragebogens, die Berechnungen von Scores und Subskalen sowie der Umgang mit fehlenden Werten erfolgten nach den Vorgaben des Anbieters (Galen Research).
	Als absolute Veränderung des CAMPHOR „Symptome“-Scores vom Basiswert bis zu Woche 26 wurde die Differenz vom Wert in Woche 26 minus Wert zum Zeitpunkt der Basiswerterhebung ausgewertet.
	Auswertungspopulation für die Hauptanalyse war die QAS-Population. Hierbei handelt es sich um diejenigen Patienten in der FAS-Population, in deren Zentren Versionen des Fragebogens vorlagen, die für die jeweilige Sprache validiert waren.
	Ersetzungsregeln
	Es wurden in der Auswertung nur diejenigen Patienten der QAS-Population berücksichtigt, bei denen ein Basiswert des CAMPHOR „Symptome“-Scores vorlag.
	Bei Patienten, bei denen der Basiswert des CAMPHOR „Symptome“-Scores, aber nicht der Wert für Woche 26 vorlag, wurden die folgenden Imputationsalgorithmen angewendet:
	• Regel 1: Für Patienten, die vor Tag 271 verstorben waren und für die keine Untersuchungsergebnisse im Woche-26-Zeitfenster vorlagen, wurde der fehlende Score-Wert durch 25 (schlechtestmöglicher Wert) ersetzt.
	• Regel 2 (sofern Regel 1 nicht anwendbar war): Für Patienten, die vor Tag 271 die Behandlung mit der Studienmedikation abgebrochen hatten und für die keine Untersuchungsergebnisse im Woche-26-Zeitfenster vorlagen:
	  o der letzte verfügbare Wert der WHO-/NYHA-Klasse war I oder II: der fehlende Wert wurde ersetzt durch das 88. Perzentil der CAMPHOR „Symptome“-Score-Werte aus Woche 26, die für Patienten in den WHO-/NYHA-Klassen I oder II beobachtet worden waren (unabhängig von der Behandlungsgruppe).
	  o der letzte verfügbare Wert der WHO-/NYHA-Klasse war III oder IV: der fehlende Wert wurde ersetzt durch das 88. Perzentil der CAMPHOR „Symptome“-Score-Werte aus Woche 26, die für Patienten in den WHO-/NYHA-Klassen III oder IV beobachtet worden waren (unabhängig von der Behandlungsgruppe).
	• Regel 3 (sofern Regel 1 und 2 nicht anwendbar waren): der fehlende Wert wird durch den letzten Wert des CAMPHOR „Symptome“-Scores ersetzt, der für die Zeit von vor Woche 26 verfügbar war.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-77: Bewertung des Verzerrungspotentials für „Absolute Veränderung des CAMPHOR „Symptome“-Scores“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt „Absolute Veränderung des CAMPHOR „Symptome“-Scores“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-78: Ergebnisse für „Absolute Veränderung des CAMPHOR ‚Symptome‘-Scores“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: QAS**					
Intervention	Selexipag (N=241)			Placebo (N=250)		
Veränderung des CAMPHOR „Symptome“-Scores vom Basiswert bis zu Woche 26 (Mittelwert)						
	n	MW	SD	N	MW	SD
Basiswert	239	10,44	5,64	240	11,70	5,60
Woche 26	239	10,67	7,00	240	12,11	6,89
Veränderung Woche 26 gegenüber Basiswert	**239**	**0,23**	**5,69**	**240**	**0,41**	**6,19**
Veränderung des CAMPHOR „Symptome“-Scores vom Basiswert bis zu Woche 26 (Median)						
	n	Median	Min; Max	N	Median	Min; Max
Basiswert	239	10,0	0,0; 24,0	240	11,0	0,0; 25,0
Woche 26	239	10,0	0,0; 25,0	240	12,0	0,0; 25,0
Veränderung Woche 26 gegenüber Basiswert	**239**	**-1,0**	**-16,0; 20,0**	**240**	**0,0**	**-18,0; 21,0**
Imputation fehlender Werte für Woche 26						
	n	%		n	%	
Insgesamt ersetzt	53	22,0		57	22,8	
Grund: Tod[1]	7	2,9		12	4,8	
Therapieabbruch vor Woche 26[2]	39	16,2		41	16,4	
Werte der Woche-26-Untersuchung fehlen[3]	7	2,9		4	1,6	
Behandlungseffekt (Veränderung des CAMPHOR „Symptome“-Scores vom Basiswert bis Woche 26)						
Punktschätzer [99 % KI][4]				0,0 [-1,0; 1,0]		
p-Wert (Wilcoxon-Mann-Whitney)[5]				0,2185		

[1]) Fehlende Werte wurden mit einem Score von 25 ersetzt

[2]) Fehlende Werte von Patienten in den WHO-/NYHA-Klassen I/II wurden durch das 88. Perzentil der insgesamt bei Patienten in den WHO-/NYHA-Klassen I/II beobachteten Werte ersetzt. Bei Patienten in den WHO-/NYHA-Klassen III/IV wurde entsprechend das 88. Perzentil der insgesamt bei Patienten in den WHO-/NYHA-Klassen III/IV beobachteten Werte verwendet

[3]) Fehlende Werte wurden durch den letzten verfügbaren Wert ersetzt (Patienten, die an der Woche-26-Untersuchung teilgenommen hatten, für die aber Werte fehlten, und Patienten, die an der Untersuchung nicht teilgenommen hatten)

[4]) Nach der Hodges-Lehmann-Methode

[5]) Aus nicht-parametrischer ANCOVA mit CAMPHOR „Symptome“-Score als Kovariate

MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max –Maximalwert, KI – Konfidenzintervall, ANCOVA – Kovarianzanalyse, QAS – *quality of life analysis set*

Hauptanalyse

Die Analyse der absoluten Veränderung des CAMPHOR „Symptome"-Scores wurde auf Grundlage der QAS-Population durchgeführt. Diese Population besteht aus denjenigen Patienten der FAS-Population, für die eine validierte Übersetzung des Fragebogens in ihrer Sprache vorliegt. Es wurden jedoch nur die Patienten aus der QAS-Population zur Auswertung herangezogen, bei denen ein Basiswert des CAMPHOR „Symptome"-Scores vorlag.

In Bezug auf die Erhebung der Lebensqualität wurde vom G-BA im Beratungsgespräch zu Selexipag gefordert, dass „*alle Anstrengungen unternommen werden [sollten], eine möglichst hohe Rücklaufquote (≥70 %) bezogen auf die Intention-to-Treat(ITT)-Population* [hier: QAS-Population] *zu erzielen*" [3]. Von den Patienten in der QAS-Population lagen für 2 Patienten im Selexipag-Arm und für 10 Patienten im Placebo-Arm keine Ergebnisse zum Basiswert des CAMPHOR „Symptome"-Scores vor; sie wurden daher nicht in die Auswertung eingeschlossen. Weiterhin mussten bei den Patienten, bei denen ein Basiswert vorlag, 22,0 % bzw. 22,8 % der Werte in Woche 26 in den beiden Therapiegruppen ersetzt werden. Es wurden demnach Rücklaufquoten von 77,1 %[33] (Selexipag) bzw. 73,2 %[34] (Placebo) erreicht.

Die mediane absolute Veränderung des CAMPHOR „Symptome"-Scores vom Basiswert bis Woche 26 betrug −1,0 (d. h. sie verbesserte sich) bei Patienten in der Selexipag-Gruppe und 0,0 in der Placebo-Gruppe. Der Behandlungseffekt [99 % KI] von Selexipag im Vergleich zu Placebo betrug 0,0 [−1,0; 1,0], der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Wilcoxon-Mann-Whitney-Test aus nicht-parametrischer ANCOVA mit CAMPHOR „Symptome"-Score-Basiswert als Kovariate: p=0,2185).

Die Tatsache, dass in dieser Analyse kein Effekt von Selexipag auf die Lebensqualität der Patienten nachgewiesen werden konnte, kann möglicherweise zum Teil durch die Verwendung des CAMPHOR-Fragebogens erklärt werden. Wie bereits in Abschnitt 4.2.5.2 beschrieben, wurden in die Validierungsstudie des CAMPHOR-Fragebogens nicht ausschließlich Patienten mit PAH, sondern auch Patienten mit PH eingeschlossen. Der CAMPHOR-Fragebogen könnte damit möglicherweise nicht ausreichend spezifisch sein, um Probleme bezüglich der Lebensqualität zu addressieren, die für PAH-Patienten wichtig sind [89]. Als Folge hieraus sollten die Ergebnisse des CAMPHOR, die in Studien mit einer reinen PAH-Population beobachtet werden – wie z. B. der Population in GRIPHON also –, mit Vorsicht interpretiert werden [89].

Ein weiterer Faktor, der zur Erklärung des nicht nachweisbaren Therapieeffektes beitragen könnte, ist die Tatsache, dass der CAMPHOR-Fragebogen in GRIPHON nur in der Subgruppe der Patienten untersucht wurde, für deren Sprachen/Länder der Fragebogen validiert wurde. Hierbei handelt es sich im Großen und Ganzen um westliche Länder, in denen die Wahrscheinlichkeit hoch war, dass Patienten zum Zeitpunkt des Beginns der Studie

[33] 53 von 241 Patienten
[34] 67 von 250 Patienten

bereits mit zwei PAH-spezifischen Arzneimitteln behandelt wurden und damit mit geringerer Wahrscheinlichkeit symptomatisch waren als therapienaive Patienten. Weiterhin mussten sich die Patienten, um eingeschlossen werden zu können, zum Zeitpunkt des Einschlusses in die Studie in einem stabilen Zustand befinden. Beide Sachverhalte sind potentiell aber dazu in der Lage, die Wahrscheinlichkeit zum Nachweis einer Verbesserung der Lebensqualität zu vermindern.

Insgesamt fällt jedoch auch positiv auf, dass sich die Selexipag-Patienten in ihrer Lebensqualität nicht verschlechterten. Dies ist insbesondere daher bemerkenswert, da aufgrund der Dosierungsstrategie und der damit fast zwangsläufig einhergehenden Prostacyclin-assoziierten Nebenwirkungen tendenziell eher mit einer Abnahme der Lebensqualität zu rechnen gewesen wäre.

Nebenanalysen

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf die absolute Veränderung des CAMPHOR „Symptome“-Scores in Woche 26 im Vergleich zum Basiswert.

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nicht-stratifizierten Analyse des Endpunkts „Absolute Veränderung des CAMPHOR ‚Symptome‘-Scores“ mit Analysen verglichen, die nach WHO-/NYHA-Klasse bzw. PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. In keiner der Analysen war der Unterschied zwischen den Behandlungsgruppen statistisch signifikant (siehe Tabelle 4-79).

Tabelle 4-79: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Absolute Veränderung des CAMPHOR ‚Symptome'-Scores"

Modell	Behandlungseffekt von Selexipag gegenüber Placebo
Nicht-stratifiziertes Modell mit Basiswert des CAMPHOR „Symptome"-Scores als Kovariate[1]	p=0,2185
Modell mit Basiswert des CAMPHOR „Symptome"-Scores als Kovariate, stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I-II/III-IV)[2]	p=0,4224
Modell mit Basiswert des CAMPHOR „Symptome"-Scores als Kovariate, stratifiziert nach PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)[2]	p=0,4562

[1]) p-Wert aus einseitigem Wilcoxon-Mann-Whitney-Test
[2]) p-Wert aus zweiseitigem van-Elteren-Test
Grundlage der Analyse war die QAS-Population.
QAS – *quality of life analysis set*

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Tabelle 4-80 zeigt in einer Übersicht die absolute Veränderung des CAMPHOR „Symptome"-Scores zu verschiedenen Studienzeitpunkten (einschließlich der EOS-Untersuchung) im Vergleich zum Basiswert. Zu keinem der Zeitpunkte ist ein wesentlicher Unterschied zwischen den Behandlungsgruppen in Bezug auf die Veränderung des CAMPHOR „Symptome"-Scores zu beobachten (vgl. auch Abbildung 4-18).

Tabelle 4-80: Analyse des Endpunkts „Absolute Veränderung des CAMPHOR ‚Symptome'-Scores" zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

Intervention	**Selexipag N=241**			**Placebo N=240**		
Veränderung des CAMPHOR „Symptome"-Scores im Vergleich zum Basiswert (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	239	10,44	5,64	240	11,70	5,60
Veränderung gegenüber Basiswert						
Woche 16	205	-0,89	4,05	205	-1,26	4,29
Woche 26	186	-1,37	4,48	183	-1,44	4,53
Monat 12	161	-1,59	4,88	159	-1,27	5,17
Monat 24	75	-1,26	5,39	75	-2,04	4,22
EOS-Untersuchung	208	-0,10	5,19	201	-0,62	5,37
Veränderung des CAMPHOR „Symptome"-Scores im Vergleich zum Basiswert (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	239	10,0	0,0; 24,0	240	11,0	0,0; 25,0
Veränderung gegenüber Basiswert						
Woche 16	205	-1,0	-14,0; 15,0	205	-1,0	-16,0; 13,0
Woche 26	186	-1,0	-16,0; 10,0	183	-1,0	-18,0; 11,0
Monat 12	161	-1,0	-19,0; 13,0	159	-1,0	-20,0; 20,0
Monat 24	75	-1,0	-17,0; 15,0	75	-1,0	-14,0; 10,0
EOS-Untersuchung	208	0,0	-15,7; 17,0	201	-0,4	-17,0; 18,0

Grundlage der Analyse war die QAS-Population.

MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max – Maximalwert, EOS – *end-of-study*, QAS – *quality of life analysis set*

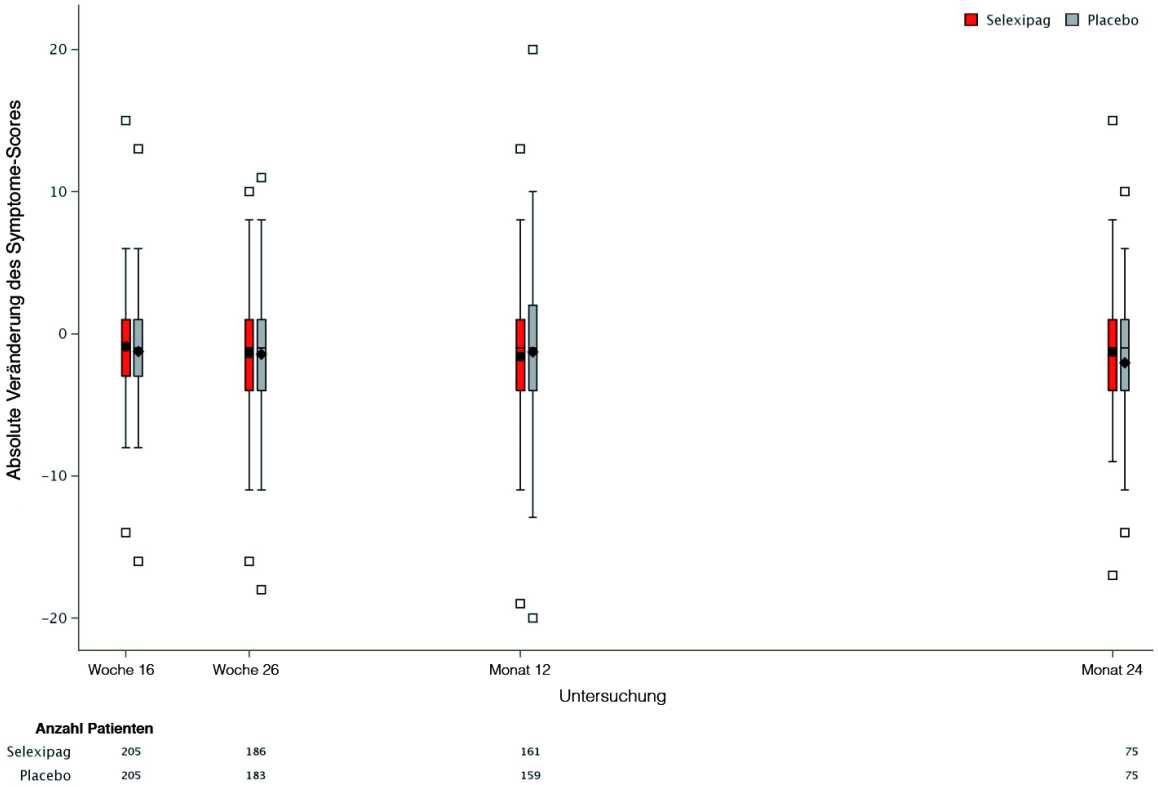


Abbildung 4-18: Absolute Veränderung des CAMPHOR „Symptome“-Scores (reguläre Untersuchungstermine im Vergleich zum Basiswert)
Auswertungspopulation war die QAS-Population. QAS – *quality of life analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen

Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-30). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effekt-modifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.12 Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-81: Operationalisierung von „Absolute Veränderung der ‚Kurzatmigkeit‘-Subskala des CAMPHOR ‚Symptome‘-Scores“

Studie	Operationalisierung

GRIPHON	*Definition des Endpunkts*
	Absolute Veränderung in der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores vom Basiswert bis zu Woche 26.
	Die „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores ist definiert als die Summe der „Kurzatmigkeit“-Items 11-18 des CAMPHOR-Fragebogens. Die Spannweite der „Kurzatmigkeit“-Subskala reicht von 0 (gut) bis 8 (schlecht). Die Erhebung des CAMPHOR-Fragebogens, die Berechnungen von Scores und Subskalen sowie der Umgang mit fehlenden Werten erfolgten nach den Vorgaben des Anbieters (Galen Research).
	Als absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores vom Basiswert bis zu Woche 26 wurde die Differenz vom Wert in Woche 26 minus Wert zum Zeitpunkt der Basiswerterhebung ausgewertet.
	Auswertungspopulation für die Hauptanalyse war die QAS-Population. Hierbei handelt es sich um diejenigen Patienten in der FAS-Population, in deren Zentren Versionen des Fragebogens vorlagen, die für die jeweilige Sprache validiert waren.
	Ersetzungsregeln
	Es wurden in der Auswertung nur diejenigen Patienten der QAS-Population berücksichtigt, bei denen ein Basiswert des CAMPHOR „Symptome“-Scores vorlag.
	Bei Patienten, bei denen der Basiswert der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores, aber nicht der Wert für Woche 26 vorlag, wurden die folgenden Imputationsalgorithmen angewendet:
	• Regel 1: Für Patienten, die vor Tag 271 verstorben waren und für die keine Untersuchungsergebnisse im Woche-26-Zeitfenster vorlagen, wurde der fehlende Score-Wert durch 8 (schlechtestmöglicher Wert) ersetzt.
	• Regel 2 (sofern Regel 1 nicht anwendbar war): Für Patienten, die vor Tag 271 die Behandlung mit der Studienmedikation abgebrochen hatten und für die keine Untersuchungsergebnisse im Woche-26-Zeitfenster vorlagen:
	  o der letzte verfügbare Wert der WHO-/NYHA-Klasse war I oder II: der fehlende Wert wurde ersetzt durch das 88. Perzentil der Werte der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores aus Woche 26, die für Patienten in den WHO-/NYHA-Klassen I oder II beobachtet worden waren (unabhängig von der Behandlungsgruppe).
	  o der letzte verfügbare Wert der WHO-/NYHA-Klasse war III oder IV: der fehlende Wert wurde ersetzt durch das 88. Perzentil der Werte der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores aus Woche 26, die für Patienten in den WHO-/NYHA-Klassen III oder IV beobachtet worden waren (unabhängig von der Behandlungsgruppe).
	• Regel 3 (sofern Regel 1 und 2 nicht anwendbar waren): der fehlende Wert wurde durch den letzten Wert der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores ersetzt, der vor Woche 26 verfügbar war.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-82: Bewertung des Verzerrungspotentials für „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.

Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Stellen Sie die Ergebnisse für den Endpunkt Absolute Veränderung der „Kurzatmigkeit"-Subskala des CAMPHOR „Symptome"-Scores" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-83: Ergebnisse für „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: QAS**					
Intervention	Selexipag (N=241)			Placebo (N=250)		
Veränderung der „Kurzatmigkeit"-Subskala des CAMPHOR „Symptome"-Scores vom Basiswert bis zu Woche 26 (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	239	3,83	1,79	240	4,18	1,84
Woche 26	239	3,78	2,21	240	4,23	2,19
Veränderung Woche 26 gegenüber Basiswert	**239**	**-0,05**	**1,92**	**240**	**0,05**	**1,91**
Veränderung der „Kurzatmigkeit"-Subskala des CAMPHOR „Symptome"-Scores vom Basiswert bis zu Woche 26 (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	239	4,0	0,0; 8,0	240	4,0	0,0; 8,0
Woche 26	239	4,0	0,0; 8,0	240	4,0	0,0; 8,0
Veränderung Woche 26 gegenüber Basiswert	**239**	**0,0**	**-6,0; 5,0**	**240**	**0,0**	**-5,0; 6,0**
Imputation fehlender Werte für Woche 26						
	n		%	n		%
Insgesamt ersetzt	53		22,0	57		22,8
Grund: Tod[1]	7		2,9	12		4,8
Therapieabbruch vor Woche 26[2]	39		16,2	41		16,4
Werte der Woche-26-Untersuchung fehlen[3]	7		2,9	4		1,6

Behandlungseffekt (Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores vom Basiswert bis Woche 26)		
Punktschätzer [99 % KI][4]		0,0 [-0,4; 0,0]
p-Wert (Wilcoxon-Mann-Whitney)[5]		0,1700

[1]) Fehlende Werte wurden mit einem Score von 8 ersetzt

[2]) Fehlende Werte von Patienten in den WHO-/NYHA-Klassen I/II wurden durch das 88. Perzentil der insgesamt bei Patienten in den WHO-/NYHA-Klassen I/II beobachteten Werte ersetzt. Bei Patienten in den WHO-/NYHA-Klassen III/IV wurde entsprechend das 88. Perzentil der insgesamt bei Patienten in den WHO-/NYHA-Klassen III/IV beobachteten Werte verwendet

[3]) Fehlende Werte wurden durch den letzten verfügbaren Wert ersetzt (Patienten, die an der Woche-26-Untersuchung teilgenommen hatten, für die aber Werte fehlten, und Patienten, die an der Untersuchung nicht teilgenommen hatten)

[4]) Nach der Hodges-Lehmann-Methode

[5]) Aus nicht-parametrischer ANCOVA mit CAMPHOR „Symptome“-Score als Kovariate

MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max –Maximalwert, KI – Konfidenzintervall, ANCOVA – Kovarianzanalyse, QAS – *quality of life analysis set*

Hauptanalyse

Die Analyse der absoluten Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores wurde auf Grundlage der QAS-Population durchgeführt. Diese Population besteht aus denjenigen Patienten der FAS-Population, für die eine validierte Übersetzung des Fragebogens in ihrer Sprache vorliegt. Es wurden jedoch nur die Patienten aus der QAS-Population zur Auswertung herangezogen, bei denen ein Basiswert des CAMPHOR „Symptome“-Scores vorlag.

In Bezug auf die Erhebung der Lebensqualität wurde vom G-BA im Beratungsgespräch zu Selexipag gefordert, dass „*alle Anstrengungen unternommen werden [sollten], eine möglichst hohe Rücklaufquote (≥70 %) bezogen auf die Intention-to-Treat(ITT)-Population* [hier: QAS-Population] *zu erzielen*“ [3]. Von den Patienten in der QAS-Population lagen für 2 Patienten im Selexipag-Arm und für 10 Patienten im Placebo-Arm keine Ergebnisse zum Basiswert des CAMPHOR „Symptome“-Scores vor; sie wurden daher nicht in die Auswertung eingeschlossen. Weiterhin mussten bei den Patienten, bei denen ein Basiswert vorlag, 22,0 % bzw. 22,8 % der Werte in Woche 26 in den beiden Therapiegruppen ersetzt werden. Es wurden demnach Rücklaufquoten von 77,1 %[35] (Selexipag) bzw. 73,2 %[36] (Placebo) erreicht.

Die mediane absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores vom Basiswert bis Woche 26 betrug 0,0 bei Patienten in der Selexipag-Gruppe und 0,0 in der Placebo-Gruppe. Der Behandlungseffekt [99 % KI] von Selexipag im Vergleich zu Placebo betrug 0,0 [−0,4; 0,0], der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Wilcoxon-Mann-Whitney-Test aus

[35] 53 von 241 Patienten.
[36] 67 von 250 Patienten.

nicht-parametrischer ANCOVA mit dem Basiswert der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores als Kovariate: p=0,1700).

Die möglichen Gründe, die zur Erklärung des fehlenden statistisch signifikanten Unterschieds zwischen den Behandlungsgruppen beitragen könnten, wurden bereits im vorangegangenen Abschnitt für den CAMPHOR „Symptome“-Score diskutiert (vgl. Abschnitt 4.3.1.3.1.11). Die dort genannten Punkte sind auch für die „Kurzatmigkeit“-Subskala relevant.

Nebenanalysen

Sofern nicht anders angegeben, beziehen sich die Nebenanalysen auf die absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores in Woche 26 im Vergleich zum Basiswert.

Stratifizierte Modelle

In einer supportiven Analyse wurden die Ergebnisse der nicht-stratifizierten Analyse des Endpunkts „Absolute Veränderung der ‚Kurzatmigkeit‘-Subskala des CAMPHOR ‚Symptome‘-Scores“ mit Analysen verglichen, die nach WHO-/NYHA-Klasse bzw. PAH-spezifischer Begleittherapie am Basiswert stratifiziert wurden. In keiner der Analysen war der Unterschied zwischen den Behandlungsgruppen statistisch signifikant (siehe Tabelle 4-84).

Tabelle 4-84: Nicht-stratifizierte und stratifizierte Analysen des Endpunkts „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores"

Modell	Behandlungseffekt von Selexipag gegenüber Placebo
Nicht-stratifiziertes Modell mit Basiswert der „Atemlosigkeit"-Subskala des CAMPHOR „Symptome"-Scores als Kovariate[1]	p=0,1700
Nicht-stratifiziertes Modell mit Basiswert der „Atemlosigkeit"-Subskala des CAMPHOR „Symptome"-Scores als Kovariate, stratifiziert nach WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I-II/III-IV)[2]	p=0,2740
Nicht-stratifiziertes Modell mit Basiswert der „Atemlosigkeit"-Subskala des CAMPHOR „Symptome"-Scores als Kovariate, stratifiziert nach PAH-spezifischer Begleittherapie zum Zeitpunkt der Basiswerterhebung (ja/nein)[2]	p=0,2695

[1]) p-Wert aus einseitigem Wilcoxon-Mann-Whitney-Test
[2]) p-Wert aus zweiseitigem van-Elteren-Test
Grundlage der Analyse war die QAS-Population.
QAS – *quality of life analysis set*

Andere Zeitpunkte/Zeiträume als für die Hauptanalyse

Tabelle 4-85 zeigt in einer Übersicht die absolute Veränderung der „Kurzatmigkeit"-Subskala des CAMPHOR „Symptome"-Scores zu verschiedenen Studienzeitpunkten (einschließlich der EOS-Untersuchung) im Vergleich zum Basiswert. Zu keinem der Zeitpunkte ist ein wesentlicher Unterschied zwischen den Behandlungsgruppen in Bezug auf die Veränderung der „Kurzatmigkeit"-Subskala des CAMPHOR „Symptome"-Scores zu beobachten (vgl. auch Abbildung 4-19).

Tabelle 4-85: Analyse des Endpunkts „Absolute Veränderung der ‚Kurzatmigkeit'-Subskala des CAMPHOR ‚Symptome'-Scores" zu anderen Zeitpunkten als dem für die Hauptanalyse verwendeten

Intervention	**Selexipag N=241**			**Placebo N=240**		
Veränderung der „Atemlosigkeit"-Subskala des CAMPHOR „Symptome"-Scores im Vergleich zum Basiswert (Mittelwert)						
	n	MW	SD	n	MW	SD
Basiswert	239	3,83	1,79	240	4,18	1,84
Veränderung gegenüber Basiswert						
Woche 16	205	-0,39	1,56	206	-0,34	1,38
Woche 26	186	-0,47	1,72	183	-0,42	1,55
Monat 12	160	-0,55	1,74	159	-0,36	1,67
Monat 24	76	-0,50	1,80	76	-0,51	1,59
EOS-Untersuchung	208	-0,19	1,76	203	-0,15	1,74
Veränderung der „Atemlosigkeit"-Subskala des CAMPHOR „Symptome"-Scores im Vergleich zum Basiswert (Median)						
	n	Median	Min; Max	n	Median	Min; Max
Basiswert	239	4,0	0,0; 8,0	240	4,0	0,0; 8,0
Veränderung gegenüber Basiswert						
Woche 16	205	0,0	-6,0; 6,0	206	0,0	-4,0; 5,0
Woche 26	186	0,0	-6,0; 4,0	183	0,0	-5,0; 3,0
Monat 12	160	0,0	-6,0; 6,0	159	0,0	-5,0; 6,0
Monat 24	76	0,0	-4,0; 5,0	76	0,0	-5,0; 3,0
EOS-Untersuchung	208	0,0	-7,0; 7,0	203	0,0	-5,0; 6,0

Grundlage der Analyse war die QAS-Population.

MW – Mittelwert, SD – Standardabweichung, Min – Minimalwert, Max – Maximalwert, EOS – *end-of-study*, QAS – *quality of life analysis set*

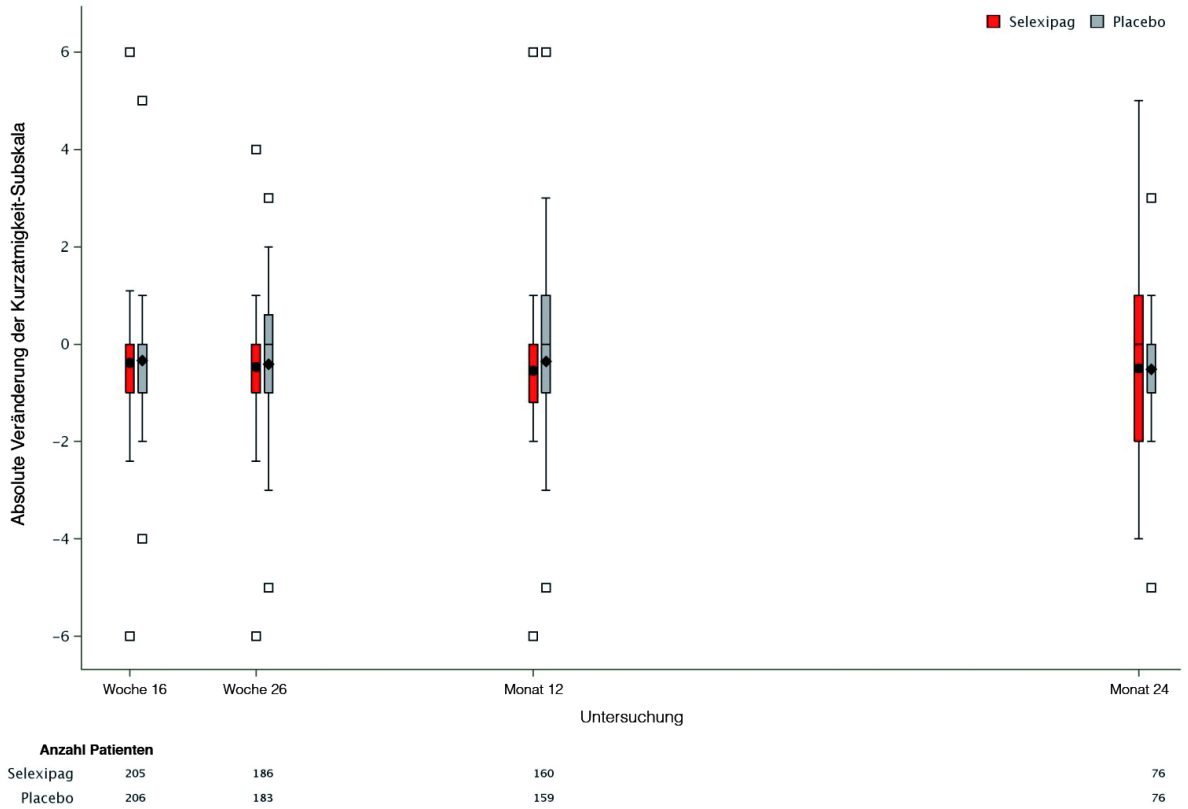


Abbildung 4-19: Absolute Veränderung der „Kurzatmigkeit"-Subskala des CAMPHOR „Symptome"-Scores (reguläre Untersuchungstermine im Vergleich zum Basiswert)
Auswertungspopulation war die QAS-Population. QAS – *quality of life analysis set*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen

Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-31). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa stärker ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effekt-modifikation tendenziell von einer Unterschätzung des Therapieffekts für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.13 Unerwünschte Ereignisse

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-86: Operationalisierung von „Unerwünschte Ereignisse“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Anteil der Patienten mit mindestens einem unerwünschten Ereignis im Behandlungszeitraum. Unerwünschte Ereignisse (UE) waren definiert als jede unerwünschte Veränderung des Basiswertzustands des Patienten, d. h. jedes nachteilige oder ungewollte Anzeichen (einschließlich abnormer Laborergebnisse), jedes Symptom oder jede Krankheit, die im Studienverlauf auftrat, unabhängig davon, ob sie in ursächlicher Beziehung zur Studienmedikation steht oder nicht. Als therapiebedingtes (*treatment-emergent*) UE wurde jegliches UE bezeichnet, das zeitgleich mit der Anwendung der Studienmedikation auftrat (d. h. von Beginn der Einnahme der Studienmedikation bis zu 7 Tage nach Ende der Einnahme oder bis zur Einnahme der ersten Dosis der Studienmedikation in der offenen Extensionsstudie AC065A303 [je nachdem, was zu erst eintrat], unabhängig davon, ob sein Auftreten in ursächlicher Beziehung zur Studienmedikation stand). Folgende Ereignisse wurden als UE gezählt: • Verschlechterung einer bestehenden Erkrankung • Anstieg in Frequenz oder Intensität einer bereits bestehenden episodischen Erkrankung oder bereits bestehender Beschwerden • Erkrankungen oder Beschwerden, die während der Anwendung der Studienmedikation bemerkt oder diagnostiziert wurden, auch wenn sie bereits vor dem Beginn der Studie vorgelegen haben könnten • Chronische, persistierende Erkrankungen oder Symptome, die bereits zum Zeitpunkt der Basiswerterhebung vorlagen, aber sich nach dem Beginn der Studie verschlechterten • Fehlende Wirksamkeit in der Notfallbehandlung einer lebensbedrohlichen Krankheit • Ungewöhnliche Ergebnisse, z. B. EKG-Befunde, mussten als UE berichtet werden, wenn sie einen klinisch signifikanten Befund darstellten, der zum Zeitpunkt der Basiswertbestimmung nicht vorlag oder sich während der Studie verschlechterte • Ungewöhnliche Ergebnisse von Laborbefunden mussten als UE berichtet werden, wenn sie einen klinisch signifikanten Befund darstellten – unabhängig von einer möglichen Symptomatik –, der zum Zeitpunkt der Basiswertbestimmung nicht vorlag, sich während der Studie verschlechterte oder zu einer Reduktion, zeitweiligen Unterbrechung oder permanenten Unterbrechung der Studienmedikation führte Folgende Ereignisse wurden <u>nicht</u> als UE gezählt: • Medizinische oder chirurgische Prozeduren, z. B. chirurgische Eingriffe, Endoskopie, Zahnextraktion oder Transfusion. Im Gegensatz dazu wurde jedoch das Ereignis, das zur jeweiligen Prozedur führte, als UE gezählt. Handelte es sich hierbei um ein schwerwiegendes Ereignis, musste es als SUE beschrieben werden. • Vorbestehende Erkrankungen oder Beschwerden, die sich nicht verschlechterten. • Situationen, in denen keine unerwünschte Veränderung aufgetreten war, z. B. Krankenhausaufenthalte für selbstgewählte kosmetische Eingriffe oder aus sozialen oder Zweckmäßigkeitsgründen. • Überdosierung der Studienmedikation oder der Begleittherapie, die nicht zu Anzeichen oder Symptomen führten.

Unerwünschte Ereignisse wurden zusätzlich nach System-Organ-Klasse (*system organ class*, SOC) und bevorzugter Bezeichnung (*preferred term*, PT) der MedDRA-Klassifikation, Version 16, zusammengefasst.

Einteilung nach Schweregraden

Die Intensität der UE wurde durch den Prüfarzt mittels einer dreistufigen Klassifikation als mild, moderat oder schwer bewertet. Erhöhte sich die Intensität eines UE während des Verlaufs der Studie, wurde lediglich die höchste Intensität berichtet. Verringerte sich die Intensität, wurde keine Veränderung der Schwere vorgenommen.

Mild

Das Ereignis ist möglicherweise für den Patienten wahrnehmbar, beeinflusst seine täglichen Aktivitäten aber nicht und erfordert üblicherweise keine Intervention.

Moderat

Das Ereignis führt möglicherweise zu Unbehagen beim Patienten, die Ausführung der täglichen Aktivitäten kann eingeschränkt sein, eine Intervention ist möglicherweise notwendig.

Schwer

Das Ereignis löst möglicherweise bemerkenswertes Unbehagen beim Patienten aus, schränkt ihn meist in seinen täglichen Aktivitäten ein und beeinträchtigt möglicherweise die Fortsetzung der Studie. Eine Behandlung oder Intervention ist üblicherweise notwendig.

Ein mildes, moderates oder schweres UE kann, muss aber nicht gleichzeitig schwerwiegend sein. Diese Begriffe werden verwendet, um die Intensität eines spezifischen Ereignisses zu beschreiben (wie z. B. bei mildem, moderatem oder schwerem Myokardinfarkt). Ein schweres Ereignis kann jedoch auch von relativ geringfügiger medizinischer Bedeutung sein (wie z. B. schwere Kopfschmerzen) und muss nicht notwendigerweise schwerwiegend sein. Fieber von 39 °C, das nicht als schwer betrachtet wird, kann sich aber per Definition zu einem schwerwiegenden Ereignis entwickeln, wenn es einen Krankenhausaufenthalt um einen Tag verlängert.

Unterteilung in Titrations- und Erhaltungsphase

Die Unterteilung der UE nach Titrations- und Erhaltungsphase erfolgte nach den folgenden Regeln:

- Titrationsphase:
  - Sämtliche UE, die an oder vor Tag 87 beginnen und an oder vor Tag 87 wieder enden.
  - Sämtliche UE, die an oder vor Tag 87 beginnen und sich über Tag 87 hinaus fortsetzen[1]
- Erhaltungsphase:
  - Sämtliche UE, die nach Tag 87 beginnen
  - Sämtliche UE, die an oder vor Tag 87 beginnen und sich über Tag 87 hinaus fortsetzen[1]

[1]) UE, deren zeitlicher Verlauf sich über Titrations- und Erhaltungsphase erstreckt, werden in beiden Phasen als UE berücksichtigt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-87: Bewertung des Verzerrungspotentials für „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es liegen jedoch sonstige das Verzerrungspotential beeinflussende Aspekte vor. So setzt die Auswertung der Ereignisse des Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen gemäß dem Cochrane Handbook for Systematic Reviews of Interventions jedoch voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden [83]. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer bzw. nicht mehr handhabbarer prostacyclinspezifischer Nebenwirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben.

Insgesamt ist das Verzerrungspotential für diesen Endpunkt als hoch anzusehen. Aufgrund der vergleichsweise längeren medianen Behandlungsdauer in der Selexipag-Gruppe (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchstens von einer Verzerrung der Ergebnisse zuungunsten von Selexipag auszugehen.

Stellen Sie die Ergebnisse für den Endpunkt „Unerwünschte Ereignisse" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-88: Ergebnisse für „Unerwünschte Ereignisse" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: SAF**			
Intervention	Selexipag N=575		Placebo N=577	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n	%	n	%
Patienten mit mindestens einem unerwünschten Ereignis	565	98,3	559	96,9
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität	282	49,0	260	45,1
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem unerwünschten Ereignis	1,01 [1,00; 1,03]	1,82 [0,83; 3,98]	1,4 [-0,4; 3,2]	0,1329
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität	1,09 [0,96; 1,23]	1,17 [0,93; 1,48]	4,0 [-1,8; 9,7]	0,2124
Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt. KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*				

Hauptanalyse

Die Analyse der unerwünschten Ereignisse (UE) wurde auf Grundlage der SAF-Population durchgeführt.

Insgesamt trat im Verlauf der Behandlungsphase von GRIPHON bei 98,3 % der Patienten im Selexipag-Arm und bei 96,9 % der Patienten im Placebo-Arm mindestens ein UE auf. Dies entspricht einer RR [95 % KI] im Vergleich von Selexipag zu Placebo von 1,01 [1,00; 1,03] und einer absoluten RD [95 % KI] von 1,4 Prozentpunkten [−0,4 %-P.; 3,2 %-P.]. Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,1329).

Die hohe Zahl an UE, die während des Studienverlaufs von GRIPHON in den beiden Studiengruppen aufgetreten waren, ist aufgrund der langen Behandlungsdauer in der Studie (im Median 70,7 Wochen in der Selexipag-Gruppe und 63,7 Wochen in der Placebo-Gruppe)

nicht außergewöhnlich. Zudem sei hier an das grundsätzliche Studiendesign von GRIPHON erinnert, das in den ersten 12 Wochen die sukzessive Erhöhung der Dosierung der Studienmedikation bis zum Auftreten nicht mehr tolerierbarer UE vorsah, um damit die patientenindividuell höchste noch vertragene Dosierung festzustellen. Diese Vorgehensweise benachteiligt üblicherweise jedoch die aktive Therapie und führt dadurch zu einer Verzerrung der Ergebnisse in der Titrationsphase zuungunsten von Selexipag. Tatsächlich ist die Rate unerwünschter Ereignisse unter Selexipag nur in der Titrationsphase höher, während sie in der Erhaltungsphase ausgeglichen ist (vgl. Tabelle 4-89).

Neben der Gesamtrate der UE wurde auch der Anteil der Patienten ausgewertet, bei denen mindestens ein UE schwerer Intensität aufgetreten war. Als UE schwerer Intensität wurde ein UE dann eingestuft, wenn es ein bemerkenswertes Unbehagen beim Patienten auslöst, ihn meist bei seinen täglichen Aktivitäten einschränkt oder möglicherweise die Fortsetzung der Studie beeinträchtigt. Üblicherweise ist eine Behandlung oder eine Intervention im Zusammenhang mit schweren UE notwendig. Es ist zudem wichtig, darauf hinzuweisen, dass ein schweres UE auch von geringer klinischer Relevanz sein kann und nicht zwangsläufig mit einem schwerwiegenden UE gleichzusetzen ist.

Schwere UE traten in GRIPHON bei 49,0 % der Patienten im Selexipag-Arm auf und damit vergleichsweise häufiger als bei Patienten im Placebo-Arm (44,9 %). Dies entspricht einer numerischen Erhöhung des relativen Risikos von 9 % (RR [95 % KI]: 1,09 [0,96; 1,23]) und des absoluten Risikos von 4,0 Prozentpunkten (RD [95 % KI]: 4,0 %-P. [−1,8 %-P.; 9,7 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war jedoch statistisch nicht signifikant (Barnards exakter Test: p=0,2124).

Die folgende Tabelle 4-89 listet den Anteil der Patienten mit unerwünschtem Ereignis bzw. mit schwerem unerwünschten Ereignis in Abhängigkeit von der Studienphase auf, in der das entsprechende Ereignis aufgetreten war. Nicht eingeschlossen in diese Auswertung wurden Patienten, die bereits in der Titrationsphase die Behandlung beendet hatten.

Die Auswertung zeigt, dass – wie aufgrund des Studiendesigns zu erwarten war – unter Selexipag in der Titrationsphase ein größerer Anteil der Patienten ein UE (93,3 %) oder ein schweres UE (21,8 %) hatten als unter Placebo (76,8 % bzw. 10,4 %). In der Erhaltungsphase wiederum war der Anteil der Patienten mit einem UE oder einem schweren UE in beiden Gruppen ausgeglichen oder nahezu ausgeglichen (93,5 % vs. 93,1 % bzw. 42,4 % vs. 40,9 %).

Tabelle 4-89: Aufteilung der unerwünschten Ereignisse nach Titrations- und Erhaltungsphase

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag N=509		Placebo N=508	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n	%	n	%
Patienten mit mindestens einem unerwünschten Ereignis in der **Titrationsphase**	475	93,3	390	76,8
Patienten mit mindestens einem unerwünschten Ereignis in der **Erhaltungssphase**	476	93,5	473	93,1
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität in der **Titrationsphase**	111	21,8	53	10,4
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität in der **Erhaltungssphase**	216	42,4	208	40,9
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem unerwünschten Ereignis in der **Titrationsphase**	1,22 [1,15; 1,28]	4,23 [2,82; 6,33]	16,5 [12,3; 20,8]	<0,0001
Patienten mit mindestens einem unerwünschten Ereignis in der **Erhaltungsphase**	1,00 [0,97; 1,04]	1,07 [0,65; 1,75]	0,4 [-2,7; 3,5]	0,8588
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität in der **Titrationsphase**	2,09 [1,54; 2,83]	2,39 [1,68; 3,41]	11,4 [6,9; 15,8]	<0,0001
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität in der **Erhaltungssphase**	1,04 [0,90; 1,20]	1,06 [0,83; 1,36]	1,5 [-4,6; 7,6]	0,6830

[1]) Ausschließlich der Patienten, die die Therapie während der Titrationsphase abgebrochen hatten.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Nebenanalysen

Sensitivitätsanalyse unter Ausschluss der Ereignisse des primären Endpunkts

Die folgende Tabelle 4-90 zeigt die vom G-BA im Beratungsgespräch zu Selexipag [3] geforderten Sensitivitätsanalysen der UE unter Ausschluss der Ereignisse des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“. Da ein Ausschluss der UE, die mit dem primären Endpunkt assoziiert sind, auf Ebene der bevorzugten Bezeichnungen nicht möglich ist, werden für die Sensitivitätsanalyse sämtliche UE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen.

Werden die mutmaßlichen Ereignisse des primären Endpunkts aus der Analyse der unerwünschten Ereignisse ausgeschlossen, so ergibt sich in der Betrachtung der Raten über den gesamten Behandlungszeitraum zunächst ein ähnliches Bild wie ohne Ausschluss dieser Ereignisse (vgl. Tabelle 4-90). Bei 97,2 % der Patienten im Selexipag-Arm trat im Laufe der gesamten Behandlungsphase ein UE auf, während im Placebo-Arm nur bei 93,2 % der Patienten ein UE beobachtet wurde. Im Gegensatz zur Analyse der UE unter Einschluss der Ereignisse des primären Endpunkts ist der Unterschied hier jedoch statistisch signifikant (Barnards exakter Text: p=0,0016) und entspricht einer Erhöhung des relativen Risikos von 4 % (RR [95 % KI]: 1,04 [1,02; 1,07]) und einer Erhöhung des absoluten Risikos von 4,0 Prozentpunkten (RD [95 % KI]: 4,0 %-P. [1,5 %-P.; 6,4 %-P.].

Wie aber auch bereits in der Analyse der UE unter Einschluss der Ereignisse des primären Endpunkts geht der Unterschied zwischen den Behandlungsgruppen insbesondere auf die unterschiedlichen Raten in der Titrationsphase zurück, in der das Studiendesign von Selexipag eine Titration bis zu der Konzentration vorsah, bei der nicht mehr tolerierbare unerwünschte Ereignisse auftraten. Die beobachtete Erhöhung des relativen (RR [95 % KI]: 1,23 [1,16; 1,29]) wie auch des absoluten Risikos (RD [95 % KI]: 17,1 %-P. [12,8 %-P.; 21,5 %-P.] in dieser Phase war unter diesen Bedingungen somit zu erwarten.

Ebenso wie die Gesamtrate der UE war auch die Rate der UE schwerer Intensität in der Analyse unter Ausschluss der mutmaßlichen Ereignisse des primären Endpunkts in der Selexipag-Gruppe erhöht im Vergleich zur Placebo-Gruppe (RR [95 % KI]: 1,18 [1,02; 1,37], RD [95 % KI]: 6,4 %-P. [0,8 %-P.; 12,0 %-P.]. Auch hier ging der Unterschied zwischen den Behandlungsgruppen überwiegend auf die Unterschiede in der Titrationsphase zurück, während die Raten in der Erhaltungsphase vergleichbar waren.

Tabelle 4-90: Sensitivitätsanalyse für den Endpunkt „Unerwünschte Ereignisse“ unter Ausschluss der Ereignisse des primären Endpunkts in GRIPHON

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag		Placebo	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n/N	%	n/N	%
Patienten mit mindestens einem unerwünschten Ereignis	559/575	97,2	538/577	93,2
In der Titrationsphase[1]	474/509	93,1	386/508	76,0
In der Erhaltungsphase[1]	470/509	92,3	453/508	89,2
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität	241/575	41,9	205/577	35,5
In der Titrationsphase[1]	110/509	21,6	47/508	9,3
In der Erhaltungsphase[1]	182/509	35,8	156/508	30,7
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem unerwünschten Ereignis	1,04 [1,02; 1,07]	2,53 [1,40; 4,59]	4,0 [1,5; 6,4]	0,0016
In der Titrationsphase[1]	1,23 [1,16; 1,29]	4,28 [2,87; 6,38]	17,1 [12,8; 21,5]	<0,0001
In der Erhaltungsphase[1]	1,04 [1,00; 1,08]	1,46 [0,95; 2,25]	3,2 [-0,4; 6,7]	0,0834
Patienten mit mindestens einem unerwünschten Ereignis schwerer Intensität	1,18 [1,02; 1,37]	1,31 [1,03; 1,66]	6,4 [0,8; 12,0]	0,0271
In der Titrationsphase[1]	2,34 [1,70; 3,21]	2,70 [1,87; 3,90]	12,4 [8,0; 16,7]	<0,0001
In der Erhaltungsphase[1]	1,16 [0,98; 1,39]	1,26 [0,97; 1,63]	5,0 [-0,7; 10,8]	0,0968

[1]) Für die Analyse der UE während Titrations- und Erhaltungsphase wurden diejenigen Patienten ausgeschlossen, die die Therapie während der Titrationsphase abgebrochen hatten.

Da es nicht möglich ist, UE, die mit dem primären Endpunkt „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ assoziiert sind, auf Ebene der PT auszuschließen, werden in der Sensitivitätsanalyse sämtliche UE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Aufteilung nach bevorzugter Bezeichnung oder System-Organ-Klasse

Tabelle 4-91 zeigt eine Auswertung der in der gesamten Behandlungsphase aufgetretenen UE nach bevorzugter Bezeichnung (*preferred term*, PT), der sie zuzuorden sind. Aus dieser Auswertung wird klar, dass insbesondere solche UE vermehrt bei Patienten unter Selexipag aufgetreten waren, die üblicherweise mit dem Einsatz von Arzneimitteln assoziiert werden, die den Prostacyclin-Signalweg addressieren. Zu dieser Gruppe unerwünschter Ereignisse, die in Abschnitt 4.3.1.3.1.17 detailliert diskutiert werden, gehören u. a. die PT „Kopfschmerz“ (Selexipag: 65,2 % vs. Placebo: 32,8 %), „Diarrhöe“ (42,4 % vs. 19,1 %), „Übelkeit“ (33,6 % vs. 18,5 %), „Kieferschmerz“ (25,7 % vs. 6,2 %), „Erbrechen“ (18,1 % vs. 8,5 %), „Extremitätenschmerz“ (16,9 % vs. 8,0 %), „Myalgie“ (16,0 % vs. 5,9 %), „Flush“ (12,2 % vs. 5,0 %), „Abdominalschmerz“ (8,3 % vs. 5,7 %) und „Arthralgien (10,8 % vs. 7,6 %).

Im Gegensatz dazu handelte es sich bei UE, die vermehrt bei Patienten im Placebo-Arm aufgetreten waren, überwiegend um UE, die entweder der PAH direkt (21,9 % vs. 35,7 %) oder den vom G-BA bestimmten Symptomen der PAH zuzurechnen waren. Im Einzelnen ergaben sich für die Symptome der PAH, die bereits in Abschnitt 4.3.1.3.1.7 zusammengefasst analysiert wurden, folgende Unterschiede: „Dyspnoe“ (16,0 % vs. 21,0 %), „Synkope“ (6,4 % vs. 8,8 %), „Erschöpfung“ (8,0 % vs. 10,2 %), „periphere Ödeme“ (13,9 % vs. 18,0 %) und „Schwindel“ (15,0 % vs. 14,7 %). Somit war das Risiko des Auftretens bei allen vom G-BA definierten Symptomen der PAH mit Ausnahme des Schwindels im Selexipag-Arm niedriger als im Placebo-Arm.

Tabelle 4-91: Auflistung der unerwünschten Ereignisse nach bevorzugter Bezeichnung (MedDRA Version 16), die bei mindestens 5 % der Patienten in mindestens einem Studienarm aufgetreten waren (gesamte Behandlungsphase)

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem UE	565	98,3 %	559	96,9 %
Kopfschmerz	375	65,2	189	32,8
Diarrhö	244	42,4	110	19,1
Übelkeit	193	33,6	107	18,5
Kieferschmerz	148	25,7	36	6,2
Pulmonal arterielle Hypertonie	126	21,9	206	35,7
Erbrechen	104	18,1	49	8,5
Extremitätenschmerz	97	16,9	46	8,0
Dyspnoe	92	16,0	121	21,0
Myalgie	92	16,0	34	5,9
Schwindel	86	15,0	85	14,7
Peripheres Ödem	80	13,9	104	18,0
Infektion der oberen Atemwege	75	13,0	80	13,9
Nasopharyngitis	75	13,0	63	10,9
Flush	70	12,2	29	5,0
Arthralgien	62	10,8	44	7,6
Husten	56	9,7	67	11,6
Abdominalschmerz	48	8,3	33	5,7
Anämie	48	8,3	31	5,4
Bronchitis	47	8,2	43	7,5
Erschöpfung	46	8,0	59	10,2
Rechtsherzversagen	46	8,0	58	10,1
Brustschmerz	39	6,8	42	7,3
Synkope	37	6,4	51	8,8
Rückenschmerz	35	6,1	35	6,1
Schmerz des oberen Abdomens	34	5,9	32	5,5
Palpitationen	34	5,9	32	5,5
Verminderter Appetit	34	5,9	19	3,3
Asthenie	31	5,4	24	4,2
Pneumonie	30	5,2	33	5,7
Epistaxis	30	5,2	29	5,0

Hypotonie	29	5,0	18	3,1
Infektion der Harnwege	26	4,5	30	5,2

Bevorzugte Bezeichnungen, bei denen der Unterschied zwischen den Behandlungsgruppen ≥2 Prozentpunkte war, sind grau hinterlegt (hellgrau – Inzidenz in Selexipag-Gruppe höher, mittelgrau – Inzidenz in Placebo-Gruppe höher). Grundlage der Analyse war die SAF-Population.

UE – Unerwünschtes Ereignis, SAF – *safety* (population)

Das Ergebnis der Auswertung der UE nach System-Organ-Klasse (*system organ class*, SOC) in Tabelle 4-92 spiegelt letztlich das Ergebnis der Aufteilung nach PT wider. Die Unterschiede in SOC, bei denen ein häufigeres Auftreten entsprechender UE im Selexipag-Arm beobachtet wurde, wie z. B. Erkrankungen des Nervensystems (73,7 % vs. 49,0 %) und Erkrankungen des Gastrointestinaltrakts (65,4 % vs. 48,9 %), waren hierbei insbesondere durch Unterschiede prostacyclinspezifischer UE geprägt.

Umgekehrt war die SOC „Erkrankungen der Atemwege, des Brustraums und Mediastinums“, in der mehr UE im Placebo-Arm als im Selexipag-Arm beobachtet wurden, überwiegend durch die Unterschiede der PT „pulmonal arterielle Hypertonie“ und „Dyspnoe“ geprägt.

Tabelle 4-92: Auflistung der unerwünschten Ereignisse nach System-Organ-Klasse (MedDRA Version 16)

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem UE	565	98,3 %	559	96,9 %
Erkrankungen des Nervensystems	424	73,7	283	49,0
Erkrankungen des Gastrointestinaltrakts	376	65,4	282	48,9
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	332	57,7	213	36,9
Erkrankungen des Immunsystems	294	51,1	294	51,0
Erkrankungen der Atemwege, des Brustraums und Mediastinums	266	46,3	345	59,8
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	226	39,3	237	41,1
Herzerkrankungen	142	24,7	160	27,7
Gefäßerkrankungen	138	24,0	88	15,3
Untersuchungen	115	20,0	130	22,5
Erkrankungen der Haut und des Unterhautzellgewebes	113	19,7	91	15,8
Stoffwechsel- und Ernährungsstörungen	100	17,4	97	16,8
Erkrankungen des Blutes und des Lymphsystems	70	12,2	63	10,9
Augenerkrankungen	63	11,0	46	8,0
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	55	9,6	73	12,7
Psychiatrische Erkrankungen	48	8,3	61	10,6
Erkrankungen der Nieren und Harnwege	46	8,0	29	5,0
Affektionen der Leber und der Gallenblase	29	5,0	20	3,5
Erkrankungen des Ohrs und des Labyrinths	28	4,9	19	3,3
Erkrankungen der Geschlechtsorgane und der Brustdrüse	26	4,5	27	4,7
Chirurgische und medizinische Eingriffe	21	3,7	14	2,4
Gutartige, bösartige und unspezifische Neubildungen (einschl. Zysten und Polypen)	19	3,3	17	2,9
Endokrine Erkrankungen	18	3,1	12	2,1
Infektionen und parasitäre Erkrankungen	11	1,9	22	3,8

Kongenitale, familiäre und genetische Erkrankungen	1	0,2	4	0,7
Schwangerschaft, Wochenbett und perinatale Erkrankungen	1	0,2	2	0,3
Bevorzugte Bezeichnungen, bei denen der Unterschied zwischen den Behandlungsgruppen ≥5 Prozentpunkte war, sind grau hinterlegt (hellgrau – Inzidenz in Selexipag-Gruppe höher, mittelgrau – Inzidenz in Placebo-Gruppe höher). Grundlage der Analyse war die SAF-Population. UE – Unerwünschtes Ereignis, SAF – *safety* (population)				

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-32 und Abbildung 4-33). Zudem ist der Unterschied zwischen den Behandlungsgruppen zumindest für die Auswertung der schweren UE in der für Deutschland repräsentativen Region Westeuropa weniger stark ausgeprägt als in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation tendenziell von einer Überschätzung des Schadenspotentials für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.14 Schwerwiegende unerwünschte Ereignisse

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-93: Operationalisierung von „Schwerwiegende unerwünschte Ereignisse“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Anteil der Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis im Behandlungszeitraum. Schwerwiegende unerwünschte Ereignisse (SUE) waren entsprechend den Leitlinien der International Conference on Harmonisation (ICH) als jegliches UE, das zumindest eines der folgenden Kriterien erfüllt: - Es ist tödlich. - Es ist lebensbedrohlich. - Es führt zur Hospitalisierung des Patienten oder zur Verlängerung eines bereits bestehenden stationären Aufenthalts. - Es führt zu anhaltender oder bedeutender Einschränkung oder Behinderung. - Es handelt sich um eine angeborene Anomalie oder einen Geburtsfehler. - Es ist medizinisch bedeutsam oder setzt eine Intervention voraus, um zumindest eines der oben gelisteten Ereignisse zu verhindern. Lebensbedrohlich bezieht sich in diesem Zusammenhang auf ein Ereignis, aufgrund dessen der Patient zum Zeitpunkt des Ereignisses zu versterben drohte, jedoch nicht auf ein Ereignis, das hypothetischerweise zum Tod hätte führen können, wenn es schwerer gewesen wäre. Wichtige medizinische Ereignisse, die nicht unmittelbar zum Tod führen, lebensbedrohlich sind oder zur Hospitalisierung führen, können ebenfalls als SUE angesehen werden, sofern sie ausgehend von einem angemessenen medizinischen Urteil den Patienten in Gefahr bringen und medizinische oder chirurgische Intervention erforderlich machen könnten, um eines der oben aufgeführten Ereignisse abzuwenden. Die SUE werden analog zur Darstellung der UE nach System-Organ-Klasse (*system organ class*, SOC) und bevorzugter Bezeichnung (*preferred term*, PT) der MedDRA-Klassifikation, Version 16, zusammengefasst. Zur Operationalisierung der UE siehe auch Tabelle 4-86.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-94: Bewertung des Verzerrungspotentials für „Schwerwiegende unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es liegen jedoch sonstige das Verzerrungspotential beeinflussende Aspekte vor. So setzt die Auswertung der Ereignisse des Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen gemäß dem Cochrane Handbook for Systematic Reviews of Interventions jedoch voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden [83]. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer bzw. nicht mehr handhabbarer prostacyclinspezifischer Nebenwirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben.

Insgesamt ist das Verzerrungspotential für diesen Endpunkt als hoch anzusehen. Aufgrund der vergleichsweise längeren medianen Behandlungsdauer in der Selexipag-Gruppe (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchstens von einer Verzerrung der Ergebnisse zuungunsten von Selexipag auszugehen.

Stellen Sie die Ergebnisse für den Endpunkt „Schwerwiegende unerwünschte Ereignisse" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-95: Ergebnisse für „Schwerwiegende unerwünschte Ereignisse" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: SAF**			
Intervention	Selexipag N=575		Placebo N=577	
Anzahl Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis				
	n	%	n	%
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis	252	43,8	272	47,1
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis	0,93 [0,82; 1,06]	0,88 [0,69; 1,10]	-3,3 [-9,1; 2,4]	0,2902
Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt. KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*				

Hauptanalyse

Die Analyse der unerwünschten Ereignisse (UE) wurde auf Grundlage der SAF-Population durchgeführt.

Insgesamt trat im Verlauf der Behandlungsphase von GRIPHON bei 43,8 % der Patienten im Selexipag-Arm und bei 47,1 % der Patienten im Placebo-Arm mindestens ein SUE auf. Dies entspricht einer numerischen Verringerung des relativen Risikos von 7 % (RR [95 % KI]: 0,93 [0,82; 1,06]) bzw. einer numerischen Verringerung des absoluten Risikos von 3,3 Prozentpunkten (RD [95 % KI]: −3,3 %-P. [−9,1 %-P.; 2,4 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war jedoch statistisch nicht signifikant (Barnards exakter Test, p=0,2902).

Die folgende Tabelle 4-96 zeigt eine Aufteilung der schwerwiegenden unerwünschten Ereignisse in Abhängigkeit von der Behandlungsphase, in der das SUE aufgetreten ist. Nicht eingeschlossen in diese Auswertung wurden Patienten, die bereits in der Titrationsphase die Behandlung beendet hatten.

Im Gegensatz zur Auswertung der UE, bei der in der Titrationsphase eine erhöhte Ereignisrate im Selexipag-Arm beobachtet wurde, sind die Anteile der Patienten mit mindestens einem SUE in der gleichen Phase ausgeglichen (Selexipag: 7,1 %, Placebo: 7,9 %). In der Erhaltungsphase wiederum ist eine leichte numerische Tendenz zu einer geringeren Ereignisrate in der Selexipag-Gruppe zu beobachten (Selexipag: 42,4 %, Placebo: 44,5 %).

Tabelle 4-96: Aufteilung der schwerwiegenden unerwünschten Ereignisse nach Titrations- und Erhaltungsphase

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag N=509		Placebo N=508	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n	%	n	%
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis in der **Titrationsphase**	36	7,1	40	7,9
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis in der **Erhaltungssphase**	216	42,4	226	44,5
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis in der **Titrationsphase**	0,90 [0,58; 1,39]	0,89 [0,56; 1,42]	-0,8 [-4,0; 2,4]	0,6830
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis in der **Erhaltungsphase**	0,95 [0,83; 1,10]	0,92 [0,72; 1,18]	-2,1 [-8,1; 4,0]	0,5329

[1]) Ausschließlich der Patienten, die die Therapie während der Titrationsphase abgebrochen hatten.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Nebenanalysen

Sensitivitätsanalyse unter Ausschluss der Ereignisse des primären Endpunkts

Die folgende Tabelle 4-97 zeigt die vom G-BA im Beratungsgespräch zu Selexipag [3] geforderten Sensitivitätsanalysen der SUE unter Ausschluss der Ereignisse des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“. Da ein Ausschluss der SUE, die mit dem primären Endpunkt assoziiert sind, auf Ebene der

bevorzugten Bezeichnungen nicht möglich ist, werden für die Sensitivitätsanalyse sämtliche SUE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen.

Die Sensitivitätsanalyse unter Ausschluss der mutmaßlichen Ereignisse des primären Endpunkts zeigt ein zur Hauptanalyse konsistentes Ergebnis. Es ergeben sich weder in der gesamten Behandlungsphase noch in der Titrations- oder der Erhaltungsphase bedeutsame Unterschiede zwischen den Behandlungsgruppen.

In der gesamten Behandlungsphase trat nach Ausschluss der Ereignisse des primären Endpunkts bei 35,1 % der Patienten im Selexipag-Arm und bei 37,3 % im Placebo-Arm mindestens ein SUE auf. Dies entspricht einer numerischen Reduktion des relativen Risikos von 6 % (RR [95 % KI]: 0,94 [0,81; 1,10]) und einer numerischen Reduktion des absoluten Risikos um 2,1 Prozentpunkte (RD [95 % KI]: −2,1 %-P. [−7,7 %-P.; 3,4 %-P.]. Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,5328).

Sowohl in der Titrations- als auch in der Erhaltungsphase waren die beobachteten Ereignisraten nahezu ausgeglichen, ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen besteht nicht (Barnards exakter Test: p=0,9436 (Titrationsphase), p=0,9557 (Erhaltungsphase)).

Tabelle 4-97: Sensitivitätsanalyse für den Endpunkt „Schwerwiegende unerwünschte Ereignisse“ unter Ausschluss der Ereignisse des primären Endpunkts in GRIPHON

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag		Placebo	
Anzahl Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis				
	n/N	%	n/N	%
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis	202/575	35,1	215/577	37,3
In der Titrationsphase[1]	35/509	6,9	34/508	6,7
In der Erhaltungsphase[1]	176/509	34,6	177/508	34,8
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem schwerwiegenden unerwünschten Ereignis	0,94 [0,81; 1,10]	0,91 [0,72; 1,16]	-2,1 [-7,7; 3,4]	0,5328
In der Titrationsphase[1]	1,03 [0,65; 1,62]	1,03 [0,63; 1,68]	0,2 [-2,9; 3,3]	0,9436
In der Erhaltungsphase[1]	0,99 [0,84; 1,18]	0,99 [0,76; 1,28]	-0,3 [-6,1; 5,6]	0,9557

[1]) Für die Analyse der UE während Titrations- und Erhaltungsphase wurden diejenigen Patienten ausgeschlossen, die die Therapie während der Titrationsphase abgebrochen hatten.

Da es nicht möglich ist, UE, die mit dem primären Endpunkt „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ assoziiert sind, auf Ebene der PT auszuschließen, werden in der Sensitivitätsanalyse sämtliche UE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Aufteilung nach bevorzugter Bezeichnung oder System-Organ-Klasse

Tabelle 4-98 zeigt eine Auswertung der in der gesamten Behandlungsphase aufgetretenen SUE nach PT, der sie zuzuordnen sind. Der Großteil der aufgetretenen SUE ist primär der zugrunde liegenden Erkrankung und damit voraussichtlich nicht der Studienmedikation zuzurechnen. Die am häufigsten nach PT aufgetretenen SUE in GRIPHON waren die Verschlechterung der pulmonal arteriellen Hypertonie (Selexipag: 14,4 % vs. Placebo: 22,0 %) und das Rechtsherzversagen (5,9 % vs. 7,1 %). Ebenfalls häufig aufgetreten waren Pneumonie (3,0 % vs. 4,3 %), Dyspnoe (3,0 % vs. 2,3 %) und Synkope (1,7 % vs. 3,5 %).

Tabelle 4-98: Auflistung der schwerwiegenden unerwünschten Ereignisse nach bevorzugter Bezeichnung (MedDRA Version 16), die bei mindestens 5 Patienten in mindestens einem Studienarm aufgetreten waren (gesamte Behandlungsphase)

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem SUE	252	43,8	272	47,1
Pulmonal arterielle Hypertonie	83	14,4	127	22,0
Rechtsherzversagen	34	5,9	41	7,1
Pneumonie	17	3,0	25	4,3
Dyspnoe	17	3,0	13	2,3
Synkope	10	1,7	20	3,5
Vorhofflimmern	7	1,2	4	0,7
Brustschmerz	6	1,0	6	1,0
Akutes Nierenversagen	6	1,0	6	1,0
Bronchitis	6	1,0	4	0,7
Hämoptysis	5	0,9	5	0,9
Abdominalschmerz	5	0,9	4	0,7
Plötzlicher Tod	5	0,9	4	0,7
Anämie	5	0,9	3	0,5
Systemischer Lupus erythematodes	5	0,9	1	0,2
Sturz	3	0,5	6	1,0
Vorhofflattern	3	0,5	5	0,9
Akutes Rechtsherzversagen	2	0,3	5	0,9

Bevorzugte Bezeichnungen, bei denen der Unterschied zwischen den Behandlungsgruppen ≥1 Prozentpunkte war, sind grau hinterlegt (hellgrau – Inzidenz in Selexipag-Gruppe höher, mittelgrau – Inzidenz in Placebo-Gruppe höher). Grundlage der Analyse war die SAF-Population.

SUE – schwerwiegendes unerwünschtes Ereignis, SAF – *safety* (population)

Die Häufigkeit schwerwiegender unerwünschter Ereignisse aufgeteilt nach SOC zeigt Tabelle 4-99. Am häufigsten traten in der Behandlungsphase von GRIPHON SUE in der SOC „Erkrankungen der Atemwege, des Brustraums und Mediastinums“ (19,3 % vs. 26,0 %) auf. Ebenfalls häufig waren „Herzerkrankungen“ (11,8 % vs. 13,7 %) und „Infektionen und parasitäre Erkrankungen“ (10,8 % vs. 10,1 %).

Tabelle 4-99: Auflistung der schwerwiegenden unerwünschten Ereignisse nach System-Organ-Klasse (MedDRA Version 16)

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem SUE	252	43,8	272	47,1
Erkrankungen der Atemwege, des Brustraums und Mediastinums	111	19,3	150	26,0
Herzerkrankungen	68	11,8	79	13,7
Infektionen und parasitäre Erkrankungen	62	10,8	58	10,1
Erkrankungen des Nervensystems	26	4,5	28	4,9
Erkrankungen des Gastrointestinaltrakts	24	4,2	22	3,8
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	19	3,3	23	4,0
Gefäßerkrankungen	16	2,8	11	1,9
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	15	2,6	7	1,2
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	12	2,1	15	2,6
Untersuchungen	11	1,9	10	1,7
Chirurgische und medizinische Eingriffe	10	1,7	7	1,2
Erkrankungen der Nieren und Harnwege	10	1,7	7	1,2
Erkrankungen des Blutes und des Lymphsystems	9	1,6	4	0,7
Gutartige, bösartige und unspezifische Neubildungen (einschl. Zysten und Polypen)	8	1,4	5	0,9
Stoffwechsel- und Ernährungsstörungen	8	1,4	8	1,4
Affektionen der Leber und der Gallenblase	5	0,9	4	0,7
Erkrankungen der Geschlechtsorgane und der Brustdrüse	4	0,7	3	0,5
Augenerkrankungen	3	0,5	-	
Endokrine Erkrankungen	3	0,5	-	
Erkrankungen der Haut und des Unterhautzellgewebes	3	0,5	5	0,9
Erkrankungen des Immunsystems	2	0,3	2	0,3
Erkrankungen des Ohrs und des Labyrinths	1	0,2	-	
Psychiatrische Erkrankungen	1	0,2	4	0,7

Schwangerschaft, Wochenbett und perinatale Erkrankungen	1	0,2	2	0,3
Bevorzugte Bezeichnungen, bei denen der Unterschied zwischen den Behandlungsgruppen ≥2 Prozentpunkte war, sind grau hinterlegt (hellgrau – Inzidenz in Selexipag-Gruppe höher, mittelgrau – Inzidenz in Placebo-Gruppe höher). Grundlage der Analyse war die SAF-Population. SUE – Unerwünschtes Ereignis, SAF – *safety* (population)				

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-34). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa ähnlich ausgeprägt wie in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation nicht von einem unterschiedlichen Therapieffekt für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.15 Therapieabbrüche wegen unerwünschter Ereignisse

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-100: Operationalisierung von „Therapieabbrüche wegen unerwünschter Ereignisse“

Studie	Operationalisierung
GRIPHON	Definition des Endpunkts Anteil der Patienten, die die Behandlung aufgrund eines unerwünschten Ereignisses vorzeitig abbrachen. Die Therapieabbrüche wegen unerwünschter Ereignisse werden analog zur Darstellung der UE nach System-Organ-Klasse (*system organ class*, SOC) und bevorzugter Bezeichnung (*preferred term*, PT) der MedDRA-Klassifikation, Version 16, zusammengefasst. Zur Operationalisierung der UE siehe auch Tabelle 4-86.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-101: Bewertung des Verzerrungspotentials für „Therapieabbrüche wegen unerwünschter Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotential auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotential Endpunkt
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es liegen jedoch sonstige das Verzerrungspotential beeinflussende Aspekte vor. So setzt die Auswertung der Ereignisse des Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen gemäß dem Cochrane Handbook for Systematic Reviews of Interventions jedoch voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden [83]. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende

der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer bzw. nicht mehr handhabbarer prostacyclinspezifischer Nebenwirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses und damit für einen Therapieabbruch aufgrund eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben.

Insgesamt ist das Verzerrungspotential für diesen Endpunkt als hoch anzusehen. Aufgrund der vergleichsweise längeren medianen Behandlungsdauer in der Selexipag-Gruppe (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchstens von einer Verzerrung der Ergebnisse zuungunsten von Selexipag auszugehen.

Stellen Sie die Ergebnisse für den Endpunkt „Therapieabbrüche wegen unerwünschter Ereignisse“ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-102: Ergebnisse für „Therapieabbrüche wegen unerwünschter Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: SAF**			
Intervention	Selexipag N=575		Placebo N=577	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n	%	n	%
Therapieabbrecher wegen eines unerwünschten Ereignisses	182	31,7	214	37,1
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Therapieabbrecher wegen eines unerwünschten Ereignisses	0,85 [0,73; 1,00]	0,79 [0,62; 1,00]	-5,4 [-10,9; 0,0]	0,0554

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Hauptanalyse

Die Analyse der Therapieabbrüche wegen unerwünschter Ereignisse wurde auf Grundlage der SAF-Population durchgeführt.

Insgesamt brachen im Verlauf der Behandlungsphase von GRIPHON 31,7 % der Patienten im Selexipag-Arm und 37,1 % der Patienten im Placebo-Arm die Therapie mit der Studienmedikation aufgrund eines UE vorzeitig ab. Dies entspricht einer numerischen Verringerung des relativen Risikos von 15 % (RR [95 % KI]: 0,85 [0,73; 1,00]) bzw. einer numerischen Verringerung des absoluten Risikos von 5,4 Prozentpunkten (RD [95 % KI]: −5,4 %-P. [−10,9 %-P.; 0,0 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,0554), jedoch stellt die Obergrenze des Konfidenzintervalls beider Effektschätzer gleichzeitig auch den Nulleffekt-Wert dar.

Die folgende Tabelle 4-103 zeigt eine Aufteilung der Therapieabbrüche aufgrund unerwünschter Ereignisse in Abhängigkeit von der Behandlungsphase, in der das verursachende UE aufgetreten ist. Nicht eingeschlossen in diese Auswertung wurden Patienten, die bereits in der Titrationsphase die Behandlung beendet hatten.

In dieser Analyse zeigt sich, dass in der Titrationsphase die Anzahl der Therapieabbrecher im Selexipag-Arm numerisch höher war als im Placebo-Arm (Selexipag: 3,9 %, Placebo: 2,4 %). Der Unterschied zwischen den Behandlungsgruppen war jedoch statistisch nicht signifikant (RR [95 % KI]: 1,66 [0,82; 3,37], RD [95 % KI]: 1,6 %-P. [−0,6 %-P.; 3,7 %-P.], Barnards exakter Test: p=0,1555).

Im Gegensatz zur Titrationsphase war die Rate der Therapieabbrecher aufgrund eines UE in der Erhaltungsphase unter Selexipag numerisch niedriger als unter Placebo (Selexipag: 26,7 %, Placebo: 30,9 %). Aber auch hier war der Unterschied zwischen den Behandlungsgruppen statistisch nicht signifikant (RR [95 % KI]: 0,87 [0,71; 1,05], RD [95 % KI]: −4,2 %-P. [−9,7 %-P.; 1,4 %-P.], Barnards exakter Test: p=0,1467).

Tabelle 4-103: Aufteilung der Therapieabbrüche wegen eines unerwünschten Ereignisses nach Titrations- und Erhaltungsphase

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag N=509		Placebo N=508	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n	%	n	%
Therapieabbrecher wegen eines unerwünschten Ereignisses in der **Titrationsphase**	20	3,9	12	2,4
Therapieabbrecher wegen eines unerwünschten Ereignisses in der **Erhaltungsphase**	136	26,7	157	30,9
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Therapieabbrecher wegen eines unerwünschten Ereignisses in der **Titrationsphase**	1,66 [0,82; 3,37]	1,69 [0,82; 3,50]	1,6 [-0,6; 3,7]	0,1555
Therapieabbrecher wegen eines unerwünschten Ereignisses in der **Erhaltungsphase**	0,87 [0,71; 1,05]	0,82 [0,62; 1,07]	-4,2 [-9,7; 1,4]	0,1467

[1]) Ausschließlich der Patienten, die die Therapie während der Titrationsphase abgebrochen hatten.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Nebenanalysen

Sensitivitätsanalyse unter Ausschluss der Ereignisse des primären Endpunkts

Die folgende Tabelle 4-104 zeigt die vom G-BA im Beratungsgespräch zu Selexipag [3] geforderten Sensitivitätsanalysen der Therapieabbrüche aufgrund eines UE unter Ausschluss der Ereignisse des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“. Da ein Ausschluss der Therapieabbrüche aufgrund von UE, die mit dem primären Endpunkt assoziiert sind, auf Ebene der bevorzugten Bezeichnungen nicht möglich ist, werden für die Sensitivitätsanalyse sämtliche Therapieabbrüche wegen UE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen.

Die Sensitivitätsanalyse unter Ausschluss der mutmaßlichen Ereignisse des primären Endpunkts zeigt ein zur Hauptanalyse konsistentes Ergebnis, wenngleich sich der (numerische) Unterschied zwischen den Behandlungsgruppen nivelliert. Es ergeben sich weder in der gesamten Behandlungsphase noch in der Titrations- oder der Erhaltungsphase bedeutsame Unterschiede zwischen den Behandlungsgruppen.

In der gesamten Behandlungsphase brachen nach Ausschluss der Ereignisse des primären Endpunkts 22,1 % der Patienten im Selexipag-Arm und 22,7 % im Placebo-Arm die Therapie wegen eines UE ab. Dies entspricht einer numerischen Reduktion des relativen Risikos von 3 % (RR [95 % KI]: 0,97 [0,79; 1,21]) und einer numerischen Reduktion des absoluten Risikos um 0,6 Prozentpunkte (RD [95 % KI]: −0,6 %-P. [−5,4 %-P.; 4,2 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,8694).

Wie auch in der Analyse unter Einschluss der Ereignisse des primären Endpunkts war auch nach deren Ausschluss in der Titrationsphase die Rate der Therapieabbrecher in der Selexipag-Gruppe höher als in der Placebo-Gruppe. Es ergab sich eine numerische Erhöhung des relativen Risikos um 111 % (RR [95 % KI]: 2,11 [0,96; 4,61]) bzw. eine numerische Erhöhung des absoluten Risikos von 2,0 Prozentpunkten (RD [95 % KI]: 2,0 %-P. [0,0 %-P.; 4,0 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,0599).

In der Erhaltungsphase waren die beobachteten Ereignisraten in beiden Behandlungsgruppen nahezu ausgeglichen, ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen wurde nicht beobachtet (Barnards exakter Test: p=1,000).

Tabelle 4-104: Sensitivitätsanalyse für den Endpunkt „Therapieabbrüche wegen unerwünschter Ereignisse“ unter Ausschluss der Ereignisse des primären Endpunkts in GRIPHON

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag		Placebo	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n/N	%	n/N	%
Therapieabbrecher wegen eines unerwünschten Ereignisses	127/575	22,1	131/577	22,7
In der Titrationsphase[1]	19/509	3,7	9/508	1,8
In der Erhaltungsphase[1]	90/509	17,7	90/508	17,7
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Therapieabbrecher wegen eines unerwünschten Ereignisses	0,97 [0,79; 1,21]	0,97 [0,73; 1,27]	-0,6 [-5,4; 4,2]	0,8694
In der Titrationsphase[1]	2,11 [0,96; 4,61]	2,15 [0,96; 4,80]	2,0 [0,0; 4,0]	0,0599
In der Erhaltungsphase[1]	1,00 [0,77; 1,30]	1,00 [0,72; 1,38]	0,0 [-4,7; 4,7]	1,0000

[1]) Für die Analyse der UE während Titrations- und Erhaltungsphase wurden diejenigen Patienten ausgeschlossen, die die Therapie während der Titrationsphase abgebrochen hatten.

Da es nicht möglich ist, UE, die mit dem primären Endpunkt „Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis“ assoziiert sind, auf Ebene der PT auszuschließen, werden in der Sensitivitätsanalyse sämtliche UE ausgeschlossen, deren Start-Datum mit dem Start-Datum des jeweiligen M/M-Ereignisses zusammenfällt oder die später begannen.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand von Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Aufteilung nach bevorzugter Bezeichnung oder System-Organ-Klasse

Tabelle 4-105 zeigt eine Auswertung der Therapieabbrüche wegen unerwünschter Ereignisse, aufgeteilt nach PT des zugrunde liegenden UE. Am häufigsten wurde die Behandlung mit der Studienmedikation in GRIPHON aufgrund einer Verschlechterung der PAH abgebrochen (Selexipag: 13,6 % vs. 23,4 %). Ebenfalls häufig waren Therapieabbrüche wegen Kopfschmerz (3,3 % vs. 0,7 %), Rechtsherzversagen (2,4 % vs. 4,0 %), Diarrhö (2,3 % vs. 0 %) und Übelkeit (1,7 % vs. 0,5 %).

Tabelle 4-105: Auflistung der Therapieabbrüche wegen unerwünschter Ereignisse nach bevorzugter Bezeichnung des zugrunde liegenden unerwünschten Ereignisses (MedDRA Version 16), die bei mindestens 5 Patienten in mindestens einem Studienarm aufgetreten waren (gesamte Behandlungsphase)

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Therapieabbrecher aufgrund eines UE	182	31,7	214	37,1
Pulmonal arterielle Hypertonie	78	13,6	135	23,4
Kopfschmerz	19	3,3	4	0,7
Rechtsherzversagen	14	2,4	23	4,0
Diarrhö	13	2,3	-	
Übelkeit	10	1,7	3	0,5
Dyspnoe	7	1,2	10	1,7
Extremitätenschmerz	6	1,0	2	0,3
Plötzlicher Tod	5	0,9	1	0,2
Myalgie	5	0,9	-	

Bevorzugte Bezeichnungen, bei denen der Unterschied zwischen den Behandlungsgruppen ≥1 % war, sind grau hinterlegt (hellgrau – Inzidenz in Selexipag-Gruppe höher, mittelgrau – Inzidenz in Placebo-Gruppe höher). Grundlage der Analyse war die SAF-Population.
SUE – schwerwiegendes unerwünschtes Ereignis, SAF – *safety* (population)

Die Häufigkeit der Therapieabbrüche wegen unerwünschter Ereignisse aufgeteilt nach SOC zeigt Tabelle 4-106. Am häufigsten wurde die Behandlung in GRIPHON wegen UE der SOC „Erkrankungen der Atemwege, des Brustraums und Mediastinums“ (19,3 % vs. 26,0 %) abgebrochen. UE der SOC „Erkrankungen des Nervensystems“ (5,0 % vs. 1,6 %), „Herzerkrankungen“ (4,3 % vs. 5,9 %) und „Erkrankungen des Gastrointestinaltrakts“ (4,3 % vs. 1,4 %) waren ebenfalls häufig der Grund für den Therapieabbruch.

Tabelle 4-106: Auflistung der Therapieabbrüche wegen unerwünschter Ereignisse nach System-Organ-Klasse (MedDRA Version 16)

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem SUE	182	31,7	214	37,1
Erkrankungen der Atemwege, des Brustraums und Mediastinums	91	15,8	149	25,8
Erkrankungen des Nervensystems	29	5,0	9	1,6
Herzerkrankungen	25	4,3	34	5,9
Erkrankungen des Gastrointestinaltrakts	25	4,3	8	1,4
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	16	2,8	6	1,0
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	14	2,4	9	1,6
Infektionen und parasitäre Erkrankungen	10	1,7	6	1,0
Gefäßerkrankungen	5	0,9	4	0,7
Chirurgische und medizinische Eingriffe	4	0,7	2	0,3
Untersuchungen	3	0,5	9	1,6
Erkrankungen der Nieren und Harnwege	3	0,5	2	0,3
Erkrankungen der Haut und des Unterhautzellgewebes	3	0,5	-	
Psychiatrische Erkrankungen	2	0,3	1	0,2
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	2	0,3	1	0,2
Augenerkrankungen	2	0,3	-	
Affektionen der Leber und der Gallenblase	1	0,2	2	0,3
Stoffwechsel- und Ernährungsstörungen	1	0,2	2	0,3
Gutartige, bösartige und unspezifische Neubildungen (einschl. Zysten und Polypen)	1	0,2	1	0,2
Endokrine Erkrankungen	1	0,2	-	
Erkrankungen des Immunsystems	-		2	0,3
Kongenitale, familiäre und genetische Erkrankungen	-		1	0,2

Bevorzugte Bezeichnungen, bei denen der Unterschied zwischen den Behandlungsgruppen ≥2 Prozentpunkte war, sind grau hinterlegt (hellgrau – Inzidenz in Selexipag-Gruppe höher, mittelgrau – Inzidenz in Placebo-Gruppe höher). Grundlage der Analyse war die SAF-Population.

SUE – Schwerwiegendes unerwünschtes Ereignis, SAF – *safety* (population)

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-35). Zudem ist der Therapieeffekt in der für Deutschland repräsentativen Region Westeuropa ähnlich ausgeprägt wie in der Gesamtstudienpopulation von GRIPHON, sodass bei einer tatsächlichen Effektmodifikation nicht von einem unterschiedlichen Therapieffekt für Deutschland auszugehen wäre.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.16 Unerwünschte Ereignisse von besonderem Interesse

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-107: Operationalisierung von „Unerwünschte Ereignisse von besonderem Interesse“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Anteil der Patienten mit mindestens einem unerwünschten Ereignis von besonderem Interesse im Behandlungszeitraum. Zur Operationalisierung der UE siehe auch Tabelle 4-86. Unerwünschte Ereignisse von besonderem Interesse waren „Hämorrhagie“, „Hypotonie“ und „Störungen des Auges“. Zur Auswertung der unerwünschten Ereignisse von besonderem Interesse wurde folgendermaßen vorgegangen: <u>Hämorrhagie</u> Die Auswertung der Blutungsereignisse erfolgte anhand der *Standardized MedDRA Queries* (SMQ) „Hämorrhagie“ (ausschließlich von Laborwerten) und „Gastrointestinale Hämorrhagie“. Ausreichend war jeweils mindestens ein UE, das als PT in einer der beiden SMQ gelistet war. <u>Hypotonie</u> Die Auswertung der Hypotonie-Ereignisse erfolgte anhand einer vorab definierten Liste von UE mit Hypotonie-assoziierten PT: „Verringerter ambulatorischer Blutdruck“, „Verringerter Blutdruck“, „Verringerter diastolischer Blutdruck“, „Verringerter orthostatischer Blutdruck“, „Verringerter systolischer Blutdruck“, „Diastolische Hypotonie“, „Hypotonie“, „Verringerter arterieller Blutdruck“, „Orthostatische Hypotonie“, „Prozedurale Hypotonie“ Ausreichend war jeweils mindestens ein UE der vorab definierten Liste. <u>Störungen des Auges</u> Die Auswertung der Störungen des Auges erfolgte anhand der System-Organ-Klasse (*system organ class*, SOC) „Augenerkrankungen“ und der SMQ „Retinale Störungen“. Ausreichend war jeweils mindestens ein UE, das als PT in der SOC oder der SMQ gelistet war.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-108: Bewertung des Verzerrungspotentials für „Unerwünschte Ereignisse von besonderem Interesse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es liegen jedoch sonstige das Verzerrungspotential beeinflussende Aspekte vor. So setzt die Auswertung der Ereignisse des Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen gemäß dem Cochrane Handbook for Systematic Reviews of Interventions jedoch voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden [83]. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer bzw. nicht mehr handhabbarer prostacyclinspezifischer Nebenwirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben.

Insgesamt ist das Verzerrungspotential für diesen Endpunkt als hoch anzusehen. Aufgrund der vergleichsweise längeren medianen Behandlungsdauer in der Selexipag-Gruppe (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchstens von einer Verzerrung der Ergebnisse zuungunsten von Selexipag auszugehen.

Stellen Sie die Ergebnisse für den Endpunkt „Unerwünschte Ereignisse von besonderem Interesse" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-109: Ergebnisse für „Unerwünschte Ereignisse von besonderem Interesse" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: SAF**			
Intervention	Selexipag N=575		Placebo N=577	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
Unerwünschtes Ereignis von besonderem Interesse	n	%	n	%
Hämorrhagie	90	15,7	91	15,8
Hypotonie	34	5,9	22	3,8
Störung des Auges	63	11,0	46	8,0
Behandlungseffekt (Selexipag gegenüber Placebo)				
Unerwünschtes Ereignis von besonderem Interesse	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Hämorrhagie	0,99 [0,76; 1,30]	0,99 [0,72; 1,36]	-0,1 [-4,3; 4,1]	0,9756
Hypotonie	1,55 [0,92; 2,62]	1,59 [0,92; 2,75]	2,1 [-0,4; 4,6]	0,1025
Störung des Auges	1,37 [0,96; 1,97]	1,42 [0,95; 2,12]	3,0 [-0,4; 6,4]	0,0968

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen mit Ausnahme der Bestimmung der RR wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Hauptanalyse

Die Analyse der unerwünschten Ereignisse von besonderem Interesse wurde auf Grundlage der SAF-Population durchgeführt.

Hämorrhagie

Die Auswertung von Blutungsereignissen als unerwünschte Ereignisse von besonderem Interesse erfolgte aufgrund der bereits vor GRIPHON bekannten Eigenschaft von Prostacyclin-Rezeptor-Antagonisten, die Thrombozytenaggregation hemmen zu können (vgl. Abschnitt 4.2.5.2).

Der Anteil der Patienten mit mindestens einem Hämorrhagie-assoziiertem UE in der Behandlungsphase betrug 15,7 % im Selexipag-Arm und 15,8 % im Placebo-Arm. Dies entspricht einem nahezu ausgeglichenen relativen (RR [95 % KI]: 0,99 [0,76; 1,30]) und absoluten Risiko (RD [95 % KI]: −0,1 %-P. [−4,3 %-P.; 4,1 %-P.]. Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,9756).

Weiterhin ergab auch eine detaillierte Auswertung der einzelnen Blutungsereignisse in GRIPHON keinen Anhaltspunkt für ein häufigeres Auftreten nach der Anwendung von Selexipag. Für die Ergebnisse dieser Analysen sei jedoch auf den Studienbericht von GRIPHON verwiesen.

Hypotonie

Hypotonie ist ein möglicher Klasseneffekt der Agonisten des IP-Rezeptors und wurde daher ebenfalls als unerwünschtes Ereignis von besonderem Interesse in GRIPHON ausgewertet (vgl. Abschnitt 4.2.5.2).

Der Anteil der Patienten mit mindestens einem Hypotonie-assoziierten UE in der Behandlungsphase lag im Selexipag-Arm bei 5,9 % während im Placebo-Arm bei 3,8 % der Patienten ein solches UE aufgetreten war. Dies entspricht einer numerischen Erhöhung des relativen Risikos um 55 % (RR [95 % KI]: 1,55 [0,92; 2,62] und einer numerischen Erhöhung des absoluten Risikos um 2,1 Prozentpunkt (RD [95 % KI]: 2,1 %-P. [−0,4 %-P.; 4,6 %-P.]). Der Unterschied zwischen den Behandlungsgruppen ist statistisch nicht signifikant (Barnards exakter Test: p=0,1025).

Die Aufteilung der Hypotonie-assoziierten UE nach PT ist in Tabelle 4-110 dargestellt. Demnach ist der Großteil der aufgetretenen Hypotonie-assoziierten UE (und auch der numerische Unterschied zwischen den Behandlungsgruppen) der PT „Hypotonie“ zuzuordnen.

Weiterhin ergab auch die detaillierte Auswertung der einzelnen Hypotonie-Ereignisse, dass höchstens ein geringer Effekt von Selexipag auf den systemischen Blutdruck besteht. Für die Ergebnisse dieser Analysen sei jedoch auf den Studienbericht von GRIPHON verwiesen.

Tabelle 4-110: Hypotonie-assoziierte unerwünschte Ereignisse in GRIPHON, aufgeteilt nach bevorzugter Bezeichnung

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem Hypotonie-assoziierten UE	90	15,7	91	15,8
Hypotonie	29	5,0	18	3,1
Orthostatische Hypotonie	5	0,9	3	0,5
Verringerter Blutdruck	-		1	0,2

Grundlage der Analyse war die SAF-Population.
UE – Unerwünschtes Ereignis, SAF – *safety* (population)

Störung des Auges

Der Grund für die Auswertung von Störungen des Auges als unerwünschte Ereignisse von besonderem Interesse waren die Ergebnisse einer präklinischen Studie, in denen am Ende einer zweijährigen Karzinogenität-Studie Tortuosität und Dilatation retinaler Blutgefäße bei Ratten beobachtet worden waren [72]. Auch wenn die Übertragbarkeit dieser Beobachtung auf den Menschen als unwahrscheinlich angesehen wurde, wurden Störungen des Auges dennoch als besonderer Sicherheitsschwerpunkt in GRIPHON untersucht.

Während der Behandlungsphase von GRIPHON trat bei insgesamt 11,0 % der Patienten im Selexipag-Arm und bei 8,0 % der Patienten im Placebo-Arm mindestens eine Störung des Auges auf. Dies entspricht einer numerischen Erhöhung des relativen Risikos um 37 % (RR [95 % KI]: 1,37 [0,96; 1,97]) und einer numerischen Erhöhung des absoluten Risikos um 3,0 Prozentpunkte (RD [95 % KI]: 3,0 %-P. [−0,4 %-P.; 6,4 %-P.]). Der zwischen den Behandlungsgruppen beobachtete Unterschied war jedoch statistisch nicht signifikant (Barnards exakter Test: p=0,0968).

Auch weitere detaillierte Untersuchungen, die zur Abklärung des Risikos von Störungen des Auges durch den Einsatz von Selexipag durchgeführt wurden, konnten die ursprüngliche Befürchtung nicht bestätigen. Für die Ergebnisse dieser Analysen sei jedoch auf den Studienbericht von GRIPHON verwiesen.

Aufteilung der unerwünschten Ereignisse nach Titrations- und Erhaltungsphase

Die folgende Tabelle 4-103 zeigt eine Aufteilung der Therapieabbrüche aufgrund unerwünschter Ereignisse in Abhängigkeit von der Behandlungsphase, in der das verursachende UE aufgetreten ist. Nicht eingeschlossen in diese Auswertung wurden Patienten, die bereits in der Titrationsphase die Behandlung beendet hatten.

Die Ergebnisse zum Auftreten der UE von besonderem Interesse in Titrations- bzw. Erhaltungsphase sind vollständig konsistent zur Hauptanalyse. Es zeigt sich weder in Bezug auf Hämorrhagie, Hypotonie oder Störungen des Auges ein signifikanter Unterschied zwischen den Behandlungsgruppen in der Titrations- oder Erhaltungsphase.

Tabelle 4-111: Aufteilung der unerwünschten Ereignisse von besonderem Interesse nach Titrations- und Erhaltungsphase

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag N=509		Placebo N=508	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
Unerwünschtes Ereignis von besonderem Interesse	n	%	n	%
Hämorrhagie **- Titrationsphase**	30	5,9	25	4,9
Hypotonie **- Titrationsphase**	16	3,1	7	1,4
Störung des Auges **- Titrationsphase**	27	5,3	27	5,3
Hämorrhagie **- Erhaltungsphase**	72	14,1	74	14,6
Hypotonie **- Erhaltungsphase**	25	4,9	14	2,8
Störung des Auges **- Erhaltungsphase**	48	9,4	32	6,3
Behandlungseffekt (Selexipag gegenüber Placebo)				
Unerwünschtes Ereignis von besonderem Interesse	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Hämorrhagie **- Titrationsphase**	1,20 [0,72; 2,01]	1,21 [0,70; 2,09]	1,0 [-1,8; 3,8]	0,5329
Hypotonie **- Titrationsphase**	2,28 [0,95; 5,50]	2,32 [0,95; 5,70]	1,8 [-0,1; 3,6]	0,0699
Störung des Auges **- Titrationsphase**	1,00 [0,59; 1,68]	1,00 [0,58; 1,73]	0,0 [-2,8; 2,7]	1,0000
Hämorrhagie **- Erhaltungsphase**	0,97 [0,72; 1,31]	0,97 [0,68; 1,37]	-0,4 [-4,7; 3,9]	0,9022
Hypotonie **- Erhaltungsphase**	1,78 [0,94; 3,39]	1,82 [0,94; 3,55]	2,2 [-0,2; 4,5]	0,0765
Störung des Auges **- Erhaltungsphase**	1,50 [0,97; 2,30]	1,55 [0,97; 2,47]	3,1 [-0,2; 6,4]	0,0699

[1]) Ausschließlich der Patienten, die die Therapie während der Titrationsphase abgebrochen hatten.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA) und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region mit Ausnahme der Analyse für Störungen des Auges keinen Beleg für eine Effektmodifikation für diesen Endpunkt (siehe Abbildung 4-36, Abbildung 4-37 und Abbildung 4-38). Für Störungen des Auges ergibt in der für Deutschland repräsentativen Subgruppe Westeuropa ein erhöhtes Risiko für das Auftreten eines entsprechenden unerwünschten Ereignisses. Da sich der Verdacht eines ursächlichen, biologischen Zusammenhangs zwischen der Anwendung von Selexipag und dem Auftreten von Augenstörungen in der Studie insgesamt nicht bestätigen konnte, ist davon auszugehen, dass das entsprechende Subgruppenergebnis als Zufallsbefund und mögliches Resultat des multiplen Testens zu bewerten ist.

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

4.3.1.3.1.17 Prostacyclin-assoziierte unerwünschte Ereignisse

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-112: Operationalisierung von „Prostacyclin-assoziierte unerwünschte Ereignisse“

Studie	Operationalisierung
GRIPHON	*Definition des Endpunkts* Anteil der Patienten mit mindestens einem Prostacyclin-assoziierten unerwünschten Ereignis im Behandlungszeitraum. Zur Operationalisierung der UE siehe auch Tabelle 4-86. Als Prostacyclin-assoziierte UE wurden UE der folgenden PT definiert: • Kieferschmerz, Kiefergelenk-Syndrom, Arthralgie, muskuloskeletaler Schmerz, Myalgie, Extremitätenschmerz • Flush • Übelkeit, Erbrechen • Durchfall • Kopfschmerzen • Schwindel

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-113: Bewertung des Verzerrungspotentials für „Prostacyclin-assoziierte unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotential auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotential Endpunkt
GRIPHON	niedrig	ja	ja	ja	nein	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes beobachtet werden, es liegen jedoch sonstige das Verzerrungspotential beeinflussende Aspekte vor. So setzt die Auswertung der Ereignisse des

Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen gemäß dem Cochrane Handbook for Systematic Reviews of Interventions jedoch voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden [83]. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer bzw. nicht mehr handhabbarer prostacyclinspezifischer Nebenwirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben.

Insgesamt ist das Verzerrungspotential für diesen Endpunkt als hoch anzusehen. Aufgrund der vergleichsweise längeren medianen Behandlungsdauer in der Selexipag-Gruppe (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchstens von einer Verzerrung der Ergebnisse zuungunsten von Selexipag auszugehen.

Stellen Sie die Ergebnisse für den Endpunkt „Prostacyclin-assoziierte unerwünschte Ereignisse" für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-114: Ergebnisse für „Prostacyclin-assoziierte unerwünschte Ereignisse" aus RCT mit dem zu bewertenden Arzneimittel

Studie: GRIPHON	**Ausgewertete Population: SAF**			
Intervention	Selexipag N=575		Placebo N=577	
Anzahl Patienten mit mindestens einem Prostacyclin-assoziierten unerwünschten Ereignis				
	n	%	n	%
Patienten mit mindestens einem Prostacyclin-assoziierten unerwünschten Ereignis	523	91,0	363	62,9
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem Prostacyclin-assoziierten UE	1,45 [1,35; 1,55]	5,93 [4,26; 8,26]	28,0 [23,5; 32,6]	<0,0001
Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt. UE – Unerwünschtes Ereignis, KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*				

Hauptanalyse

Die Analyse der Prostacyclin-assoziierten unerwünschten Ereignisse wurde auf Grundlage der SAF-Population durchgeführt.

Insgesamt trat im Verlauf der Behandlungsphase von GRIPHON bei 91,0 % der Patienten im Selexipag-Arm und bei 62,9 % der Patienten im Placebo-Arm mindestens ein UE auf. Dies entspricht einer RR [95 % KI] im Vergleich von Selexipag zu Placebo von 1,45 [1,35; 1,55] bzw. einer absoluten RD [95 % KI] von 28 Prozentpunkten [23,5 %-P.; 32,6 %-P.]. Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (Barnards exakter Test: p<0,0001).

Wie aufgrund des Studiendesigns zu erwarten war, ist das Auftreten Prostacyclin-assoziierter UE nicht gleichmäßig über die gesamte Behandlungsdauer von GRIPHON. Die Aufteilung der Prostacyclin-assoziierten UE nach Behandlungsphasen zeigt Tabelle 4-115.

In der Titrationsphase trat im Selexipag-Arm bei 441 Patienten (86,6 % der Patienten, die in beiden Studienphasen behandelt worden waren) und damit bei deutlich mehr Patienten als im

Placebo-Arm (263 Patienten, 51,8 %) ein Prostacyclin-assoziiertes UE auf. Dies entspricht einer Erhöhung des relativen Risikos von 67 % (RR [95 % KI]: 1,67 [1,53; 1,83]) und einer Erhöhung des absoluten Risikos von 34,9 Prozentpunkten (RD [95 % KI]: 34,9 %-P. [29,6 %-P.; 40,1 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant.

Mit diesem Unterschied war aufgrund der für GRIPHON gewählten Titrationsstrategie, in der das Auftreten von Prostacyclin-assoziierten UE Voraussetzung für die Bestimmung der patientenindividuell höchsten tolerierten Dosis war, bereits vor Beginn der Studie zu rechnen gewesen. Tatsächlich verringerte sich der Unterschied zwischen den Behandlungsgruppen bezüglich der Inzidenz Prostacyclin-assoziierter UE in der Erhaltungsphase deutlich. Obwohl diese Phase wesentlich länger war als die Titrationsphase[37], wurde nur bei 367 Patienten (72,1 %) im Selexipag-Arm und bei 238 Patienten (46,9 %) im Placebo-Arm ein Prostacyclin-assoziiertes UE beobachtet. Dies entspricht einer Erhöhung des relativen Risikos um 54 % (RR [95 % KI]: 1,54 [1,38; 1,71]) bzw. einer Erhöhung des absoluten Risikos von 25,3 Prozentpunkten (RD [95 % KI]: 25,3 %-P. [19,4 %-P.; 31,1 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war weiterhin statistisch signifikant (Barnards exakter Test: p<0,0001).

[37] Die Titrationsphase dauerte 12 Wochen, die Erhaltungsphase im Median etwa 59 (Selexipag) bzw. 52 Wochen (Placebo).

Tabelle 4-115: Aufteilung der Prostacyclin-assoziierten unerwünschten Ereignisse nach Titrations- und Erhaltungsphase

Studie: GRIPHON	**Ausgewertete Population: SAF[1]**			
Intervention	Selexipag N=509		Placebo N=508	
Anzahl Patienten mit mindestens einem unerwünschten Ereignis				
	n	%	n	%
Patienten mit mindestens einem Prostacyclin-assoziierten UE in der **Titrationsphase**	441	86,6	263	51,8
Patienten mit mindestens einem Prostacyclin-assoziierten UE in der **Erhaltungsphase**	367	72,1	238	46,9
Behandlungseffekt (Selexipag gegenüber Placebo)				
	RR [95 % KI]	OR [95 % KI]	RD [95 % KI] (Prozentpunkte)	p
Patienten mit mindestens einem Prostacyclin-assoziierten UE in der **Titrationsphase**	1,67 [1,53; 1,83]	6,04 [4,44; 8,23]	34,9 [29,6; 40,1]	<0,0001
Patienten mit mindestens einem Prostacyclin-assoziierten UE in der **Erhaltungsphase**	1,54 [1,38; 1,71]	2,93 [2,26; 3,81]	25,3 [19,4; 31,1]	<0,0001

[1]) Ausschließlich der Patienten, die die Therapie während der Titrationsphase abgebrochen hatten.

Dargestellt sind nicht-adjustierte, übliche RR und OR. Sofern Werte in der Vierfeldertafel niedriger als 5 waren, wurden exakte Werte berechnet. Für die Berechnung der RR wurde hierfür die Farrington-Manning-Methode verwendet, für die Berechnung der OR die PROC-FREQ-Prozedur. Die Berechnung der RD erfolgte mithilfe der exakten Clopper-Pearce-Methode. Der p-Wert schließlich wurde anhand Barnards exaktem Test berechnet. Sämtliche Analysen wurden *post hoc* durchgeführt.

KI – Konfidenzintervall, RR – Risk Ratio, OR – Odds Ratio, RD – absolute Risikodifferenz, SAF – *safety (population)*

Bei den am häufigsten aufgetretenen Prostacyclin-assoziierten UE handelt es sich durchweg um Störungen mit vergleichsweise geringer klinischer Relevanz, die (sofern sie in der Titrationsphase aufgetreten waren) bereits durch die Dosisreduktion oder mit üblichen Therapiemethoden gut zu beherrschen waren. Am weitaus häufigsten wurde in beiden Behandlungsgruppen und Behandlungsphasen das UE Kopfschmerz beobachtet (64,4 % bzw. 39,9 % (Selexipag) vs. 28,3 % bzw. 19,5 % (Placebo), vgl. auch Tabelle 4-116). Ebenfalls häufig traten auch die UE Diarrhö, Übelkeit und Kieferschmerz auf.

Tabelle 4-116: Prostacyclin-assoziierte unerwünschte Ereignisse in GRIPHON, aufgeteilt nach bevorzugter Bezeichnung

	Selexipag				**Placebo**			
Phase	Titrations-phase		Erhaltungs phase		Titrations-phase		Erhaltungs phase	
N	509		509		508		508	
	n	%	n	%	n	%	n	%
Patienten mit mindestens einem Prostacyclin-assoziierten UE	441	86,6	367	72,1	263	51,8	238	56,9
Kopfschmerz	328	64,4	203	39,9	144	28,3	99	19,5
Diarrhö	182	35,8	151	29,7	62	12,2	67	13,2
Übelkeit	148	29,1	100	19,6	65	12,8	51	10,0
Kieferschmerz	134	26,3	105	20,6	20	3,9	20	3,9
Myalgie	76	14,9	48	9,4	24	4,7	16	3,1
Extremitätenschmerz	73	14,3	66	13,0	26	5,1	31	6,1
Erbrechen	72	14,1	39	7,7	21	4,1	28	5,5
Flush	57	11,2	52	10,2	19	3,7	16	3,1
Schwindel	50	9,8	53	10,4	42	8,3	58	11,4
Arthralgie	37	7,3	47	9,2	25	4,9	27	5,3
Muskuloskeletaler Schmerz	14	2,8	11	2,2	3	0,6	10	2,0
Kiefergelenk-Syndrom	3	0,6	3	0,6	2	0,4	1	0,2

Grundlage der Analyse war die SAF-Population, wobei jedoch nur diejenigen Patienten in die Analyse eingeschlossen wurden, die sowohl in der Titrations- als auch der Erhaltungsphase behandelt wurden.
UE – Unerwünschtes Ereignis, SAF – *safety* (population)

Nur wenige der in der gesamten Behandlungsphase beobachteten Prostacyclin-assoziierten UE wurden gleichzeitig auch als SUE eingeschätzt (vgl. Tabelle 4-117). Keines dieser SUE führte zum Versterben des Patienten.

Tabelle 4-117: Prostacyclin-assoziierte schwerwiegende unerwünschte Ereignisse in GRIPHON, aufgeteilt nach bevorzugter Bezeichnung

	Selexipag		**Placebo**	
N	575		577	
	n	%	n	%
Patienten mit mindestens einem SUE	252	43,8	272	47,1
Kopfschmerz	2	0,3	-	
Diarrhö	3	0,5	3	0,5
Übelkeit	1	0,2	-	
Erbrechen	2	0,3	2	0,3
Extremitätenschmerz	2	0,3	-	
Myalgie	2	0,3	-	
Flush	1	0,2	-	

Grundlage der Analyse war die SAF-Population.
SUE – Schwerwiegendes unerwünschtes Ereignis, SAF – *safety* (population)

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Meta-Analyse

Es wurde lediglich eine Studie in die Nutzenbewertung von Selexipag aufgenommen, die Notwendigkeit einer Meta-Analyse entfällt daher.

Übertragbarkeit auf den deutschen Versorgungskontext

Bei GRIPHON handelt es sich um eine multinationale Studie, die in 181 Zentren in 39 Ländern durchgeführt wurde. Der überwiegende Anteil der Studienpopulation waren weiße bzw. kaukasische Patienten (65,5 %). Aufgrund der ebenfalls überwiegend weißen Bevölkerung der Bundesrepublik Deutschland ist daher aus physiologischer/genetischer Sicht von einer vollständigen Übertragbarkeit der Ergebnisse auszugehen.

Weiterhin zeigt ein Vergleich mit den Basiswertcharakteristika der für dieses Dossier angefertigten Analyse der deutschen PAH-Population von COMPERA [82], dass in GRIPHON zwar tendenziell etwas jüngere (48,1 Jahre gegenüber 62,7 Jahre in COMPERA)

und mehr weibliche Patienten (79,6 % gegenüber 63,5 % in COMPERA) eingeschlossen wurden, die Basiswertcharakteristika aber grundsätzlich vergleichbar sind.

Schließlich zeigen auch die Ergebnisse der Subgruppenanalyse nach Region für diesen Endpunkt keinen Beleg für eine Effektmodifikation (siehe Abbildung 4-39).

Insgesamt ist aus den genannten Gründen von einer vollständigen Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.2 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Die Darstellung der Subgruppen erfolgt für die Studie GRIPHON getrennt für jeden Endpunkt und gemäß der in Abschnitt 4.2.5.5 beschriebenen Methodik. Die Ergebnisse der Subgruppenanalysen sind zusammenfassend in Form von Forest-Plots berichtet. Analysiert wurden die folgenden Subgruppen:

- PAH-spezifische Therapie zum Zeitpunkt der Basiswerterhebung (ERA-Monotherapie, PDE-5-I-Monotherapie, ERA- und PDE-5-I-Therapie, keine PAH-spezifische Therapie)
- WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung (I/II, III/IV)
- Geschlecht (männlich, weiblich)
- Rasse (kaukasisch/hispanisch, asiatisch, schwarz)
- Alter zum Zeitpunkt des Screenings (<65 Jahre, ≥65 Jahre)
- PAH-Ätiologie zum Zeitpunkt der Basiswerterhebung (IPAH/HPAH/HIV-assoziiert/Drogen- oder Toxin-induziert, assoziiert mit Bindegewebserkrankungen, assoziiert mit angeborener Herzschwäche)

- Geographische Region (Nordamerika, Westeuropa/Australien, Osteuropa, Asien, Lateinamerika)[38]

A priori definiert, d. h. bereits im Studienprotokoll festgelegt, waren ausschließlich die Subgruppenanalysen für den primären Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“. Zusätzlich wurden im statistischen Analyseplan von GRIPHON Subgruppenanalysen für die Endpunkte „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ und „Absolute Veränderung der 6-MWD“ definiert. Die Subgruppenanalysen für alle weiteren Endpunkte wurden *post hoc* für das Nutzendossier erstellt.

Sofern sich mehrere Zeitpunkte für die Analyse eines Endpunkts anbieten, wurde derjenige Zeitpunkt für die Subgruppenanalysen angewendet, der auch bereits für die jeweilige Hauptanalyse herangezogen wurde. Beispielsweise erfolgte die Subgruppenanalyse für die absolute Veränderung der 6-MWD als Analyse der Veränderung in Woche 26 im Vergleich zum Basiswert.

Bei der Testung einer größeren Anzahl statistischer Hypothesen ist weiterhin zu beachten, dass eine gewisse Anzahl der durchgeführten Tests aufgrund der angenommenen Irrtumswahrscheinlichkeit zu einem falsch positiven Ergebnis führen. In den für dieses Dossier durchgeführten Subgruppenanalysen wurden bei 140 Testungen (20 Forest Plots à 7 Subgruppen) 7 positive Belege ($\alpha<0{,}05$) für eine Effektmodifikation beobachtet. Geht man davon aus, dass alle Testungen vollständig unabhängig sind und die Auftretens-wahrscheinlichkeit falsch positiver Tests 5 % beträgt, wäre aufgrund des Zufalls selbst bei einer vollständigen Abwesenheit tatsächlicher Effektmodifikationen mit einer Wahr-scheinlichkeit von 55 % mindestens eine solche Anzahl positiver Tests zu erwarten gewesen [90]. Die Wahrscheinlichkeit, dass kein positiver Test beobachtet würde, beträgt unter den gleichen Bedingungen 0,08 % [90].

Ein positiver Interaktionstest ($p<0{,}05$) wird in diesem Dossier daher lediglich als Ausgangspunkt für die Überprüfung einer möglichen Effektmodifikation herangezogen. Liegt ein positiver Interaktionstest vor ($p<0{,}05$), werden daher zusätzlich die folgenden Kriterien geprüft, die aus dem Entwurf der EMA-Leitlinien zur Untersuchung von Subgruppen in konfirmatorischen klinischen Studien sowie einem Gutachten des medizinischen Dienstes der Spitzenverbände der Krankenkassen zu Subgruppenanalysen abgeleitet wurden [74, 75]:

- Besteht eine medizinisch, biologisch oder physiologisch plausible Erklärung für die mögliche Effektmodifikation?

[38] Da der CAMPHOR-Fragebogen nur für eine begrenzte Anzahl von Ländern verfügbar ist, musste die Auswertung der entsprechenden Endpunkte auf die QAS-Population eingeschränkt werden. Diese Population beinhaltet jedoch keine Patienten in Asien oder Lateinamerika, sodass auch keine CAMPHOR-Subgruppen-ergebnisse für diese Regionen vorliegen.

- Liegen konsistente Ergebnisse zur möglichen Effektmodifikation über mehrere Endpunkte vor?
- Liegen zur möglichen Effektmodifikation bereits ähnliche Ergebnisse aus anderen Studien vor?

Ingesamt gesehen liegen für einzelne Endpunkte und in einzelnen Subgruppenanalysen zwar positive Interaktionstests vor, in den wenigsten dieser Fälle lässt sich jedoch eine plausible Erklärung für die Effektmodifikation finden bzw. war die mutmaßliche Effektmodifikation als solche bereits in der Literatur beschrieben. Schließlich ergibt sich auch für keinen der möglichen Effektmodifikatoren ein konsistentes Bild über mehrere Endpunkte.

Die beobachteten positiven Interaktionstests werden daher als zufällige Effekte bewertet und nicht als Nachweis einer tatsächlichen Effektmodifikation. Insgesamt wird von einem homogenen Effekt über alle Subgruppen ausgegangen.

4.3.1.3.2.1 Zeit bis zum ersten CEC-bestätigten Morbiditäts-/Mortalitätsereignis

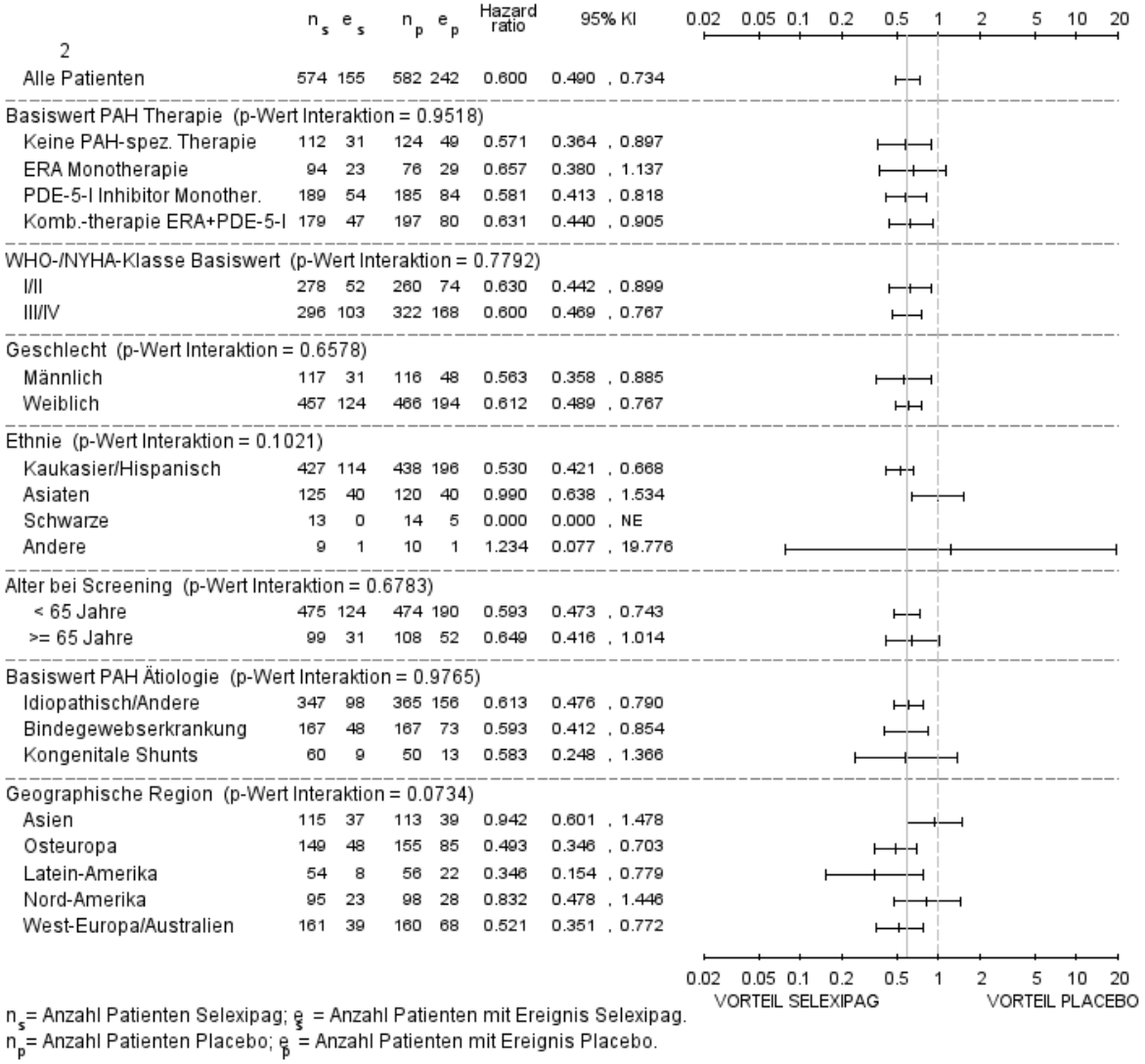
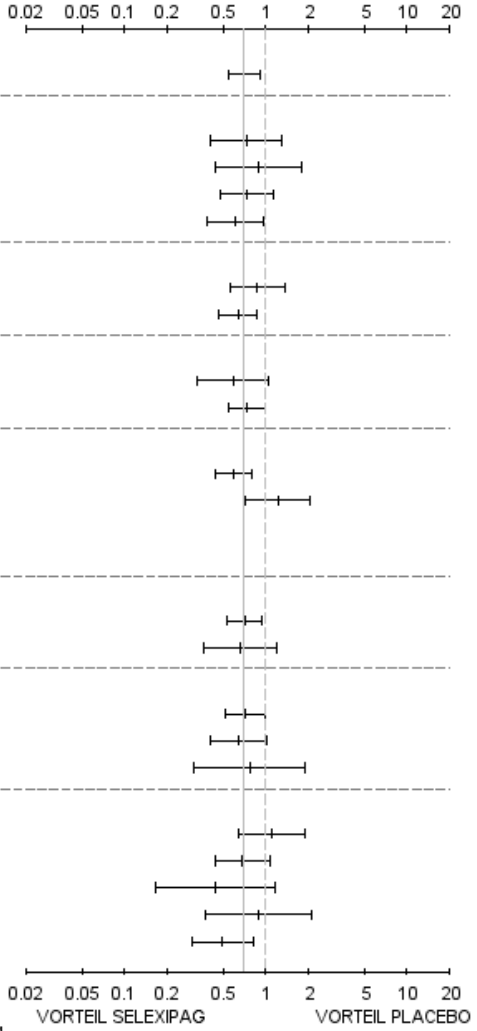


Abbildung 4-20: Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis – Forest-Plot der Subgruppenanalysen
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen waren bereits im statistischen Analyseplan prädefiniert. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.2 Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH

	n_s	e_s	n_p	e_p	Hazard ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	574	102	582	137	0.700	0.542 , 0.905
Basiswert PAH Therapie (p-Wert Interaktion = 0.8481)						
Keine PAH-spez. Therapie	112	20	124	25	0.729	0.405 , 1.313
ERA Monotherapie	94	16	76	15	0.889	0.439 , 1.802
PDE-5-I Inhibitor Monother.	189	37	185	46	0.731	0.474 , 1.127
Komb.-therapie ERA+PDE-5-I	179	29	197	51	0.608	0.385 , 0.960
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion = 0.2570)						
I/II	278	37	260	38	0.878	0.558 , 1.381
III/IV	296	65	322	99	0.642	0.470 , 0.879
Geschlecht (p-Wert Interaktion = 0.3839)						
Männlich	117	19	116	29	0.591	0.331 , 1.056
Weiblich	457	83	466	108	0.738	0.554 , 0.982
Ethnie (p-Wert Interaktion = 0.1452)						
Kaukasier/Hispanisch	427	71	438	109	0.596	0.442 , 0.804
Asiaten	125	31	120	25	1.228	0.725 , 2.080
Schwarze	13	0	14	3	0.000	0.000 , NE
Andere	9	0	10	0	NE	NE , NE
Alter bei Screening (p-Wert Interaktion = 0.8757)						
< 65 Jahre	475	85	474	109	0.711	0.535 , 0.944
>= 65 Jahre	99	17	108	28	0.659	0.360 , 1.206
Basiswert PAH Ätiologie (p-Wert Interaktion = 0.9213)						
Idiopathisch/Andere	347	62	365	84	0.724	0.521 , 1.005
Bindegewebserkrankung	167	31	167	43	0.650	0.409 , 1.032
Kongenitale Shunts	60	9	50	10	0.773	0.313 , 1.909
Geographische Region (p-Wert Interaktion = 0.2184)						
Asien	115	28	113	25	1.113	0.649 , 1.908
Osteuropa	149	34	155	44	0.684	0.437 , 1.071
Latein-Amerika	54	6	56	13	0.441	0.167 , 1.162
Nord-Amerika	95	10	98	11	0.898	0.379 , 2.126
West-Europa/Australien	161	24	160	44	0.497	0.302 , 0.817



n_s = Anzahl Patienten Selexipag; e_s = Anzahl Patienten mit Ereignis Selexipag.
n_p = Anzahl Patienten Placebo; e_p = Anzahl Patienten mit Ereignis Placebo.

Abbildung 4-21: Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH – Forest-Plot der Subgruppenanalysen CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen waren bereits im statistischen Analyseplan prädefiniert. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.3 Zeit bis zum Tod (Studienschluss)

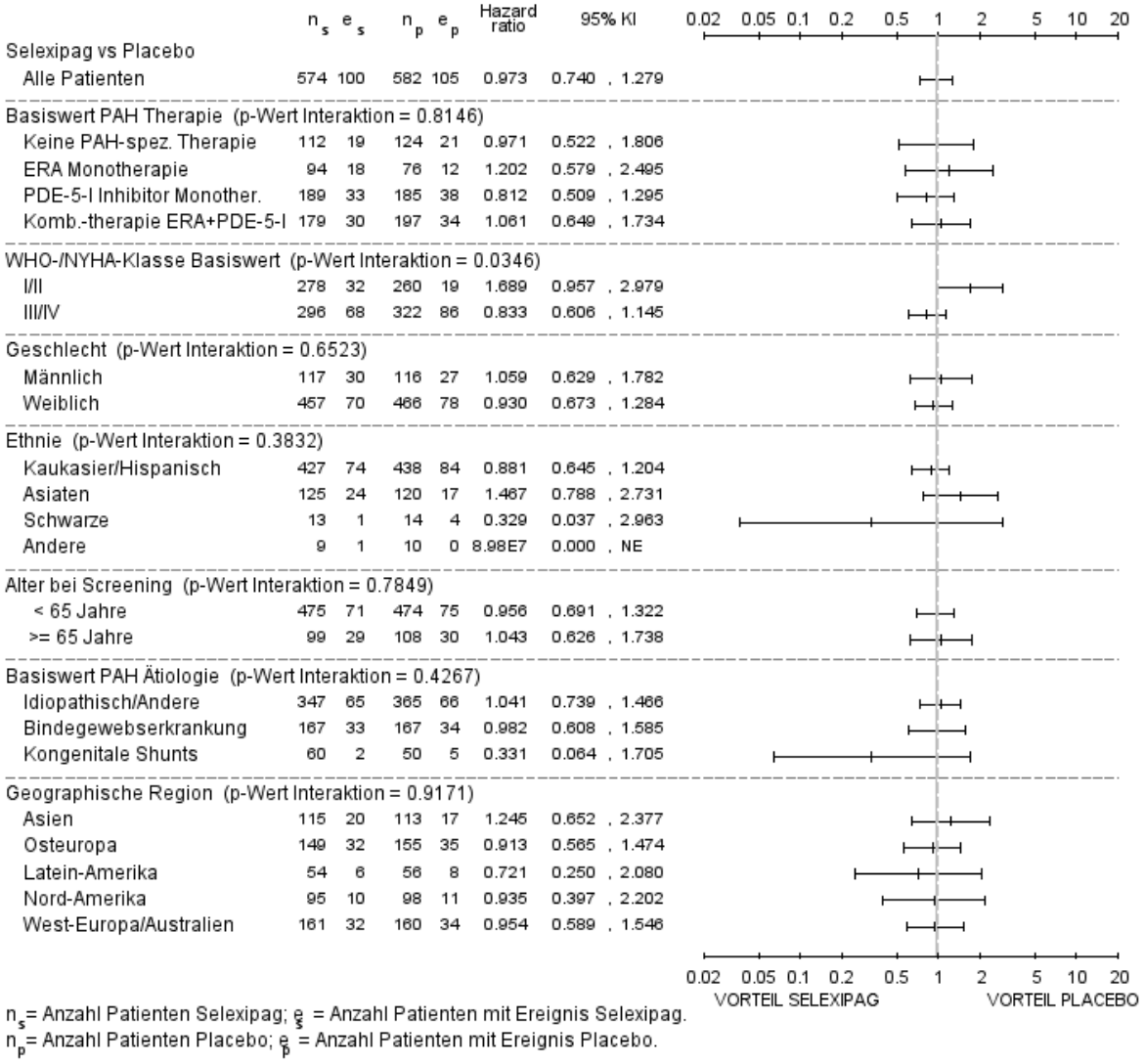


Abbildung 4-22: Zeit bis zum Tod (Studienschluss) – Forest-Plot der Subgruppenanalysen
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

Lediglich in der Subgruppenanalyse nach WHO-/NYHA-Klasse wird im Interaktionstest ein statistisch signifikantes Ergebnis (p=0,0346) erreicht, was auf eine mögliche Effektmodifikation hindeutet. In dieser Subgruppenanalyse zeigen Patienten in WHO-/NYHA-Klasse I oder II eine höhere Hazard Ratio als Patienten in den Klassen III oder IV (1,69 vs. 0,83), was einer Effektumkehr in den definierten Subgruppen entsprechen würde. Die Beobachtung, dass die Mortalität bei Patienten in den WHO-/NYHA-Klassen I und II unter Selexipag auffallend hoch war, wurde auch im Rahmen des Zulassungsprozesses extensiv diskutiert [73]. Die EMA kam hierbei zu folgendem Schluss: „*Weder bezogen auf die Gesamtpopulation noch auf Patienten in WHO-/NYHA-Klasse II wurde ein Sicherheitssignal beobachtet und es existiert auch keine klinische Rationale, warum Patienten in WHO-/NYHA-*

Klasse II mit besserer Herzfunktion für einen möglicherweise schädigenden Effekt durch Selexipag anfälliger sein sollten als Patienten in WHO-/NYHA-Klasse III/IV mit verschlechterter Herzfunktion, was damit einen ursächlichen Zusammenhang zwischen dem beobachteten Anstieg der Mortalität und Selexipag höchst unwahrscheinlich macht.“[39]

Zudem ist keine Studie in der PAH bekannt, bei der ein ähnlicher Effekt beobachtet worden wäre.

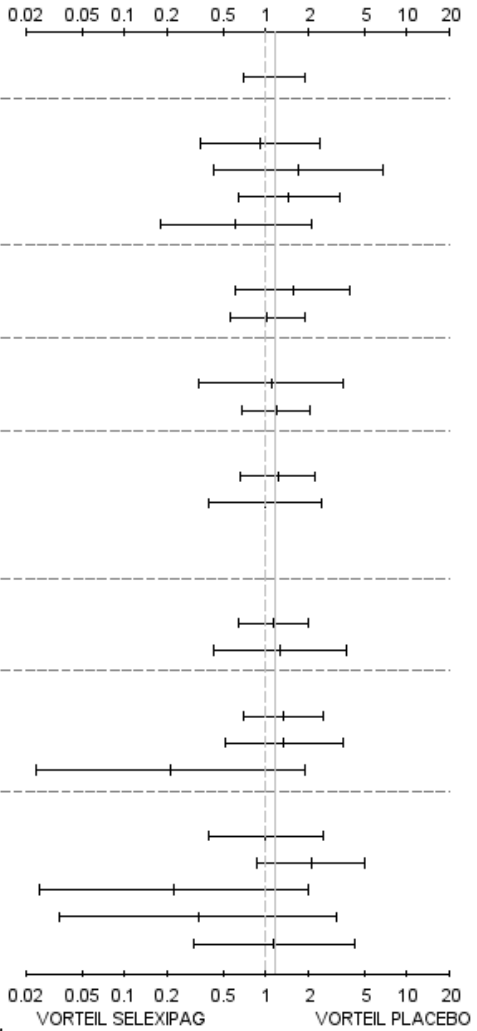
Mit Ausnahme der Analyse der Hämorrhagie-assoziierten UE (vgl. Abschnitt 4.3.1.3.2.16) wurde schließlich in keinem weiteren Endpunkt eine ähnliche Effektmodifikation beobachtet, sodass auch nicht von einem konsistenten Effekt über mehrere Endpunkte ausgegangen werden kann.

In Übereinstimmung mit der Einschätzung der Zulassungsbehörde wird das vorliegende Ergebnis daher als Ergebnis der informativen Zensierung der Todesfälle oder als zufälliger Effekt, aber nicht als Nachweis einer tatsächlichen Effektmodifikation bewertet. Insgesamt wird von einem homogenen Effekt über alle Subgruppen ausgegangen.

[39] Englischer Originaltext: “*No safety signal, overall or specifically related to FC II, was detected and there is no clinical rationale why patients in functional class II with better cardiac function should be more sensitive to a putative harmful effect of selexipag than patients in FC III/IV with deteriorated cardiac function making a causal relationship between the observed increased mortality and selexipag in the first event analysis is highly unlikely*” [73] (S. 89)

4.3.1.3.2.4 Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)

	n_s	e_s	n_p	e_p	Hazard ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	574	33	582	27	1.159	0.697 , 1.928
Basiswert PAH Therapie (p-Wert Interaktion = 0.5911)						
Keine PAH-spez. Therapie	112	8	124	8	0.920	0.345 , 2.454
ERA Monotherapie	94	6	76	3	1.725	0.431 , 6.900
PDE-5-I Inhibitor Monother.	189	15	185	9	1.460	0.638 , 3.337
Komb.-therapie ERA+PDE-5-I	179	4	197	7	0.614	0.179 , 2.098
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion = 0.4706)						
I/II	278	12	260	7	1.561	0.614 , 3.966
III/IV	296	21	322	20	1.036	0.561 , 1.912
Geschlecht (p-Wert Interaktion = 0.8326)						
Männlich	117	6	116	5	1.095	0.333 , 3.595
Weiblich	457	27	466	22	1.187	0.676 , 2.085
Ethnie (p-Wert Interaktion = 0.9835)						
Kaukasier/Hispanisch	427	24	438	18	1.224	0.664 , 2.256
Asiaten	125	9	120	9	1.000	0.397 , 2.520
Schwarze	13	0	14	0	NE	NE , NE
Andere	9	0	10	0	NE	NE , NE
Alter bei Screening (p-Wert Interaktion = 0.8225)						
< 65 Jahre	475	26	474	21	1.139	0.641 , 2.025
>= 65 Jahre	99	7	108	6	1.276	0.429 , 3.800
Basiswert PAH Ätiologie (p-Wert Interaktion = 0.2620)						
Idiopathisch/Andere	347	22	365	16	1.344	0.706 , 2.560
Bindegewebserkrankung	167	10	167	7	1.348	0.513 , 3.542
Kongenitale Shunts	60	1	50	4	0.211	0.024 , 1.887
Geographische Region (p-Wert Interaktion = 0.2804)						
Asien	115	9	113	9	1.006	0.399 , 2.535
Osteuropa	149	17	155	7	2.098	0.870 , 5.064
Latein-Amerika	54	1	56	4	0.225	0.025 , 2.025
Nord-Amerika	95	1	98	3	0.332	0.035 , 3.195
West-Europa/Australien	161	5	160	4	1.155	0.310 , 4.304



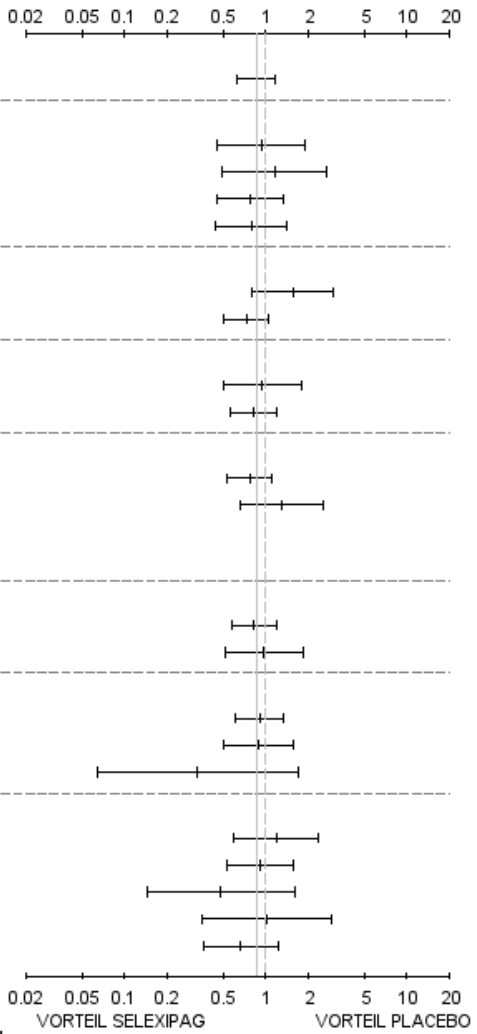
n_s = Anzahl Patienten Selexipag; e_s = Anzahl Patienten mit Ereignis Selexipag.
n_p = Anzahl Patienten Placebo; e_p = Anzahl Patienten mit Ereignis Placebo.

Abbildung 4-23: Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT) – Forest-Plot der Subgruppenanalysen
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.5 Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)

	n_s	e_s	n_p	e_p	Hazard ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	574	70	582	83	0.860	0.626 , 1.183
Basiswert PAH Therapie (p-Wert Interaktion = 0.8596)						
Keine PAH-spez. Therapie	112	14	124	16	0.940	0.459 , 1.925
ERA Monotherapie	94	13	76	9	1.161	0.496 , 2.716
PDE-5-I Inhibitor Monother.	189	24	185	29	0.774	0.451 , 1.330
Komb.-therapie ERA+PDE-5-I	179	19	197	29	0.797	0.446 , 1.421
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion = 0.0507)						
I/II	278	22	260	14	1.561	0.799 , 3.052
III/IV	296	48	322	69	0.730	0.505 , 1.055
Geschlecht (p-Wert Interaktion = 0.6770)						
Männlich	117	19	116	19	0.950	0.503 , 1.796
Weiblich	457	51	466	64	0.824	0.570 , 1.191
Ethnie (p-Wert Interaktion = 0.5988)						
Kaukasier/Hispanisch	427	51	438	66	0.772	0.536 , 1.113
Asiaten	125	19	120	15	1.310	0.665 , 2.578
Schwarze	13	0	14	2	0.000	0.000 , NE
Andere	9	0	10	0	NE	NE , NE
Alter bei Screening (p-Wert Interaktion = 0.6909)						
< 65 Jahre	475	52	474	63	0.832	0.576 , 1.202
>= 65 Jahre	99	18	108	20	0.975	0.516 , 1.844
Basiswert PAH Ätiologie (p-Wert Interaktion = 0.5161)						
Idiopathisch/Andere	347	46	365	53	0.915	0.617 , 1.358
Bindegewebserkrankung	167	22	167	25	0.890	0.502 , 1.579
Kongenitale Shunts	60	2	50	5	0.331	0.064 , 1.705
Geographische Region (p-Wert Interaktion = 0.6005)						
Asien	115	17	113	15	1.196	0.597 , 2.396
Osteuropa	149	25	155	27	0.923	0.536 , 1.590
Latein-Amerika	54	4	56	8	0.482	0.145 , 1.603
Nord-Amerika	95	7	98	7	1.024	0.359 , 2.920
West-Europa/Australien	161	17	160	26	0.664	0.360 , 1.224



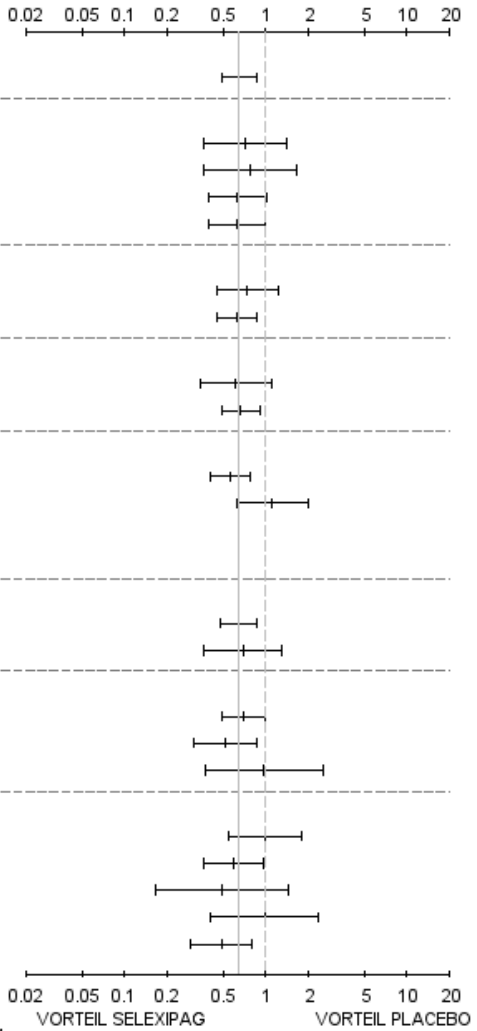
n_s = Anzahl Patienten Selexipag; e_s = Anzahl Patienten mit Ereignis Selexipag.
n_p = Anzahl Patienten Placebo; e_p = Anzahl Patienten mit Ereignis Placebo.

Abbildung 4-24: Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss) – Forest-Plot der Subgruppenanalysen
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.6 Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH

	n_s	e_s	n_p	e_p	Hazard ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	574	86	582	124	0.653	0.496 , 0.860
Basiswert PAH Therapie (p-Wert Interaktion = 0.9685)						
Keine PAH-spez. Therapie	112	15	124	19	0.717	0.364 , 1.412
ERA Monotherapie	94	13	76	14	0.771	0.361 , 1.643
PDE-5-I Inhibitor Monother.	189	29	185	42	0.632	0.394 , 1.015
Komb.-therapie ERA+PDE-5-I	179	29	197	49	0.635	0.401 , 1.005
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion = 0.5493)						
I/II	278	29	260	35	0.748	0.457 , 1.224
III/IV	296	57	322	89	0.627	0.450 , 0.875
Geschlecht (p-Wert Interaktion = 0.6517)						
Männlich	117	19	116	28	0.612	0.341 , 1.098
Weiblich	457	67	466	96	0.670	0.490 , 0.915
Ethnie (p-Wert Interaktion = 0.2408)						
Kaukasier/Hispanisch	427	61	438	99	0.565	0.410 , 0.777
Asiaten	125	25	120	22	1.124	0.634 , 1.994
Schwarze	13	0	14	3	0.000	0.000 , NE
Andere	9	0	10	0	NE	NE , NE
Alter bei Screening (p-Wert Interaktion = 0.7946)						
< 65 Jahre	475	70	474	99	0.645	0.475 , 0.876
>= 65 Jahre	99	16	108	25	0.693	0.369 , 1.299
Basiswert PAH Ätiologie (p-Wert Interaktion = 0.4710)						
Idiopathisch/Andere	347	54	365	76	0.700	0.494 , 0.992
Bindegewebserkrankung	167	23	167	40	0.513	0.307 , 0.859
Kongenitale Shunts	60	9	50	8	0.978	0.376 , 2.547
Geographische Region (p-Wert Interaktion = 0.3401)						
Asien	115	22	113	22	0.993	0.550 , 1.794
Osteuropa	149	26	155	39	0.595	0.362 , 0.978
Latein-Amerika	54	5	56	10	0.493	0.168 , 1.445
Nord-Amerika	95	10	98	10	0.991	0.410 , 2.395
West-Europa/Australien	161	23	160	43	0.488	0.294 , 0.809



n_s = Anzahl Patienten Selexipag; e_s = Anzahl Patienten mit Ereignis Selexipag.
n_p = Anzahl Patienten Placebo; e_p = Anzahl Patienten mit Ereignis Placebo.

Abbildung 4-25: Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH – Forest-Plot der Subgruppenanalysen
CEC-bestätigte Ereignisse, die vor dem 16. August 2011 begannen, sind in die Analyse eingeschlossen. Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.7 Symptome der PAH

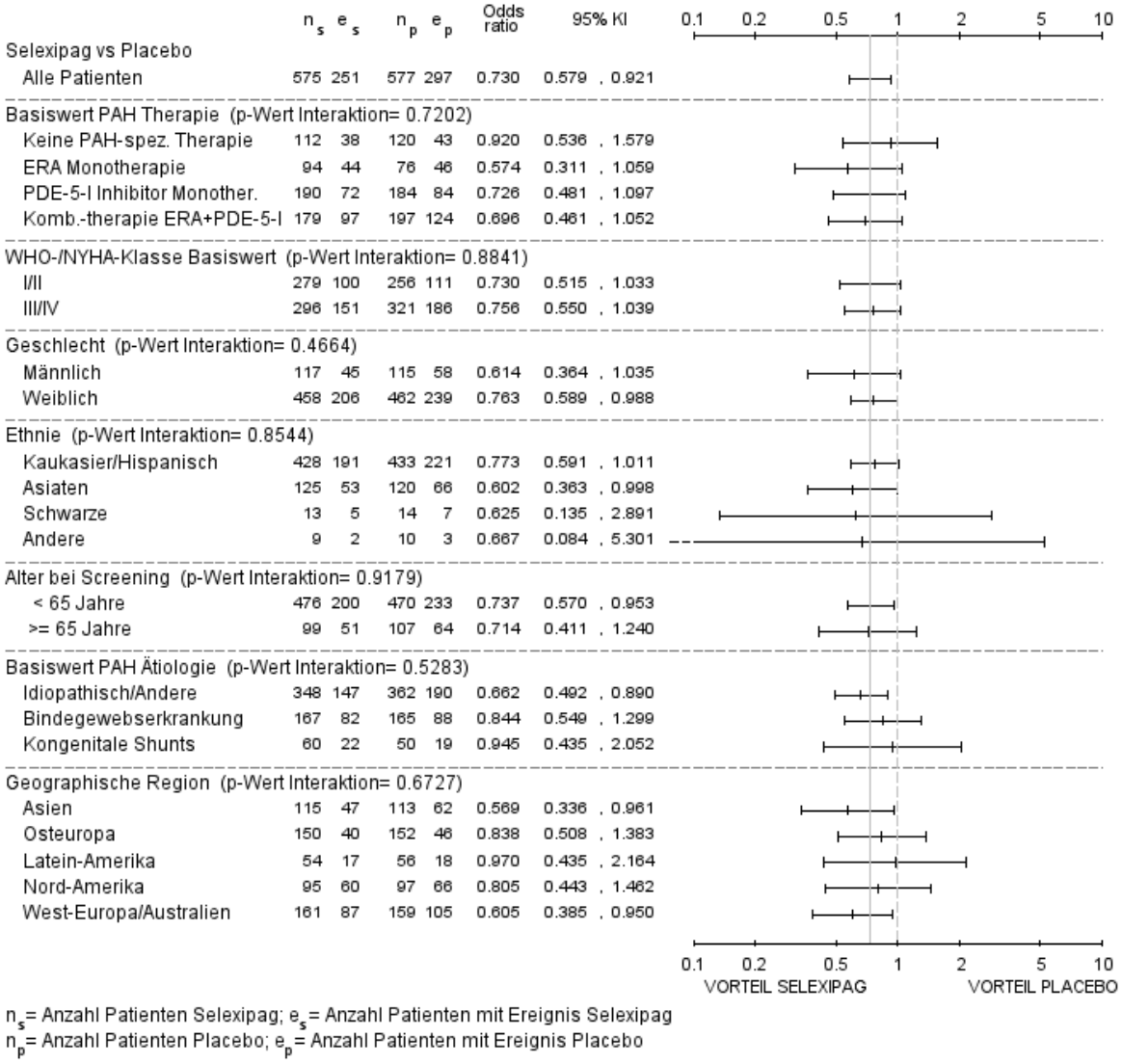


Abbildung 4-26: Symptome der PAH – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.8 Absolute Veränderung der 6-MWD

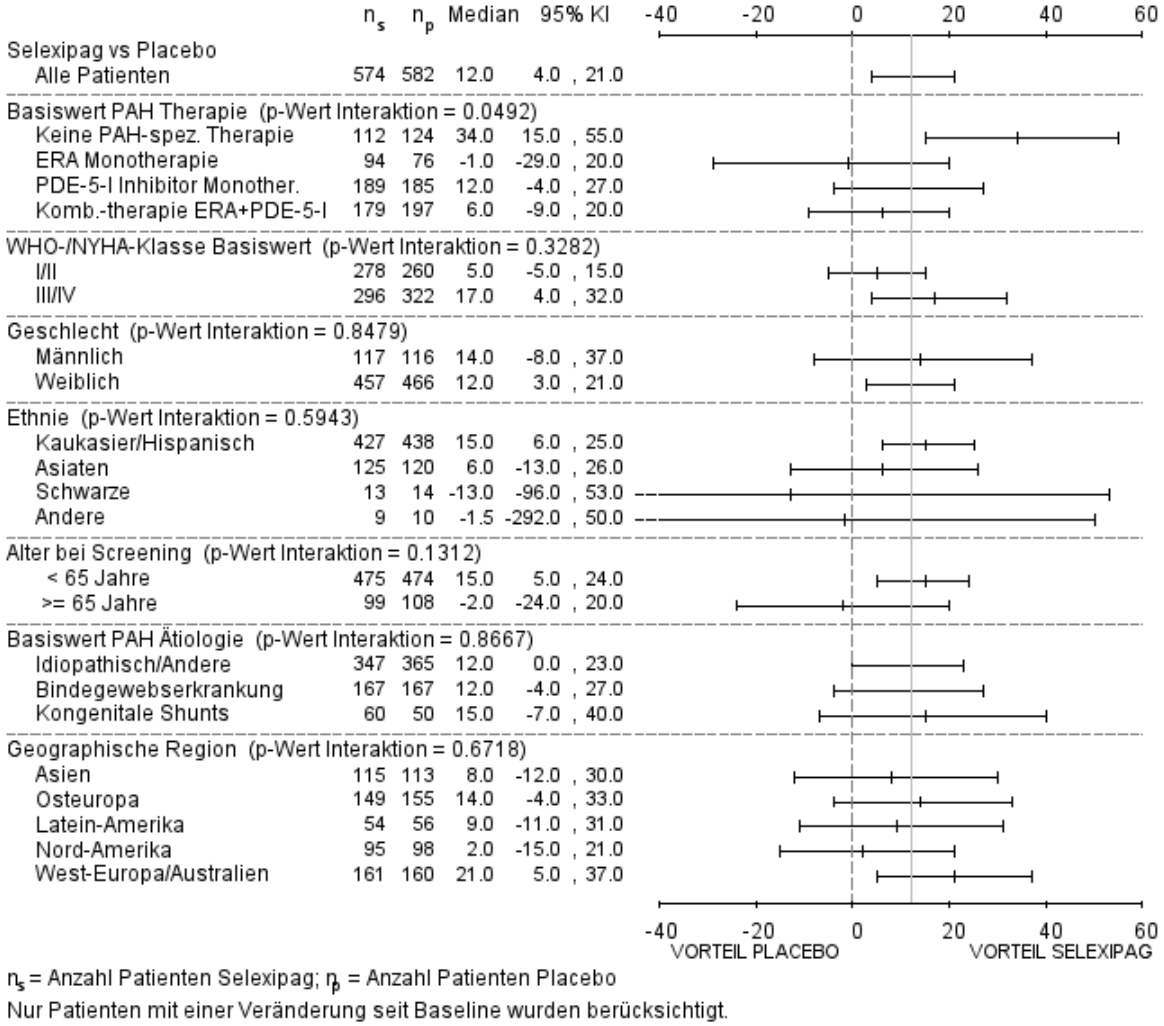


Abbildung 4-27: Absolute Veränderung der 6-MWD – Forest-Plot der Subgruppenanalysen
Analysierte Population: FAS. Mit Ausnahme der Interaktionstests waren sämtliche in der Abbildung dargestellten Analysen bereits im statistischen Analyseplan prädefiniert. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall

Bei der Auswertung der Subgruppenanalysen für den Endpunkt „Absolute Veränderung der 6-MWD“ ergibt sich in den Subgruppen nach PAH-spezifischer Therapie zum Zeitpunkt der Basiswerterhebung ein positiver Interaktionstest (p=0,0492). In dieser Subgruppenanalyse zeigen Patienten, die zum Zeitpunkt der Basiswerterhebung therapienaiv waren, einen deutlich höheren Behandlungseffekt durch Selexipag als Patienten mit PAH-spezifischer Therapie (Median: 34 m bei Patienten ohne Therapie vs. −1 m bis 12 m bei Patienten mit PAH-spezifischer Therapie).

Wie auch bereits in Abschnitt 4.3.1.3.1.8 ausgeführt, ist eine tatsächliche Effektmodifikation aufgrund des *Ceiling*-Effekts bei bereits vorbehandelten Patienten durchaus plausibel und wurde in der Vergangenheit auch schon in anderen Studien beobachtet [86]. Diese Effektmodifikation ist anzunehmenderweise vornehmlich auf die Limitationen des 6-MWT

zurückzuführen und nicht auf eine unterschiedliche Wirksamkeit von Selexipag, wenn es entweder in Monotherapie oder in Kombination mit anderen Wirkstoffen angewendet wird. Dies zeigt sich auch dadurch, dass für keinen der anderen Wirksamkeitsendpunkte ein positiver Test auf Interaktion in der Subgruppe „PAH-spezifische Therapie“ vorliegt.

4.3.1.3.2.9 Ausbleibende Verschlechterung der WHO-/NYHA-Klasse

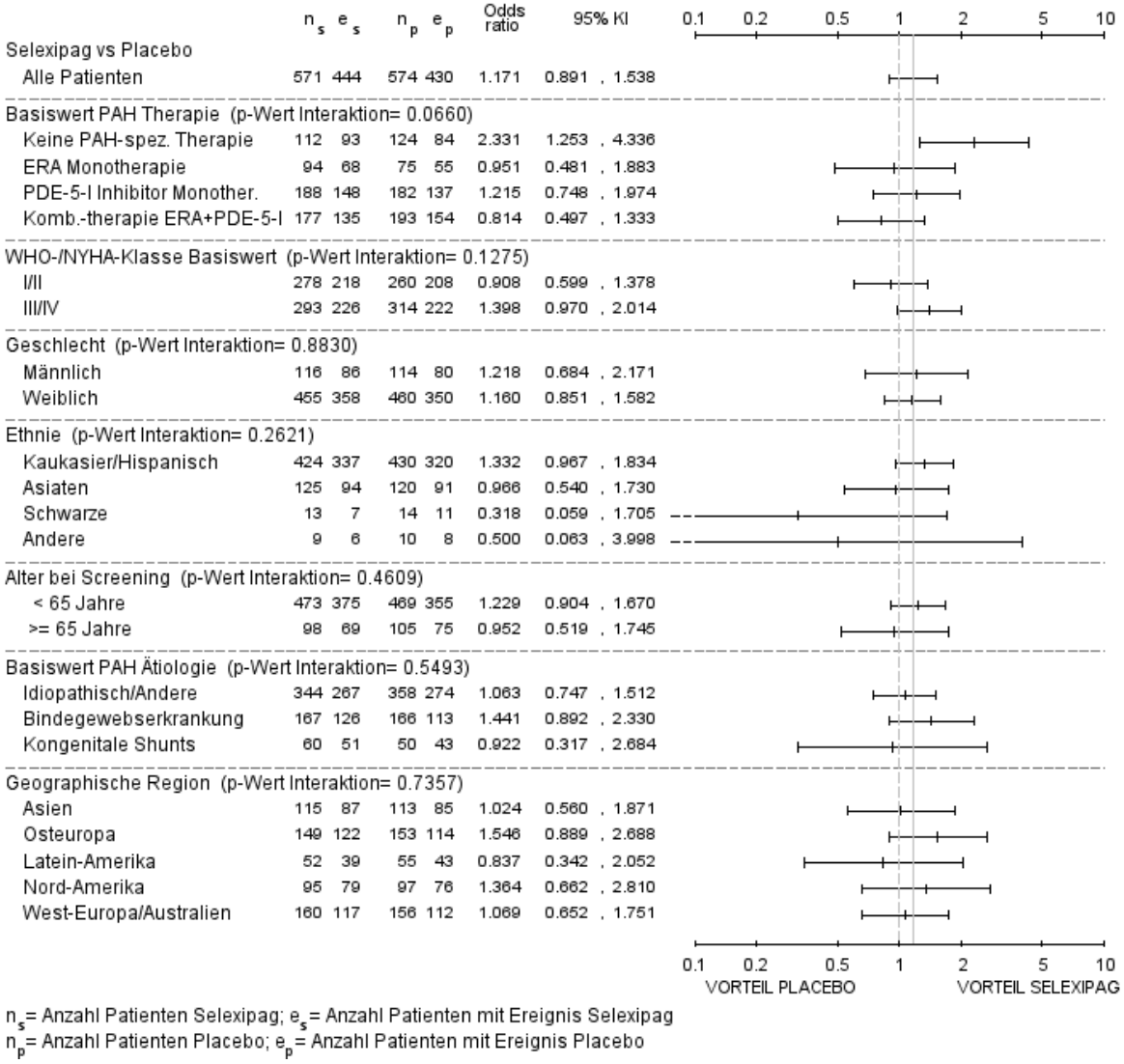


Abbildung 4-28: Ausbleibende Verschlechterung der WHO-/NYHA-Klasse – Forest-Plot der Subgruppenanalysen
Analysierte Population: FAS (ausschließlich von Patienten in WHO-/NYHA-Klasse IV). Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.10 Absolute Veränderung des Borg-Dyspnoe-Index

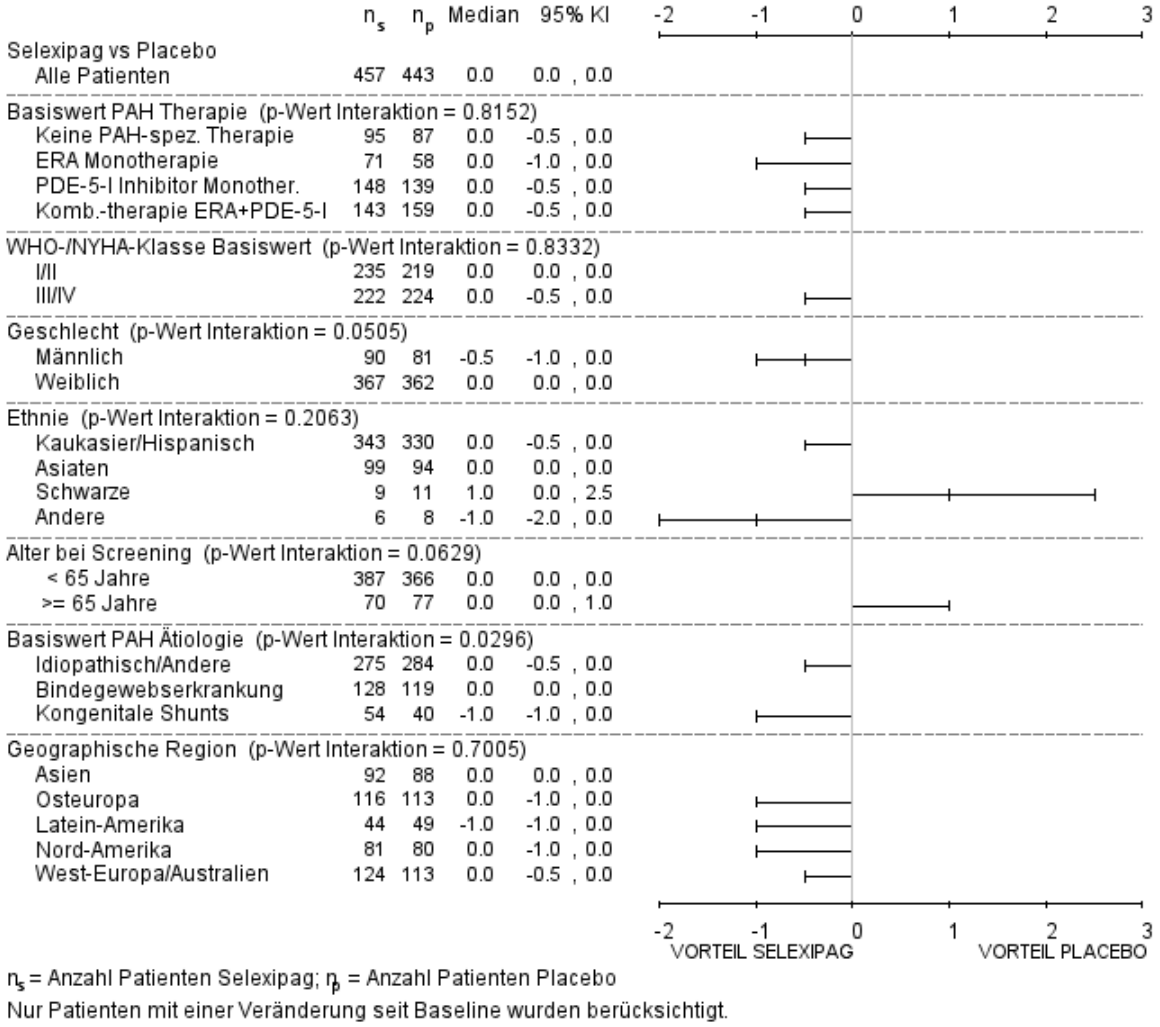


Abbildung 4-29: Absolute Veränderung des Borg-Dyspnoe-Index – Forest-Plot der Subgruppenanalysen
Analysierte Population: FAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall

Für den Endpunkt „Absolute Veränderung des Borg-Dyspnoe-Index“ deuten die Ergebnisse der Interaktionstests auf eine mögliche Effektmodifikation in den Subgruppen nach PAH-Ätiologie hin (p=0,0296). In dieser Subgruppenanalyse zeigen Patienten mit einer PAH aufgrund eines angeborenen, korrigierten Herzfehlers, die mit Selexipag behandelt wurden, eine tendenzielle Verbesserung im Borg-Dyspnoe-Index im Vergleich zu Placebo-behandelten Patienten, während der Therapieeffekt bei Patienten mit idiopathischer oder anderer PAH und bei Patienten mit einer Bindegewebsstörungs-assoziierten PAH ausgeglichen ist. Es ergibt sich jedoch kein qualitativer Unterschied zwischen den Subgruppeneffekten, da auch der Therapieeffekt bei Patienten mit PAH aufgrund eines angeborenen, korrigierten Herzfehlers statistisch nicht signifikant ist.

Weiterhin besteht keine belastbare medizinische Rationale, die die mutmaßliche Effektmodifikation erklären könnte, und es sind auch keine Studien bekannt, in denen ein ähnlicher Effekt bereits beobachtet wurde. Zudem lag nur für den Endpunkt „Absolute Veränderung des Borg-Dyspnoe-Index“ ein positiver Interaktionstest in den Subgruppen nach PAH-Ätiologie vor, sodass auch nicht von einer konsistenten Effektmodifikation über mehrere Endpunkte gesprochen werden kann.

Im Hinblick auf die Problematik der Multiplizität bei der Testung einer größeren Anzahl statistischer Hypothesen und aufgrund der Tatsache, dass mit Ausnahme des positiven Interaktionstests kein weiteres der oben definierten Kriterien erfüllt ist, wird das vorliegende Ergebnis daher als zufälliger Effekt und nicht als Nachweis einer tatsächlichen Effektmodifikation bewertet. Insgesamt wird von einem homogenen Effekt über alle Subgruppen ausgegangen.

4.3.1.3.2.11 Absolute Veränderung des CAMPHOR "Symptome"-Scores

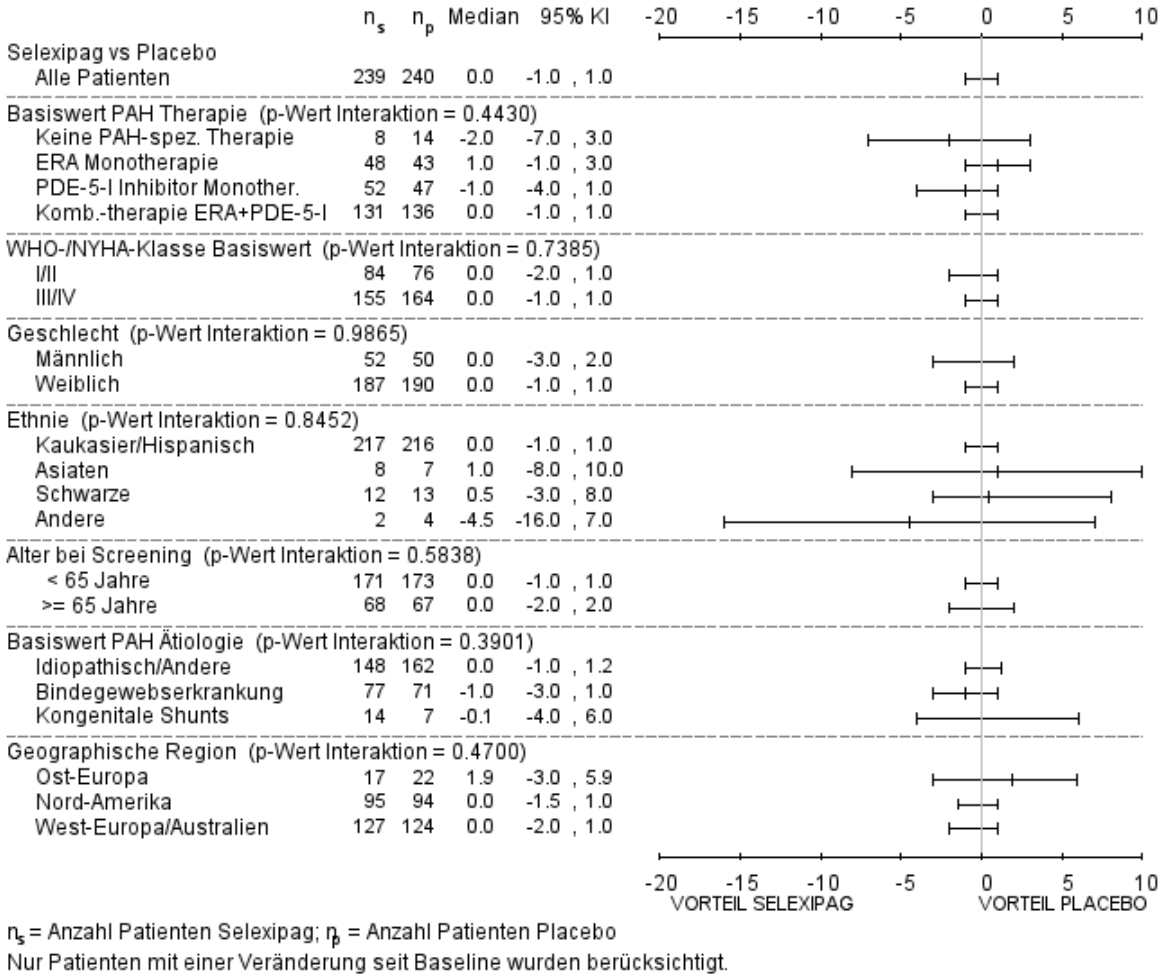


Abbildung 4-30: Absolute Veränderung des CAMPHOR „Symptome“-Scores – Forest-Plot der Subgruppenanalysen
Analysierte Population: QAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.12 Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores

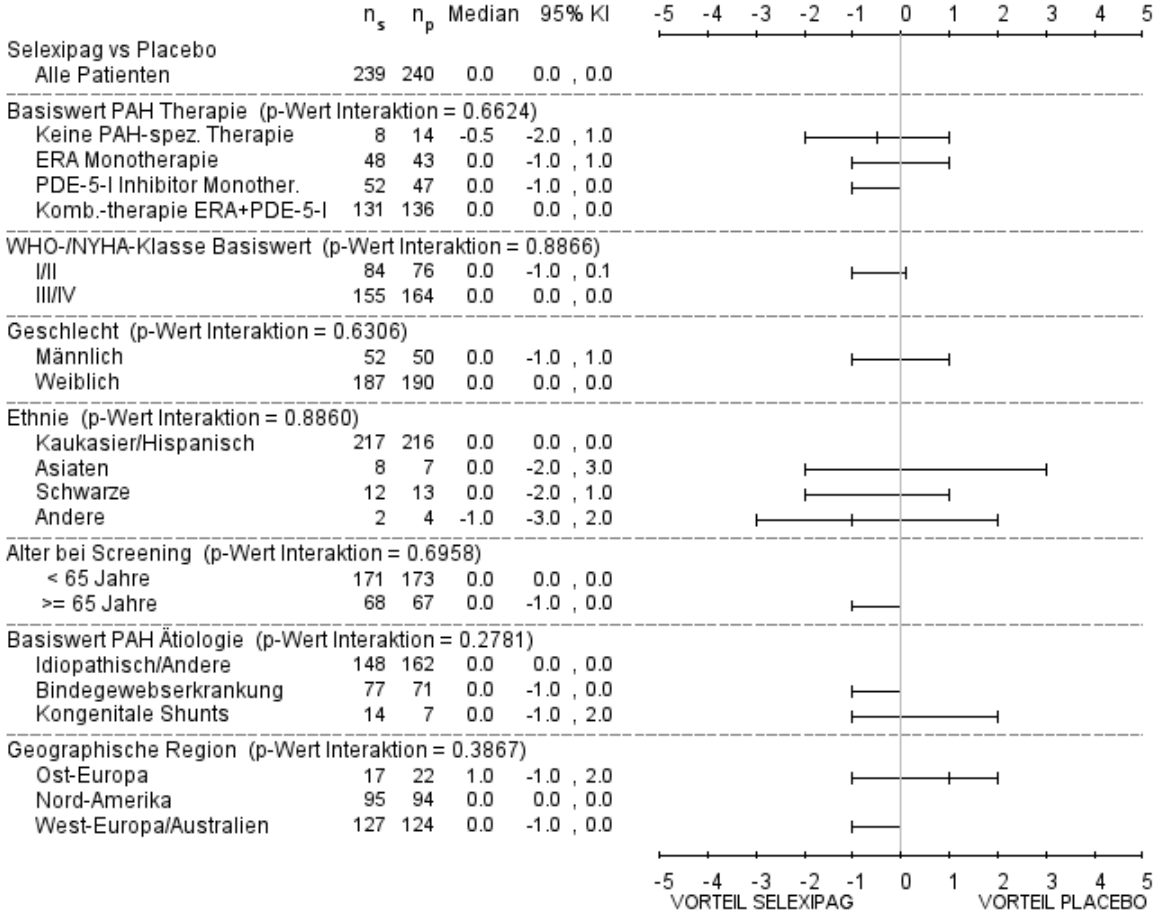


Abbildung 4-31: Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores – Forest-Plot der Subgruppenanalysen
Analysierte Population: QAS. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.13 Unerwünschte Ereignisse

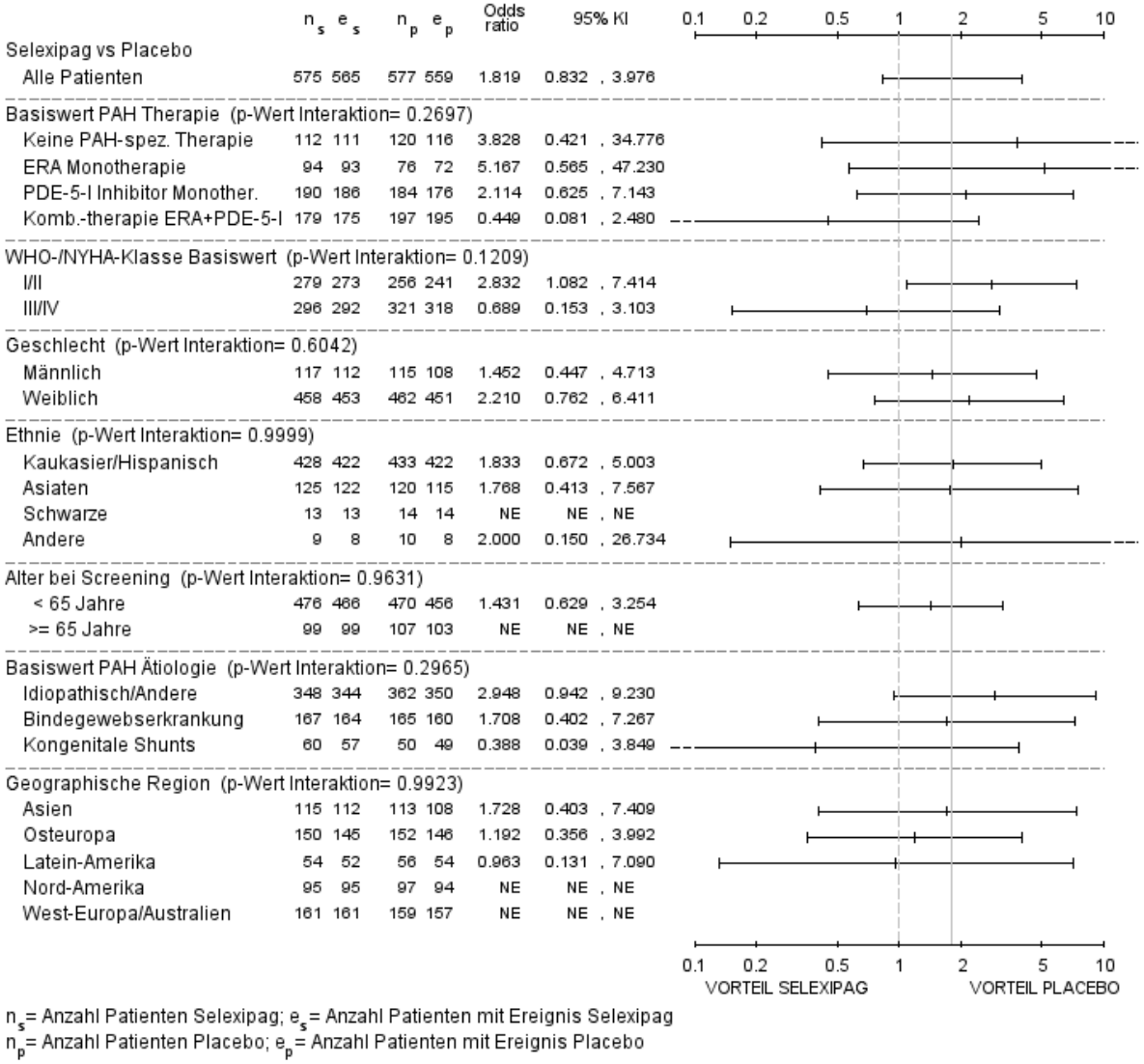


Abbildung 4-32: Unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

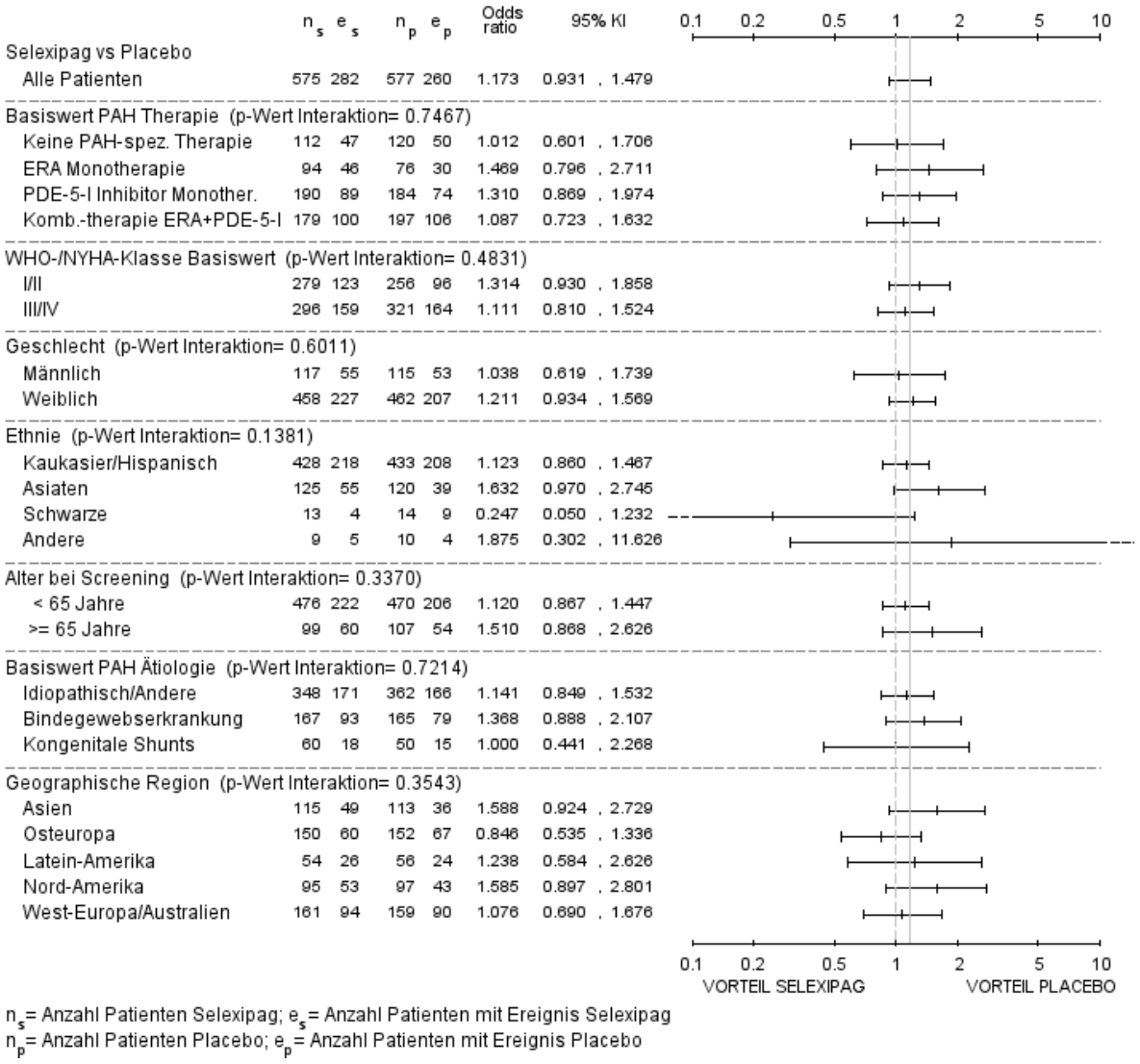


Abbildung 4-33: Unerwünschte Ereignisse schwerer Intensität – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.14 Schwerwiegende unerwünschte Ereignisse

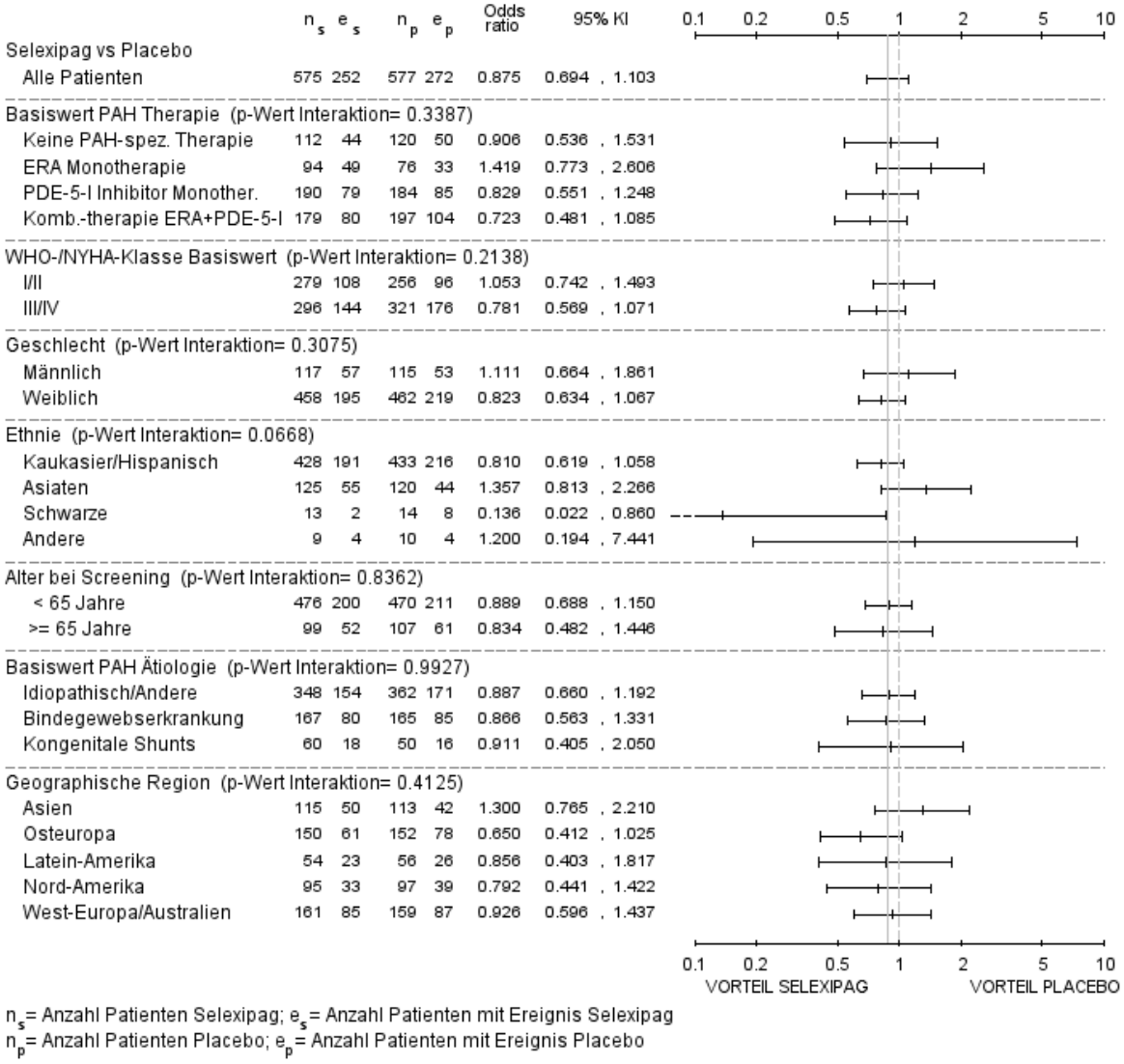


Abbildung 4-34: Schwerwiegende unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.2.15 Therapieabbrüche wegen unerwünschter Ereignisse

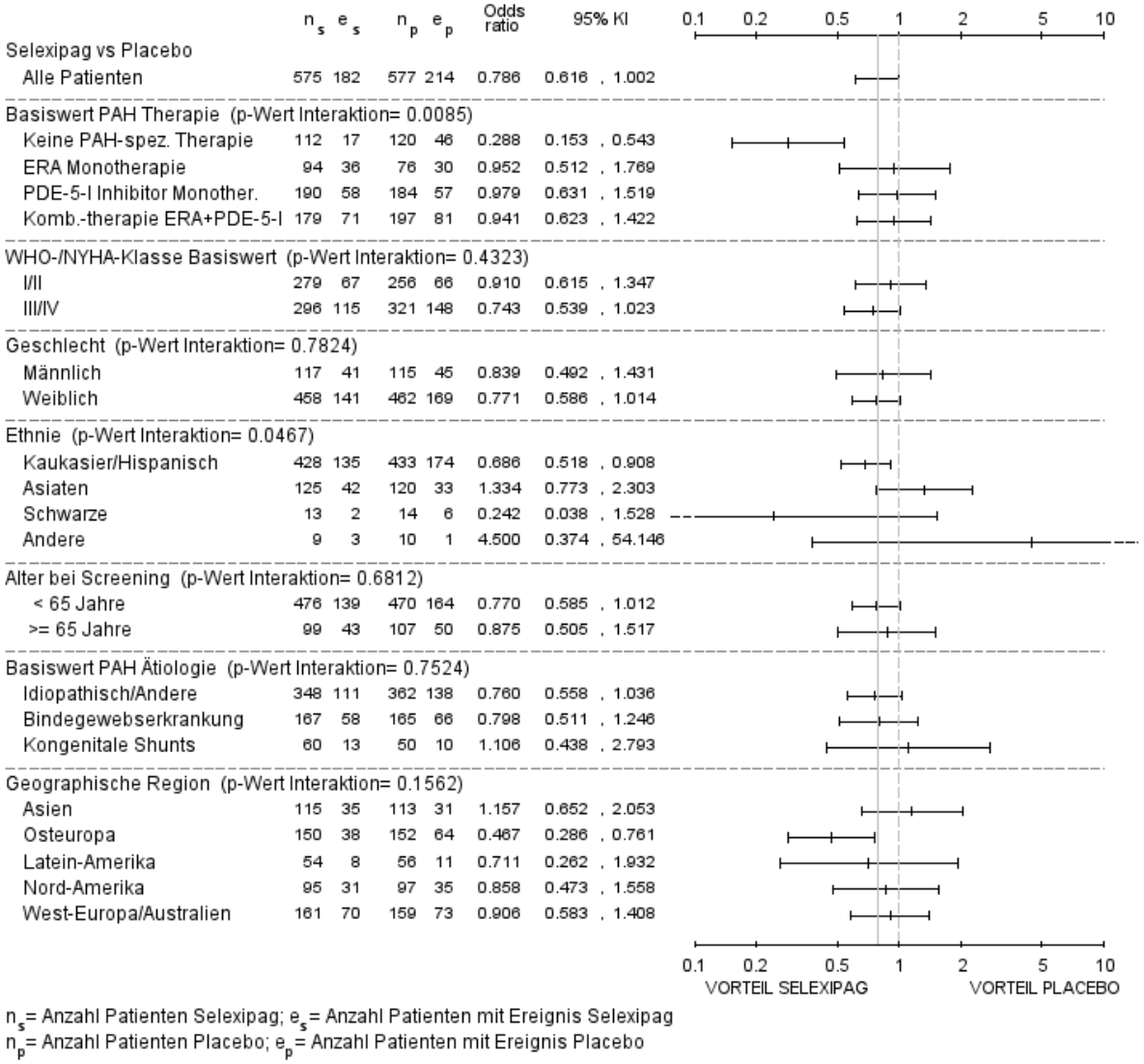


Abbildung 4-35: Therapieabbrüche wegen unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In der Auswertung der Interaktionstests für den Endpunkt „Therapieabbrüche wegen eines unerwünschten Ereignisses“ zeigen die Tests für die Subgruppen nach PAH-spezifischer Therapie zum Zeitpunkt der Basiswerterhebung sowie nach Ethnie ein positives Testergebnis.

Im Fall des Testergebnisses zu den Subgruppen nach PAH-spezifischer Therapie zum Zeitpunkt der Basiswerterhebung (p=0,0085) zeigen Patienten ohne eine solche Therapie einen deutlichen niedrigeren Effektschätzer als Patienten mit PAH-spezifischer Therapie (OR: 0,288 vs. 0,941-0,979). Eine zugrunde liegende medizinische Rationale ergibt sich nicht direkt, jedoch wäre denkbar, dass bei therapienaiven Patienten im Placebo-Arm, die somit *de facto* unbehandelt sind, häufiger PAH-bedingte UE auftreten, die schließlich zu einem

Therapieabbruch führen. Gegen diese Hypothese spricht, dass die Rate der Therapieabbrüche im Placebo-Arm in allen Subgruppen vergleichbar ist, während die Rate der Therapieabbrüche im Selexipag-Arm in der Subgruppe der therapienaiven Patienten deutlich niedriger liegt.

Mit Ausnahme des Endpunkts „Absolute Veränderung der 6-MWD“ (vgl. Abschnitt 4.3.1.3.2.8), bei dem es sich nicht um einen Sicherheits-, sondern um einen Wirksamkeitsendpunkt handelt, wurde schließlich in keinem weiteren Endpunkt eine ähnliche Effektmodifikation aufgrund PAH-spezifischer Therapie beobachtet, sodass auch nicht von einem konsistenten Effekt über mehrere Endpunkte ausgegangen werden kann.

In Bezug auf die Subgruppen nach Rasse/Ethnie (p=0,0467) zeigen Patienten, die kaukasischer oder hispanischer Abstammung sind, eine vergleichsweise niedrigere Odds Ratio als Patienten, die asiatischer Abstammung sind (OR: 0,686 vs. 1,334). Über Patienten schwarzer Abstammung oder anderer Abstammung lässt sich aufgrund der sehr geringen Patientenzahlen in den jeweiligen Subgruppen keine valide Aussage treffen.

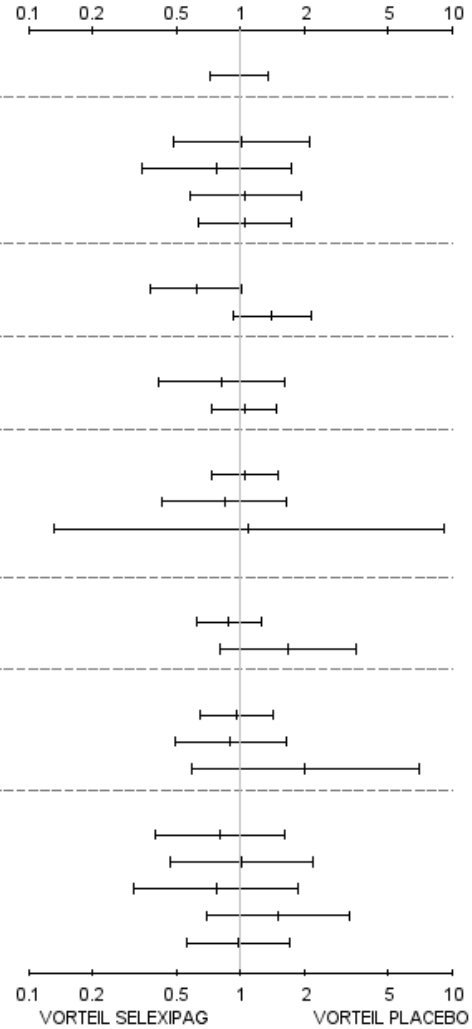
Es existiert jedoch keine medizinische Rationale, die eine solche Effektumkehr in Abhängigkeit von der kaukasischen/hispanischen bzw. asiatischen Abstammung erklären könnte. Zudem sind in der PAH keine Studien bekannt, bei denen ein ähnlicher Effekt beobachtet wurde.

Schließlich wurde in keinem weiteren Endpunkt eine ähnliche Effektmodifikation aufgrund von Ethnie/Rasse beobachtet, sodass auch nicht von einem konsistenten Effekt über mehrere Endpunkte ausgegangen werden kann.

Im Hinblick auf die Problematik der Multiplizität bei der Testung einer größeren Anzahl statistischer Hypothesen und aufgrund der Tatsache, dass mit Ausnahme des positiven Interaktionstests kein weiteres der oben definierten Kriterien sicher erfüllt ist, werden die vorliegenden Ergebnisse daher als zufällige Effekte und nicht als Nachweis einer tatsächlichen Effektmodifikation bewertet. Insgesamt wird von einem homogenen Effekt über alle Subgruppen ausgegangen.

4.3.1.3.2.16 Unerwünschte Ereignisse von besonderem Interesse

	n_s	e_s	n_p	e_p	Odds ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	575	90	577	91	0.991	0.722 , 1.361
Basiswert PAH Therapie (p-Wert Interaktion= 0.9294)						
Keine PAH-spez. Therapie	112	16	120	17	1.010	0.483 , 2.110
ERA Monotherapie	94	14	76	14	0.775	0.344 , 1.745
PDE-5-I Inhibitor Monother.	190	25	184	23	1.061	0.578 , 1.945
Komb.-therapie ERA+PDE-5-I	179	35	197	37	1.051	0.629 , 1.757
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion= 0.0126)						
I/II	279	31	256	43	0.619	0.377 , 1.018
III/IV	296	59	321	48	1.416	0.931 , 2.152
Geschlecht (p-Wert Interaktion= 0.5290)						
Männlich	117	18	115	21	0.814	0.408 , 1.622
Weiblich	458	72	462	70	1.045	0.730 , 1.494
Ethnie (p-Wert Interaktion= 0.9578)						
Kaukasier/Hispanisch	428	69	433	67	1.050	0.728 , 1.514
Asiaten	125	19	120	21	0.845	0.429 , 1.665
Schwarze	13	2	14	2	1.091	0.130 , 9.124
Andere	9	0	10	1	NE	NE , NE
Alter bei Screening (p-Wert Interaktion= 0.1237)						
< 65 Jahre	476	70	470	77	0.880	0.619 , 1.251
>= 65 Jahre	99	20	107	14	1.682	0.798 , 3.545
Basiswert PAH Ätiologie (p-Wert Interaktion= 0.4975)						
Idiopathisch/Andere	348	57	362	61	0.967	0.651 , 1.435
Bindegewebserkrankung	167	24	165	26	0.897	0.491 , 1.638
Kongenitale Shunts	60	9	50	4	2.029	0.585 , 7.037
Geographische Region (p-Wert Interaktion= 0.7777)						
Asien	115	17	113	20	0.807	0.398 , 1.634
Osteuropa	150	14	152	14	1.015	0.466 , 2.209
Latein-Amerika	54	11	56	14	0.767	0.313 , 1.882
Nord-Amerika	95	18	97	13	1.510	0.694 , 3.287
West-Europa/Australien	161	30	159	30	0.985	0.562 , 1.726



n_s= Anzahl Patienten Selexipag; e_s= Anzahl Patienten mit Ereignis Selexipag
n_p= Anzahl Patienten Placebo; e_p= Anzahl Patienten mit Ereignis Placebo

Abbildung 4-36: Unerwünschte Ereignisse von besonderem Interesse: Hämorrhagie – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

Die Auswertung der Interaktionstests für Hämorrhagie-assoziierte UE ergibt, dass in den Subgruppen nach WHO-/NYHA-Klasse eine potentielle Effektmodifikation besteht. Die Odds Ratios der beiden analysierten Subgruppen unterscheiden sich hierbei sowohl stark voneinander als auch vom Gesamteffektschätzer der Studie. Während in der Gesamtstudien-population eine ausgeglichene Odds Ratio vorliegt (0,99), ergibt sich in der Subgruppe der Patienten in den WHO-/NYHA-Klassen I und II eine Odds Ratio zugunsten von Selexipag (0,62). In der Subgruppe der Patienten in den WHO-/NYHA-Klassen III und IV wiederum wurde eine Odds Ratio zuungunsten von Selexipag beobachtet (1,42).

Eine belastbare medizinische Rationale, die eine solche Effektumkehr erklären könnte, existiert nicht. Ebensowenig sind in der PAH Studien bekannt, bei denen ein ähnlicher Effekt beobachtet wurde.

Mit Ausnahme des Endpunkts „Zeit bis zum Tod (Studienschluss)“ (vgl. Abschnitt 4.3.1.3.2.3) wurde schließlich in keinem weiteren Endpunkt eine ähnliche Effektmodifikation aufgrund PAH-spezifischer Therapie beobachtet, sodass auch nicht von einem konsistenten Effekt über mehrere Endpunkte ausgegangen werden kann.

Im Hinblick auf die Problematik der Multiplizität bei der Testung einer größeren Anzahl statistischer Hypothesen und aufgrund der Tatsache, dass mit Ausnahme des positiven Interaktionstests kein weiteres der oben definierten Kriterien erfüllt ist, wird das vorliegende Ergebnis daher als zufälliger Effekt und nicht als Nachweis einer tatsächlicher Effekt-modifikation bewertet. Insgesamt wird von einem homogenen Effekt über alle Subgruppen ausgegangen.

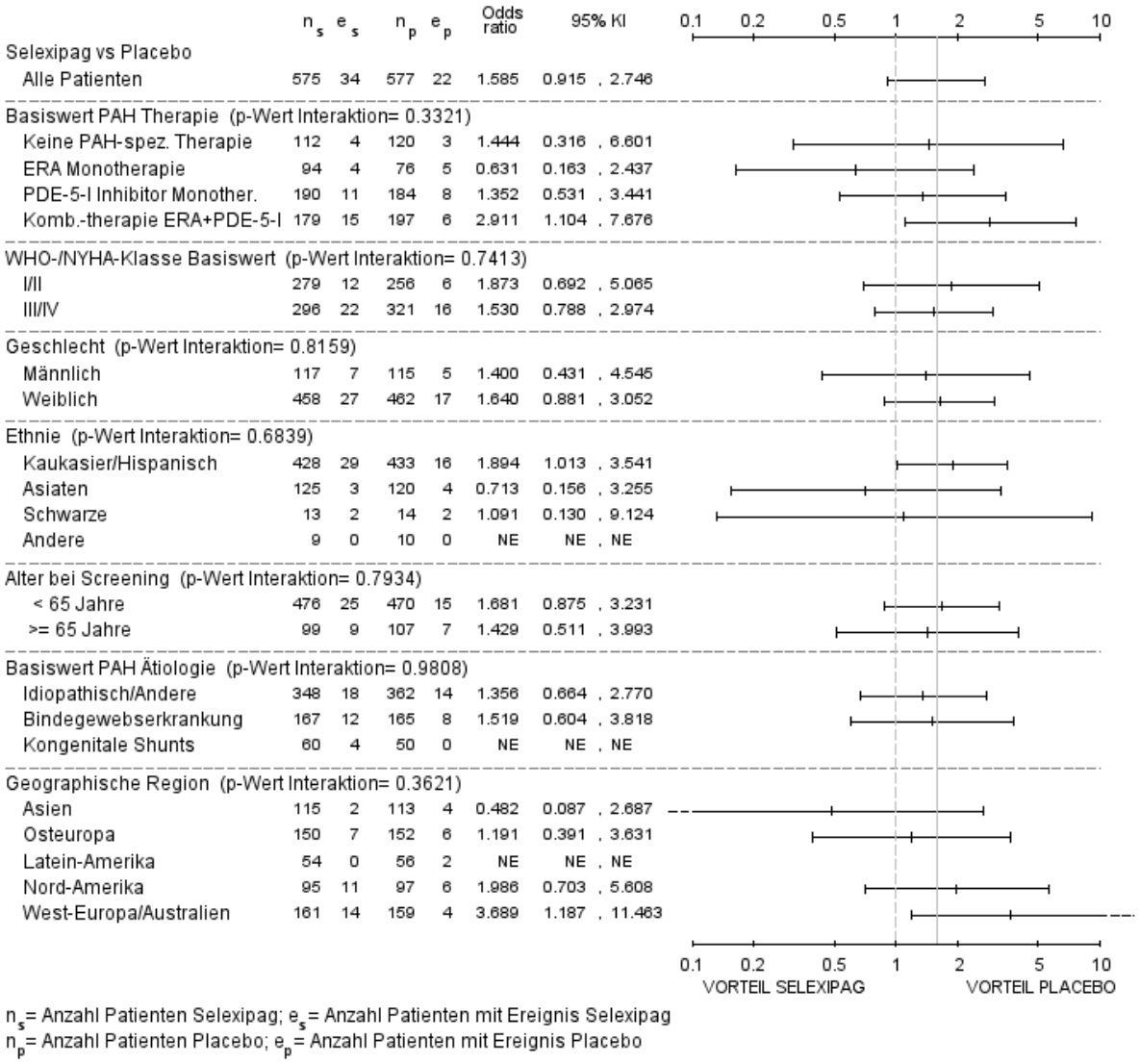
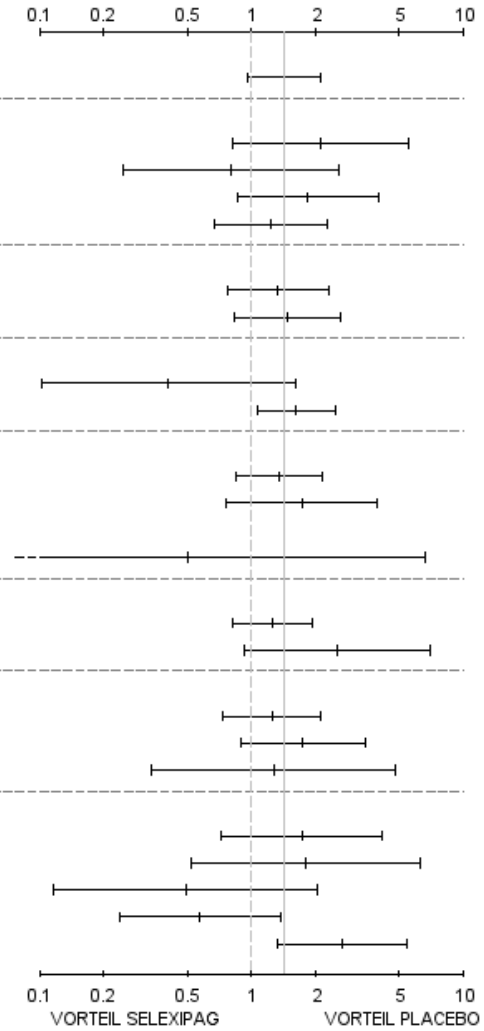


Abbildung 4-37: Unerwünschte Ereignisse von besonderem Interesse: Hypotonie – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests zu Hypotonie-assoziierten UE wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

	n_s	e_s	n_p	e_p	Odds ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	575	63	577	46	1.420	0.953 , 2.117
Basiswert PAH Therapie (p-Wert Interaktion= 0.5192)						
Keine PAH-spez. Therapie	112	13	120	7	2.120	0.814 , 5.523
ERA Monotherapie	94	6	76	6	0.795	0.246 , 2.574
PDE-5-I Inhibitor Monother.	190	20	184	11	1.850	0.861 , 3.978
Komb.-therapie ERA+PDE-5-I	179	24	197	22	1.232	0.664 , 2.284
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion= 0.8146)						
I/II	279	34	256	24	1.341	0.772 , 2.331
III/IV	296	29	321	22	1.476	0.828 , 2.632
Geschlecht (p-Wert Interaktion= 0.0583)						
Männlich	117	3	115	7	0.406	0.102 , 1.610
Weiblich	458	60	462	39	1.635	1.068 , 2.503
Ethnie (p-Wert Interaktion= 0.8280)						
Kaukasier/Hispanisch	428	44	433	34	1.345	0.841 , 2.149
Asiaten	125	17	120	10	1.731	0.759 , 3.951
Schwarze	13	1	14	0	NE	NE , NE
Andere	9	1	10	2	0.500	0.037 , 6.683
Alter bei Screening (p-Wert Interaktion= 0.2112)						
< 65 Jahre	476	50	470	40	1.262	0.815 , 1.953
>= 65 Jahre	99	13	107	6	2.545	0.928 , 6.981
Basiswert PAH Ätiologie (p-Wert Interaktion= 0.7374)						
Idiopathisch/Andere	348	32	362	27	1.256	0.736 , 2.145
Bindegewebserkrankung	167	25	165	15	1.761	0.892 , 3.475
Kongenitale Shunts	60	6	50	4	1.278	0.340 , 4.806
Geographische Region (p-Wert Interaktion= 0.0470)						
Asien	115	15	113	9	1.733	0.726 , 4.141
Osteuropa	150	7	152	4	1.811	0.519 , 6.320
Latein-Amerika	54	3	56	6	0.490	0.116 , 2.068
Nord-Amerika	95	9	97	15	0.572	0.237 , 1.379
West-Europa/Australien	161	29	159	12	2.691	1.320 , 5.488



n_s= Anzahl Patienten Selexipag; e_s= Anzahl Patienten mit Ereignis Selexipag
n_p= Anzahl Patienten Placebo; e_p= Anzahl Patienten mit Ereignis Placebo

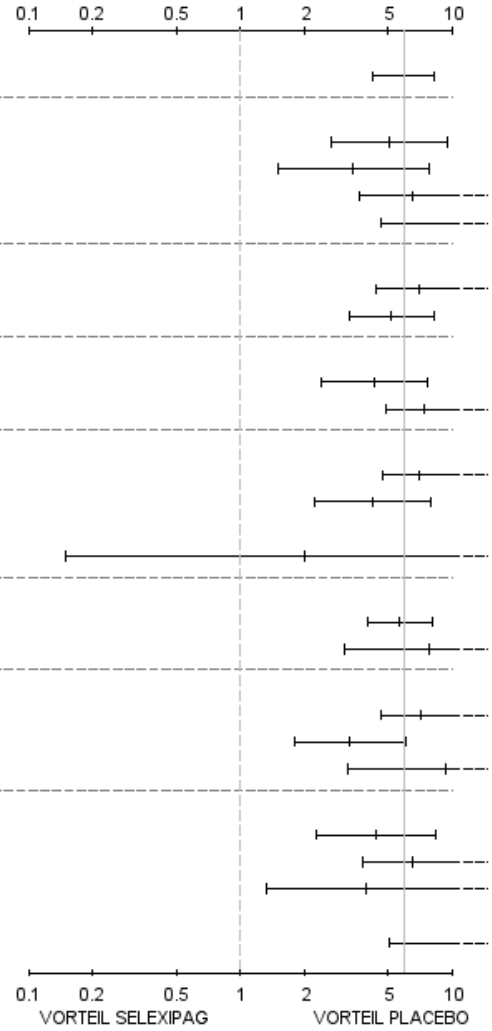
Abbildung 4-38: Unerwünschte Ereignisse von besonderem Interesse: Störungen des Auges – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

Die Auswertung der Interaktionstests für die Subgruppenanalyse der UE, die mit Störungen des Auges assoziiert sind, ergibt lediglich für die Subgruppen nach Region ein positives Testergebnis (p=0,0470).

Da sich der Verdacht eines ursächlichen biologischen Zusammenhangs zwischen der Anwendung von Selexipag und dem Auftreten von Augenstörungen in der Studie insgesamt nicht bestätigen konnte, ist davon auszugehen, dass das entsprechende Subgruppenergebnis als Zufallsbefund und mögliches Resultat des multiplen Testens zu bewerten ist.

4.3.1.3.2.17 Prostacyclin-assoziierte unerwünschte Ereignisse

	n_s	e_s	n_p	e_p	Odds ratio	95% KI
Selexipag vs Placebo						
Alle Patienten	575	523	577	363	5.929	4.257 , 8.258
Basiswert PAH Therapie (p-Wert Interaktion= 0.2748)						
Keine PAH-spez. Therapie	112	95	120	63	5.056	2.697 , 9.477
ERA Monotherapie	94	84	76	54	3.422	1.504 , 7.785
PDE-5-I Inhibitor Monother.	190	173	184	112	6.542	3.665 , 11.679
Komb.-therapie ERA+PDE-5-I	179	171	197	134	10.049	4.655 , 21.694
WHO-/NYHA-Klasse Basiswert (p-Wert Interaktion= 0.3804)						
I/II	279	253	256	149	6.988	4.350 , 11.225
III/IV	296	270	321	214	5.192	3.263 , 8.263
Geschlecht (p-Wert Interaktion= 0.1345)						
Männlich	117	94	115	56	4.306	2.400 , 7.725
Weiblich	458	429	462	307	7.469	4.894 , 11.399
Ethnie (p-Wert Interaktion= 0.4879)						
Kaukasier/Hispanisch	428	394	433	270	6.996	4.686 , 10.443
Asiaten	125	108	120	72	4.235	2.259 , 7.939
Schwarze	13	13	14	13	NE	NE , NE
Andere	9	8	10	8	2.000	0.150 , 26.734
Alter bei Screening (p-Wert Interaktion= 0.5222)						
< 65 Jahre	476	430	470	292	5.698	3.989 , 8.140
>= 65 Jahre	99	93	107	71	7.859	3.139 , 19.677
Basiswert PAH Ätiologie (p-Wert Interaktion= 0.0829)						
Idiopathisch/Andere	348	318	362	216	7.165	4.664 , 11.006
Bindegewebserkrankung	167	150	165	120	3.309	1.802 , 6.074
Kongenitale Shunts	60	55	50	27	9.370	3.210 , 27.350
Geographische Region (p-Wert Interaktion= 0.4003)						
Asien	115	99	113	66	4.406	2.307 , 8.416
Osteuropa	150	125	152	66	6.515	3.812 , 11.134
Latein-Amerika	54	49	56	40	3.920	1.321 , 11.629
Nord-Amerika	95	95	97	83	NE	NE , NE
West-Europa/Australien	161	155	159	108	12.197	5.055 , 29.430



n_s= Anzahl Patienten Selexipag; e_s= Anzahl Patienten mit Ereignis Selexipag
n_p= Anzahl Patienten Placebo; e_p= Anzahl Patienten mit Ereignis Placebo

Abbildung 4-39: Prostacyclin-assoziierte unerwünschte Ereignisse – Forest-Plot der Subgruppenanalysen
Analysierte Population: SAF. Sämtliche in der Abbildung dargestellten Analysen wurden *post hoc* durchgeführt. ERA – Endothelin-Rezeptor-Antagonist, PDE-5-I – Phosphodiesterase-5-Inhibitor, KI – Konfidenzintervall, NE – nicht bestimmbar

In keinem der durchgeführten Interaktionstests wurde ein statistisch signifikantes Ergebnis erreicht. Es wird daher von einem homogenen Effekt über alle untersuchten Subgruppen ausgegangen.

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Zusammenfassung GRIPHON

Die Analyse des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ in der GRIPHON-Studie zeigte eine deutliche und klinisch relevante Verminderung für das Risiko des Auftretens eines M/M-Ereignisses vom Zeitpunkt der Randomisierung bis zum Ende der Behandlungsphase mit einer HR [99 % KI] gegenüber Placebo von 0,60 [0,46; 0,78] (Selexipag im Vergleich zu Placebo). Dies entspricht einer Reduktion des relativen Risikos für das Auftreten eines M/M-Ereignisses von 40 %. Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p<0,0001).

Der Therapieeffekt des Endpunkts war konsistent in allen supportiven Analysen und über alle untersuchten Subgruppen. Insbesondere konnte auch eine vergleichbare Wirksamkeit von Selexipag gegenüber Placebo nachgewiesen werden, die unabhängig von der PAH-spezifischen Begleittherapie der Patienten war, d. h. insbesondere auch bei Patienten, die zum Zeitpunkt der Basiswerterhebung bereits mit einer Kombination aus ERA und PDE-5-I behandelt wurden. Ebenso war die in GRIPHON für den primären Endpunkt beobachtete Überlegenheit von Selexipag gegenüber Placebo unabhängig davon, ob die vom G-BA als weniger schwer bzw. als nicht ausreichend patientenrelevant erachteten Komponenten „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie“ in die Auswertung des kombinierten Endpunkts eingeschlossen wurden.

Das Therapiekonzept von Selexipag sieht vor, dass die Dosierung des Wirkstoffs bis zur patientenindividuell höchsten tolerierten Dosis gesteigert wird, bei der spezifische Nebenwirkungen entweder noch nicht auftreten oder ggf. durch adäquate Gegenmaßnahmen behandelt werden können. Zur Überprüfung dieses Therapiekonzepts wurde in einer supportiven Analyse die Wirksamkeit von Selexipag in Abhängigkeit von der individuellen Erhaltungsdosis ausgewertet. Die Ergebnisse dieser Analyse zeigten einen konsistenten Therapieeffekt über alle Wirkstärkekategorien und unterstreichen damit die Effektivität des gewählten Therapiekonzepts für Selexipag.

Die nach den Vorgaben des G-BA vorgenommene statistische Analyse der naiven Raten des primären Endpunkts ergab ein zur Ereigniszeitanalyse konsistentes Ergebnis, auch wenn dieses aufgrund der Nicht-Berücksichtigung der unterschiedlichen Behandlungsdauern in den beiden Studienarmen als hoch verzerrt angesehen werden muss. Die Analyse der naiven

Raten ergab eine statistisch signifikante Reduktion des relativen Risikos von 35 % (RR [95 % KI]: 0,65 [0,55; 0,77]) bzw. des absoluten Risikos von 14,6 Prozentpunkten (RD [95 % KI]: −14,6 %-P. [−20,0 %-P.; −9,2 %-P.]).

Das in beiden Behandlungsgruppen am häufigsten aufgetretene erste Ereignis war die „Hospitalisierung wegen PAH“ (13,6 % bei Patienten unter Selexipag gegenüber 18,7 % bei Patienten unter Placebo). Ebenfalls häufig wurde „Fortschreiten der Erkrankung“ als erstes Ereignis beobachtet (6,6 % gegenüber 17,2 %). Der Unterschied zwischen den Behandlungsgruppen war zudem für diese beiden Komponenten statistisch signifikant. Das Ereignis „Tod“ trat nur bei einer vergleichsweise geringen Anzahl von Patienten als erstes Ereignis auf (4,9 % gegenüber 3,1 %). Die vergleichsweise geringe Zahl ist nicht überraschend, da sich Patienten in der PAH üblicherweise erst klinisch verschlechtern (im Studienkontext also eines der anderen M/M-Ereignisse erleben), bevor sie versterben. Die Ereignisse „Lungentransplantation oder atriale Septostomie“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie“ traten jeweils nur sehr selten auf und lassen allein stehend nur eine sehr limitierte Aussage zum Behandlungseffekt zu.

In der Analyse der „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten Hospitalisierung wegen einer Verschlechterung der PAH“ konnte ebenfalls eine deutliche und klinisch relevante Reduktion des relativen Risikos für das Auftreten eines entsprechenden Ereignisses beobachtet werden. Die statistische Analyse der Patienten im Selexipag-Arm im Vergleich zum Placebo-Arm ergab eine HR [99 % KI] von 0,70 [0,50; 0,98]; dies entspricht einer Reduktion des relativen Risikos von 30 %. Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p<0,0031).

Der Therapieeffekt des Endpunkts war konsistent in allen supportiven Analysen und über alle untersuchten Subgruppen. Ebenso waren auch die Ergebnisse der Analyse der naiven Raten konsistent zu den Ergebnissen der Ereigniszeitanalyse. In der (potentiell hoch verzerrten) Auswertung der naiven Raten ergab sich eine signifikante Reduktion des relativen Risikos von 24 % (RR [95 % KI]: 0,76 [0,60; 0,95]) bzw. eine signifikante Reduktion des absoluten Risikos um 5,8 Prozentpunkte (RD [95 % KI]: −5,8 %-P. [−10,4 %-P.; −1,1 %-P.]).

Das in beiden Gruppen am häufigsten aufgetretene erste Ereignis dieses kombinierten Endpunkts war die „Hospitalisierung wegen PAH“, die bei 15,0 % der Patienten im Selexipag-Arm und bei 21,1 % der Patienten im Placebo-Arm aufgetreten war. Das Ereignis „Tod“ trat im Rahmen dieses kombinierten Endpunkts in beiden Behandlungsarmen nur bei einer geringen Anzahl von Patienten als erstes Ereignis auf (Selexipag: 2,8 %, Placebo: 2,4 %).

Das Risiko für das Auftreten eines Todesfalles jeglicher Ursache vom Zeitpunkt der Randomisierung bis Studienschluss war in GRIPHON in beiden Behandlungsgruppen vergleichbar. Die HR [99 % KI] in der Ereigniszeitanalyse der Todesfälle bis Studienschluss betrug 0,97 [0,68; 1,39] und war statistisch nicht signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p=0,4214). Die Auswertung der CEC-bestätigten Todesfälle wegen PAH bis Studienschluss ergab wiederum eine leichte numerische Reduktion des relativen Risikos

durch Selexipag von 14 % (HR [99 % KI]: 0,86 [0,63; 1,18]). Der Unterschied zwischen den Behandlungsgruppen war jedoch ebenfalls statistisch nicht signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p=0,1763). Zu beachten in der Interpretation beider Endpunkte ist das hohe Verzerrungspotential, das durch den erlaubten Therapiewechsel am Ende der verblindeten Behandlungsphase entsteht, nach der Patienten in der Placebo-Gruppe in die Selexipag-Gruppe wechseln oder mit weiteren PAH-spezifischen Arzneimitteln behandelt werden konnten. RPSFT- und MSM-Analysen, die *post hoc* durchgeführt wurden, um den Effekt durch den erlaubten Therapiewechsel zu korrigieren, bestätigen die Vermutung, dass die Ergebnisse der Auswertung der Todesfälle bis zum Studienschluss zu Ungunsten von Selexipag verzerrt sind.

Die Analyse der „Zeit bis zum CEC-bestätigten Tod wegen PAH“ bis zum Ende der Behandlungsphase (EOT)“ ergab eine leichte numerische Erhöhung des relativen Risikos von 16 % (HR [99 % KI]: 1,16 [0,70; 1,93]) durch Selexipag. Der Unterschied zwischen den Patientengruppen war jedoch auch hier statistisch nicht signifikant (einseitiger nicht-stratifizierter Log-Rang-Test: p=0,7153). Zu beachten ist hierbei, dass der Therapieeffekt aufgrund der informativen Zensierung der Todesfälle zu EOT voraussichtlich hoch verzerrt ist.

Mit Ausnahme der Subgruppen nach WHO-/NYHA-Klasse für den Endpunkt „Zeit bis zum Tod (Studienschluss)“ war der Therapieeffekt für alle drei Mortalitätsendpunkte konsistent in allen supportiven Analysen und über alle untersuchten Subgruppen. Da keine belastbare medizinische Rationale, kein konsistenter Effekt über mehrere Endpunkte und auch keine entsprechenden Ergebnisse aus früheren Studien vorlagen, wird jedoch davon ausgegangen, dass es sich beim beobachteten positiven Interaktionstest für die „Zeit bis zum Tod (Studienschluss)“ um einen zufälligen Effekt und nicht um eine tatsächliche Effekt-modifikation handelt.

Ebenso ergab auch die potentiell hoch verzerrte Analyse der naiven Raten keine von der Hauptanalyse abweichenden Ergebnisse.

In der Analyse der „Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH“ zeigte sich eine deutliche und klinisch relevante Reduktion des relativen Risikos einer Hospitalisierung wegen PAH durch Selexipag (HR [95 % KI]: 0,65 [0,50; 0,86]). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (p=0,0022).

Der Therapieeffekt war konsistent über alle untersuchten Subgruppen. Ebenso ergab auch die potentiell hoch verzerrte Analyse der naiven Raten keine von der Hauptanalyse abweichenden Ergebnisse.

Während der Behandlungsphase von GRIPHON wurde bei 43,7 % der Patienten im Selexipag-Arm und bei 51,5 % der Patienten im Placebo-Arm mindestens ein „Symptom der PAH“ beobachtet. Dies entspricht einer Reduktion des relativen Risikos von 15 % (RR [95 % KI]: 0,85 [0,75; 0,96]) bzw. einer Reduktion des absoluten Risikos von 7,8 Prozentpunkten

(RD [95 % KI]: −7,8 %-P. [−13,6 %-P.; −2,1 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (Barnards exakter Test: p=0,0081).

Der Therapieeffekt war konsistent über alle untersuchten Subgruppen.

Die mediane absolute Veränderung der 6-MWD vom Basiswert bis Woche 26 betrug 4,0 m bei Patienten in der Selexipag-Gruppe, während sich die Patienten in der Placebo-Gruppe im gleichen Zeitraum im Median um 9,0 m verschlechterten. Dies entspricht einer medianen Verbesserung [99 % KI] durch Selexipag von 12,0 m [1 m; 24 m]; der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant (Wilcoxon-Mann-Whitney-Test aus nicht-parametrischer ANCOVA mit 6-MWD-Basiswert als Kovariate: p=0,0027).

Der im Vergleich zu vergangenen Studien in der PAH relativ niedrige 6-MWD-Wert in GRIPHON lässt sich hierbei möglicherweise durch die große Anzahl an Patienten erklären, die zum Beginn der Studie bereits mit einem oder sogar zwei PAH-spezifischen Arzneimitteln behandelt worden waren. Dies könnte jedoch zu einer Limitierung der Verbesserung des einzelnen Patienten führen (*Ceiling*-Effekt), wie sie in der Vergangenheit bei Patienten mit bestehender PAH-spezifischer Therapie auch schon beobachtet wurde, und damit auch die Unterschiede zwischen den Behandlungsgruppen nivellieren. Konsistent hierzu war auch in GRIPHON der Behandlungseffekt in der Subgruppe der Patienten ohne PAH-spezifische Begleittherapie sehr viel höher als in der Gesamtstudienpopulation (Behandlungseffekt [95 % KI]: 34,0 m [15,0 m, 55,0 m]).

Mit Ausnahme von Monat 36, zu dem jedoch nur die Werte von vergleichsweise wenigen Patienten in die Analyse eingingen, war die mediane absolute Veränderung der 6-MWD zudem zu allen untersuchten Zeitpunkten im Selexipag-Arm höher als im Placebo-Arm.

Der Therapieeffekt des Endpunkts war in allen durchgeführten supportiven Analysen konsistent. Mit Ausnahme der Analyse der absoluten Veränderung der 6-MWD nach PAH-spezifischer Therapie, die oben stehend bereits berichtet wurde, war das Ergebnis zudem über alle Subgruppen konsistent.

Eine ausbleibende Verschlechterung der WHO-/NYHA-Klasse in Woche 26 im Vergleich zum Basiswert wurde in GRIPHON bei 77,8 % der Patienten im Selexipag-Arm und bei 74,9 % der Patienten im Placebo-Arm beobachtet. Dies entspricht einer numerischen Verbesserung des absoluten Risikos um 2,8 Prozentpunkte (RD [95 % KI]: 2,8 [−2,1; 7,8]) bzw. einer OR [99 % KI] von 1,16 [0,81; 1,66]. Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (zweiseitiger Cochran-Mantel-Haenszel-Test, p=0,2843).

Ähnlich wie bei der absoluten Veränderung der 6-MWD ist der fehlende Unterschied zwischen den Behandlungsgruppen möglicherweise durch die Tatsache erklärbar, dass der Großteil der Patienten zum Zeitpunkt der Basiswerterhebung bereits mit einem oder zwei PAH-spezifischen Arzneimitteln vorbehandelt wurde und daher mit hoher Wahrscheinlichkeit zu diesem Zeitpunkt verhältnismäßig klinisch stabil war.

Eine Auswertung der ausbleibenden Verschlechterung der WHO-/NYHA-Klasse zu weiteren Studienzeitpunkten zeigt zudem, dass der Anteil der Patienten ohne Verschlechterung der WHO-/NYHA-Klasse zu allen Zeitpunkten im Selexipag-Arm höher war als im Placebo-Arm. Eine Ausnahme hiervon bilden lediglich die Messungen zu Woche 26 und Monat 24, in denen der Anteil der Patienten ohne Verschlechterung in beiden Gruppen ausgeglichen war.

Der Therapieeffekt war konsistent über alle untersuchten Subgruppen.

In Bezug auf die „Absolute Veränderung des Borg-Dyspnoe-Index“ ergab die Analyse der Veränderung in Woche 26 im Vergleich zum Basiswert ein ausgeglichenes Ergebnis zwischen den Behandlungsgruppen. Der Behandlungseffekt [95 % KI] von Selexipag im Vergleich zu Placebo betrug 0,0 [0,0; 0,0], als p-Wert für den Unterschied zwischen den Behandlungsgruppen ergab sich ein Wert von p=0,0061 (Wilcoxon-Mann-Whitney-Test aus nicht-parametrischer ANCOVA mit Borg-Dyspnoe-Index-Basiswert als Kovariate).

Im Rahmen der Subgruppenanalysen ergab sich ein positiver Interaktionstest für die Subgruppenanalyse nach PAH-Ätiologie, die jedoch nicht zu einem qualitativ unterschiedlichen Ergebnis im Vergleich zur Hauptanalyse führt. Da weiterhin keine belastbare medizinische Rationale, kein konsistenter Effekt über mehrere Endpunkte und keine entsprechenden Ergebnisse aus früheren Studien vorliegen, wird davon ausgegangen, dass es sich um einen zufälligen Effekt und nicht um eine tatsächliche Effektmodifikation handelt.

In Bezug auf die gesundheitsbezogene Lebensqualität zeigte sich weder in der Auswertung der „Absoluten Veränderung des CAMPHOR ‚Symptome‘-Scores“ noch in der Auswertung der „Absoluten Veränderung der ‚Kurzatmigkeit‘-Subskala des CAMPHOR ‚Symptome‘-Scores“ ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (p=0,2185 bzw. p=0,1700). Der Behandlungseffekt [99 % KI] für Selexipag im Vergleich zu Placebo betrug hierbei 0,0 [−1,0; 1,0] für die absolute Veränderung des CAMPHOR „Symptome“-Scores und 0,0 [−0,4; 0,0] für die „Kurzatmigkeit“-Subskala.

Die Tatsache, dass die Validierung des CAMPHOR-Fragebogens in einer gemischten PH-Population durchgeführt wurde, und nicht – wie in GRIPHON – in einer reinen PAH-Gruppe, ist ein möglicher Grund, der den nicht nachweisbaren Effekt von Selexipag auf die Lebensqualität der Patienten in GRIPHON erklären könnte. Zudem stand der Fragebogen überwiegend nur in solchen Ländern/Sprachen zur Verfügung, in denen die Wahrscheinlichkeit sehr hoch war, dass Patienten zum Zeitpunkt des Studieneinschlusses bereits mit zwei PAH-spezifischen Arzneimitteln behandelt wurden und damit ihre Lebensqualität nicht im gleichen Maße steigen konnte wie bei unbehandelten und daher tendenziell symptomatischeren Patienten.

Der Therapieeffekt war konsistent über alle untersuchten Subgruppen.

Selexipag wurde in der GRIPHON-Studie insgesamt gut vertragen, wenn auch (wie aufgrund des Titrationsschemas zu erwarten war) vermehrt Prostacyclin-assoziierte UE unter Selexipag auftraten. Die Gesamtrate der unerwünschten Ereignisse war in beiden Studienarmen mit

98,3 % unter Selexipag und 96,9 % unter Placebo vergleichbar hoch. Dies entspricht einer RR [95 % KI] im Vergleich von Selexipag zu Placebo von 1,01 [1,00; 1,03] und einer absoluten RD [95 % KI] von 1,4 Prozentpunkten [−0,4 %-P.; 3,2 %-P.]. Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,1329).

Wie aufgrund des Studiendesigns von GRIPHON zu erwarten war, das in den ersten 12 Wochen die Titration des Wirkstoffs bis zur patientenindividuell höchsten noch tolerierten Dosis vorsah, war der Anteil der Patienten, die unter Selexipag ein UE hatten, ausschließlich in der Titrationsphase höher als in der Placebo-Gruppe (93,3 % vs. 76,8 %). In der zeitlich deutlich längeren Erhaltungsphase hingegen waren die Raten in beiden Gruppen ausgeglichen (93,5 % vs. 93,1 %).

Die Auswertung der UE auf Ebene der bevorzugten Bezeichnung ergab, dass UE, die in der Selexipag-Gruppe häufiger aufgetreten waren als im Placebo-Arm, solche UE sind, die üblicherweise mit dem Einsatz von Arzneimitteln assoziiert werden, die den Prostacyclin-Signalweg adressieren. Diese UE sind i. d. R. gut behandelbar und von vergleichsweise geringer klinischer Relevanz. Im Gegensatz dazu handelte es sich bei UE, die vermehrt bei Patienten im Placebo-Arm aufgetreten waren, um solche UE, die entweder der PAH direkt oder den vom G-BA bestimmten Symptomen der PAH zuzurechnen waren.

UE von schwerer Intensität traten in GRIPHON bei 49,0 % der Patienten im Selexipag-Arm auf und damit vergleichsweise häufiger als bei Patienten im Placebo-Arm (44,9 %). Dies entspricht einer numerischen Erhöhung des relativen Risikos von 9 % (RR [95 % KI]: 1,09 [0,96; 1,23]) bzw. des absoluten Risikos von 4,0 Prozentpunkten (RD [95 % KI]: 4,0 %-P. [−1,8 %-P.; 9,7 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war jedoch statistisch nicht signifikant (Barnards exakter Test: p=0,2124).

Ebenso wie auch die Gesamtrate der UE war auch die Rate der schweren UE lediglich in der Titrationsphase unter Selexipag höher als unter Placebo (21,8 % vs. 10,4 %), während in der längeren Erhaltungsphase ein nahezu ausgeglichenes Verhältnis vorlag (42,4 % vs. 40,9 %). Dies zeigt, dass der Unterschied zwischen den Behandlungsgruppen vorwiegend auf Prostacyclin-assoziierte UE zurückgeht und damit indirekt auch ein Nebeneffekt der Titrationsstrategie von Selexipag ist.

Schwerwiegende unerwünschte Ereignisse traten im Verlauf der Behandlungsphase von GRIPHON bei 43,8 % der Patienten im Selexipag-Arm und bei 47,1 % der Patienten im Placebo-Arm auf. Dies entspricht einer numerischen Verringerung des relativen Risikos von 7 % (RR [95 % KI]: 0,93 [0,82; 1,06]) bzw. einer numerischen Verringerung des absoluten Risikos von 3,3 Prozentpunkten (RD [95 % KI]: −3,3 %-P. [−9,1 %-P.; 2,4 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war jedoch statistisch nicht signifikant (Barnards exakter Test, p=0,2902).

Im Gegensatz zur Auswertung der UE veränderte sich das Verhältnis der SUE zwischen den Gruppen in Titrations- und Erhaltungsphase nicht wesentlich. Zwar waren die entsprechenden Ereignisraten in der Titratitonsphase nahezu ausgeglichen (Selexipag: 7,1 %, Placebo: 7,9 %),

genauso wie in der Gesamtphase traten in beiden Phasen aber in der Selexipag-Gruppe numerisch weniger SUE auf als in der Placebo-Gruppe.

Im Verlauf der Behandlungsphase von GRIPHON wurden insgesamt deutlich mehr Therapieabrüche aufgrund eines UE bei Patienten im Placebo-Arm als bei Patienten im Selexipag-Arm beobachtet. Der Unterschied zwischen den Behandlungsgruppen entspricht einer numerischen Verringerung des relativen Risikos von 15 % (RR [95 % KI]: 0,85 [0,73; 1,00]) bzw. einer numerischen Verringerung des absoluten Risikos von 5,4 Prozentpunkten (RD [95 % KI]: −5,4 %-P. [−10,9 %-P.; 0,0 %-P.]). Der Unterschied zwischen den Behandlungsgruppen war statistisch nicht signifikant (Barnards exakter Test: p=0,0554), jedoch stellt die Obergrenze des Konfidenzintervalls beider Effektschätzer gleichzeitig auch den Nulleffekt-Wert dar.

Bei der Aufteilung der Therapieabbrüche wegen unerwünschter Ereignisse nach der Behandlungsphase, in der sie auftraten, zeigt sich ein differenziertes Bild. Während in der Titrationsphase der Anteil der Therapieabbrüche in der Selexipag-Gruppe numerisch höher ist als in der Placebo-Gruppe (3,9 % vs. 2,4 %), ist das Verhältnis in der Erhaltungsphase umgekehrt (26,7 % vs. 30,9 %). In beiden Phasen war der Unterschied zwischen den Behandlungsphasen statistisch nicht signifikant.

Zusätzlich zur Hauptanalyse wurden für alle nicht-spezifischen Nebenwirkungsendpunkte Sensitivitätsanalysen durchgeführt, die helfen sollten, die mögliche Doppelbewertung von Morbiditätsereignissen und unerwünschten Ereignissen einzuschätzen. Hierzu wurden in der Sensitivitätsanalyse sämtliche (S)UE aus der Analyse ausgeschlossen, die am gleichen Tag oder später begannen wie das M/M-Ereignis des primären Endpunkts. Die Ergebnisse dieser Sensitivitätsanalysen zeigen für alle untersuchten Endpunkte eine tendenzielle Verschiebung der Effektschätzer zuungunsten von Selexipag, führen aber nicht zu einer grundlegend verschiedenen Interpretation des möglichen Schadenpotentials von Selexipag. Es ist auch nach Ausschluss der Ereignisse des primären Endpunkts nicht von einem erhöhten Schaden durch Selexipag auszugehen.

Mit Ausnahme der Therapieabbrüche wegen UE, bei denen positive Interaktionstests für die Subgruppen nach Therapie und nach Ethnie/Rasse vorlagen, war der Therapieeffekt für die nicht-spezifischen Nebenwirkungsendpunkte über alle untersuchten Subgruppen konsistent. Auch für die Analyse der Therapieabbrüche wegen UE wird letztlich von einem konsistenten Therapieeffekt ausgegangen, da für die beobachteten Unterschiede keine belastbare medizinische Rationale, kein konsistenter Effekt über mehrere Endpunkte und keine entsprechenden Ergebnisse aus früheren Studien vorlagen. Die beobachteten positiven Interaktionstests werden daher nicht als tatsächliche Effektmodifikation, sondern als Zufallseffekt mutmaßlich infolge des multiplen Testens gewertet.

Die Untersuchung der UE von besonderem Interesse ergab in keiner der Gruppierungen (Hämorrhagie, Hypotonie, Störung des Auges) ein statistisch signifikant höheres oder niedrigeres Risiko für das Auftreten eines entsprechenden UE.

Für Hämorrhagie und Störungen des Auges wurde ein positiver Interaktionstest in den Subgruppen nach WHO-/NYHA-Klasse bzw. nach Region beobachtet. Da keine belastbare medizinische Rationale, kein konsistenter Effekt über mehrere Endpunkte und keine entsprechenden Ergebnisse aus früheren Studien vorlagen, wird davon ausgegangen, dass es sich um einen Zufallsbefund und nicht um eine tatsächliche Effektmodifikation handelt.

Wie aufgrund der für GRIPHON gewählten Titrationsstrategie zu erwarten war, in der das Auftreten von Prostacyclin-assoziierten unerwünschten Ereignissen Voraussetzung für die Bestimmung der patientenindividuell höchsten tolerierten Dosis war, traten über den gesamten Behandlungszeitraum betrachtet im Selexipag-Arm mehr Prostacyclin-assoziierte UE auf als im Placebo-Arm (91,0 % vs. 62,9 %). Der Unterschied zwischen den Behandlungsgruppen war statistisch signifikant und entspricht einer Erhöhung des relativen Risikos um 45 % (RR [95 % KI]: 1,45 [1,35; 1,55]) bzw. einer Erhöhung des absoluten Risikos um 28 % Prozentpunkte (RD [95 % KI]: 28 %-P. [23,5 %-P.; 32,6 %-P.]).

Konsistent zur Annahme, dass die Unterschiede in Bezug auf die Prostacyclin-assoziierten UE zu einem relevanten Teil auf die Titrationsstrategie zurückzuführen sind, war der Unterschied zwischen den Behandlungsgruppen in der Titrationsphase (RD [95 % KI]: 34,9 %-P. [29,6 %-P.; 40,1 %-P.]) deutlich größer als in der wesentlich längeren Erhaltungsphase (25,3 %-P. [19,4 %-P.; 31,1 %-P.]). Dennoch war der Unterschied zwischen den Behandlungsgruppen auch in der Erhaltungsphase statistisch signifikant (Barnards exakter Test: $p<0,0001$).

Bei den am häufigsten aufgetretenen Prostacyclin-assoziierten UE handelt es sich durchweg um Störungen mit vergleichsweise geringerer klinischer Relevanz, die mit üblichen Therapiemethoden gut zu behandeln sind. Am weitaus häufigsten wurde in beiden Behandlungsgruppen und Behandlungsphasen das UE Kopfschmerz beobachtet (64,4 % bzw. 39,9 % (Selexipag) vs. 28,3 % bzw. 19,5 % (Placebo)). Ebenfalls häufig traten auch die UE Diarrhö, Übelkeit und Kieferschmerz auf. Nur wenige der in der gesamten Behandlungsphase beobachteten Prostacyclin-assoziierten UE wurden gleichzeitig auch als SUE eingeschätzt. Keines dieser SUE führte zum Versterben des Patienten.

Der Therapieeffekt war konsistent über alle untersuchten Subgruppen.

Zusammenfassend kann man sagen, dass Selexipag durch die Ergebnisse in GRIPHON das einzige den Prostacyclin-Signalweg adressierende Arzneimittel in der PAH ist, durch dessen Anwendung eine deutliche und klinisch relevante Reduktion patientenrelevanter Langzeitparameter der Erkrankung, d. h. insbesondere des primären Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ erreicht werden konnte. Begleitet wird diese Verbesserung durch eine deutliche Reduktion des Risikos für Hospitalisierungen aufgrund einer Verschlechterung der PAH, einer Reduktion des Risikos für das Auftreten von Symptomen der PAH (nach Defintion des G-BA) sowie einer signifikanten Verbesserung der 6-MWD.

Gleichzeitig erscheint Selexipag als gut verträglich. Die Behandlung mit diesem Arzneimittel führt zu einer im Vergleich zu Placebo ausgeglichenen Gesamtrate der unerwünschten Ereignisse, das Risiko für das Auftreten schwerwiegender unerwünschter Ereignisse und für Therapieabbrüche aufgrund unerwünschter Ereignisse wurde tendenziell verringert. Wie aufgrund der Titrationsstrategie und des Wirkmechanismus zu erwarten war, traten unter Selexipag insbesondere im Verlauf der Titrationsphase vermehrt Prostacyclin-assoziierte unerwünschte Ereignisse auf. Hierbei handelt es sich vorwiegend aber um Störungen mit vergleichsweise geringer klinischer Relevanz, die mit üblichen Therapiemethoden gut zu behandeln sind.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Nicht zutreffend.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-118: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	**<Mortalität>**	**<Gesundheits-bezogene Lebensqualität>**	**<Endpunkt>**	**<Endpunkt>**	**<Endpunkt>**
<Studie 1>	nein	ja	ja	ja	nein

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-119: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	**Studie**	**Intervention**	**<Vergleichs-therapie 1>**	**<Vergleichs-therapie 2>**	**<Vergleichs-therapie 3>**
1	<Studie 1>	●		●	●
2	<Studie 2>	●		●	
	<Studie 3>	●		●	
1	<Studie 4>		●	●	●
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-120: Operationalisierung von <Endpunkt xxx>

Studie	**Operationalisierung**
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-121: Bewertung des Verzerrungspotentials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	**Verzerrungspotential auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotential Endpunkt**
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-122: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	**Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)**
<Studie 1>	

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-123: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-124: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-125: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. ***Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-126: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Teilpopulation a

Die im Dossier dargelegten Nachweise zum Zusatznutzen von Selexipag basieren auf der Auswertung der pivotalen Zulassungsstudie GRIPHON. Hierbei handelte es sich um eine multizentrische, randomisierte, doppelblinde, placebokontrollierte und ereignisgetriebene Phase-3-Studie im Parallelgruppendesign zum Vergleich der Wirkung (Wirksamkeit, Sicherheit, Verträglichkeit, Pharmakokinetik, Pharmakodynamik) von Selexipag (zweimal täglich verabreicht in individuellen Dosierungen von 200 bis 1.600 µg nach Auftitration) gegenüber Placebo bei Patienten mit symptomatischer PAH.

Ebenso wie die Patienten im Selexipag-Arm konnten auch die Patienten im Placebo-Arm ihre vor dem Studienbeginn begonnene PAH-spezifische Begleittherapie aus ERA und/oder PDE-5-I weiterführen, solange die Dosierung dieser Arzneimittel in den letzten 3 Monaten stabil war. Patienten, die bereits vor Studienbeginn mit Prostanoiden (z. B. Iloprost zur Inhalation, parenterale Anwendung von Treprostinil oder Epoprostenol) behandelt worden waren, konnten an der Studie nicht teilnehmen. Nach Studienbeginn hingegen war die Initiation einer solchen Therapie bei Vorliegen einer Verschlechterung der Erkrankung jederzeit erlaubt, wurde aber direkt (Initiierung einer i. v. oder s. c. Prostanoidtherapie) oder indirekt (in Kombination mit einer Verschlechterung der WHO-/NYHA-Klasse bei Patienten in WHO-/NYHA-Klasse III) als Ereignis des primären Endpunkts gewertet und beendete damit die Behandlungsphase für den jeweiligen Patienten. Insofern ist das Vorgehen im Placebo-Arm von GRIPHON einem abwartenden Vorgehen bis zur Verschlechterung der Erkrankung gleichzusetzen.

Da es sich um eine randomisierte, kontrollierte Studie (RCT) handelt, entspricht GRIPHON dem höchsten Evidenzgrad der Evidenzklassifizierung des IQWiG [15]. Darüber hinaus weist GRIPHON sowohl endpunktübergreifend als auch in den meisten patientenrelevanten Endpunkten ein niedriges Verzerrungspotential und somit eine hohe Aussagesicherheit auf.

Die Untersuchung des Zusatznutzens wurde anhand der folgenden Endpunkte durchgeführt:

- Mortalität/Morbidität (Kombinationsendpunkte):
    - Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis (zu den Einzelkomponenten des kombinierten Endpunktes vgl. auch Abschnitt 4.2.5 bzw. Tabelle 4-17)
    - Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH

- Mortalität:
  - Zeit bis zum Tod (Studienschluss)
  - Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)
  - Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)
- Morbidität:
  - Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH
  - Symptome der PAH[40]
  - Absolute Veränderung der 6-MWD
  - Ausbleibende Verschlechterung der WHO-/NYHA-Klasse
  - Absolute Veränderung des Borg-Dyspnoe-Index
- Lebensqualität:
  - Absolute Veränderung des CAMPHOR „Symptome“-Scores
  - Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores
- Nebenwirkungen:
  - Unerwünschte Ereignisse
  - Schwerwiegende unerwünschte Ereignisse
  - Therapieabbrüche wegen unerwünschter Ereignisse
  - Unerwünschte Ereignisse von besonderem Interesse
    - Hämorrhagie
    - Hypotonie
    - Störung des Auges
  - Prostacyclin-assoziierte unerwünschte Ereignisse

Bei den genannten Endpunkten handelt es sich ausnahmslos um patientenrelevante Endpunkte. Der ebenfalls im Rahmen von GRIPHON erhobene Endpunkt „Hospitalisierungen jeglicher Ursache“ wird nicht zur Bestimmung des Ausmaßes des Zusatznutzens herangezogen, da vom G-BA ausschließlich die PAH-bedingten Hospitalisierungen als patientenrelevant angesehen werden.

[40] Der Endpunkt wurde *post hoc* auf Grundlage des Beratungsgespräches zu Selexipag [3] eingeführt. Als Symptome der PAH wurden die bevorzugten Bezeichnungen (*preferred terms*, PT) der unerwünschten Ereignisse ausgewertet, die den folgenden PAH-Symptomen zuzuordnen sind: Dyspnoe, Synkope, Ödeme, Schwindel, Müdigkeit.

Beim primären Endpunkt „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ handelt es sich um einen kombinierten Endpunkt, der sich über die beiden Endpunktkategorien Mortalität und Morbidität erstreckt. Da sich die Erkrankung bei Patienten mit PAH üblicherweise zunächst verschlechtert, bevor diese versterben, ist der primäre Endpunkt überwiegend durch die eingeschlossenen Morbiditätsereignisse getrieben (vgl. Abschnitt 4.3.1.3.1.1). Er wird daher innerhalb dieser Nutzenbewertung der Endpunktkategorie Morbidität zugerechnet.

Ähnlich wie beim primären Endpunkt handelt es sich auch beim Endpunkt „Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH“ um einen kombinierten Endpunkt, der sich ebenfalls über die beiden Endpunktkategorien Mortalität und Morbidität erstreckt. Auch hier stellen den überwiegenden Anteil der Ereignisse dieses kombinierten Endpunktes Hospitalisierungen wegen PAH, also Morbiditätsereignisse dar. Daher wird auch dieser Endpunkt innerhalb der Nutzenbewertung der Endpunktkategorie Morbidität zugerechnet.

Aufgrund der Evidenzstufe der Studie, dem niedrigen Verzerrungspotential auf Studienebene und in den meisten der patientenrelevanten Endpunkte besitzt die Studie GRIPHON somit eine hohe Aussagekraft.

Insgesamt wird mit GRIPHON lediglich eine Studie zum Nachweis des Zusatznutzens in Teilpopulation a vorgelegt. Aufgrund der überdurchschnittlichen Größe und der hohen Aussagekraft der Studie ist dennoch von einem „Hinweis“ für einen Zusatznutzen auszugehen.

Teilpopulation b

Durch die Informationsbeschaffung, die im Rahmen dieses Dossiers durchgeführt wurde, konnte keine Studie identifiziert werden, in der Selexipag mit einer patientenindividuellen Anwendung von Iloprost zur Inhalation verglichen wurde (vgl. Abschnitt 4.2.3). Insofern ist eine Bewertung des Ausmaßes des Zusatznutzens von Selexipag in Teilpopulation b anhand von RCT nicht möglich. Ein indirekter Vergleich von Selexipag mit Iloprost zur Inhalation ist ebenfalls nicht möglich, da sämtliche bisher mit diesem Wirkstoff in dieser Darreichungsform durchgeführten Studien lediglich eine Dauer von 12 Wochen hatten (für eine Auflistung der Studien siehe Galiè et al. [18]) und somit zu kurz für die Bewertung des Zusatznutzens einer Therapie in einer chronisch progredienten Erkrankung wie der PAH sind. Die Iloprost-Studien weichen in ihrem Studiendesign zudem erheblich von GRIPHON, der pivotalen Zulassungsstudie von Selexipag, ab, sodass ein indirekter Vergleich der Wirkstoffe auch aus methodischen Gründen nicht durchführbar ist.

Wenn auch eine Quantifizierung des Zusatznutzens aufgrund des Fehlens RCT-basierter klinischer Evidenz nicht möglich ist, so lassen sich dennoch verschiedene qualitative Vorteile von Selexipag gegenüber der zweckmäßigen Vergleichstherapie feststellen, die in der Folge einen Zusatznutzen begründen. Da es sich bei den vorgelegten Daten aber nicht um RCT oder indirekte Vergleiche auf Grundlage von RCT handelt, sind sie naturgemäß mit einer sehr

hohen Unsicherheit belastet. Es kann daher höchstens ein „Anhaltspunkt“ für einen Zusatznutzen aus den vorgelegten Daten abgeleitet werden.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das Risiko einer Behandlung mit Iloprost deren Nutzen noch überwiegt (Teilpopulation a)

Die folgende Tabelle 4-127 zeigt in einer Übersicht das Ausmaß des Zusatznutzens in den einzelnen Endpunktkategorien und Endpunkten sowie die zugrunde liegenden Ergebnisse gegenüber der zVT „patientenindividuell optimierte Therapie“ bei Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt.

Für diese Patientengruppe steht mit Ausnahme der Behandlung mit Selexipag keine zugelassene medikamentöse Therapie zur Verfügung, sodass das abwartende Vorgehen bis zur weiteren Verschlechterung der Erkrankung die einzige Therapieoption des behandelnden Arztes für diese Teilpopulation des Anwendungsgebietes von Selexipag ist. Der Placebo-Arm in GRIPHON entspricht durch das gewählte Studiendesign dem abwartenden Vorgehen, die Ergebnisse der Studie sind damit geeignet, eine Aussage zum Ausmaß des Zusatznutzens von Selexipag in dieser Patientenpopulation zu treffen.

Sofern auf den jeweiligen Endpunkt anwendbar, wurde das Ausmaß des Zusatznutzens anhand der Ereigniszeitanalyse, d. h. der Hazard Ratio samt zugehörigen Konfidenzintervallen bestimmt. Dieses Effektmaß wird gegenüber der Auswertung der naiven Raten aufgrund der Berücksichtigung der Zeitkomponente in der ereignisgetriebenen GRIPHON-Studie als geeignetere Auswertung für Endpunkte angesehen (vgl. auch die Ausführungen zum Verzerrungspotential der naiven Raten auf S. 124 ff.).

Wenn mehrere Zeitpunkte eines Endpunkts ausgewertet wurden, wurde i. d. R. derjenige Zeitpunkt zur Bestimmung des Ausmaßes des Zusatznutzens gewählt, der in der Endpunkthierarchie des statistischen Auswertungsplans von GRIPHON den höchsten Rang einnahm.

Tabelle 4-127: Zusammenfassung der Ergebnisse zur Bestimmung des Zusatznutzens von Macitentan

Endpunkt	**Effektschätzer [KI]** **p-Wert**	**Ausmaß des Zusatznutzens**
Morbidität		
Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis	HR [99 % KI]: **0,60 [0,46; 0,78]** p<**0,0001**	**Beträchtlich**
Tod	RR [95 % KI]: **1,58 [0,88; 2,82]** p=**0,1289**	**Zusatznutzen/größerer Schaden nicht belegt**
Hospitalisierung wegen PAH	RR [95 % KI]: **0,73 [0,56; 0,95]** p=**0,0179**	**Beträchtlich**
Lungentransplantation oder atriale Septostomie	RR [95 % KI]: **0,51 [0,05; 5,58]** p=**0,6831**	**Zusatznutzen/größerer Schaden nicht belegt**
Parenterale Prostanoid- oder dauerhafte Sauerstofftherapie	RR [95 % KI]: **0,78 [0,35; 1,76]** p=**0,5806**	**Zusatznutzen/größerer Schaden nicht belegt**
Fortschreiten der Erkrankung	RR [95 % KI]: **0,39 [0,27; 0,55]** p<**0,0001**	**Beträchtlich**
Zeit bis zum Tod wegen PAH oder Hospitalisierung wegen PAH	HR [99 % KI]: **0,70 [0,50; 0,98]** p=**0,0031**	**Beträchtlich**
Tod	RR [95 % KI]: **1,16 [0,57; 2,35]** p=**0,7534**	**Zusatznutzen/größerer Schaden nicht belegt**
Hospitalisierung	RR [95 % KI]: **0,71 [0,55; 0,91]** p=**0,0067**	**Beträchtlich**
Zeit bis zur ersten Hospitalisierungen wegen PAH	HR [99 % KI]: **0,65 [0,50; 0,86]** p=**0,0022**	**Beträchtlich**
Symptome der PAH[1]	RR [95 % KI]: **0,85 [0,75; 0,96]** p=**0,0081**	**Beträchtlich**
Absolute Veränderung der 6-MWD (vom Basiswert bis Woche 26)	Median [99 % KI]: **12,0 m [1 m, 24 m]** p=**0,0027**	**Gering**
Ausbleibende Verschlechterung der WHO-/ NYHA-Klasse (vom Basiswert bis Woche 26)	OR [99 % KI]: **1,16 [0,81; 1,66]** p=**0,2843**[2]	**Zusatznutzen/größerer Schaden nicht belegt**
Absolute Veränderung des Borg-Dyspnoe-Index (vom Basiswert bis Woche 26)	Median [95 % KI]: **0,0 [0,0; 0,0]** p=**0,0061**	**Zusatznutzen/größerer Schaden nicht belegt**

Mortalität		
Zeit bis zum Tod (Studienschluss)	HR [99 % KI]: **0,97 [0,68; 1,39]** p=**0,4214**	**Zusatznutzen/größerer Schaden nicht belegt**
Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)	HR [99 % KI]: **1,16 [0,70; 1,93]** p=**0,7153**	**Zusatznutzen/größerer Schaden nicht belegt**
Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)	HR [99 % KI]: **0,86 [0,63; 1,18]** p=**0,1763**	**Zusatznutzen/größerer Schaden nicht belegt**
Lebensqualität		
Absolute Veränderung des CAMPHOR „Symptome“-Scores (vom Basiswert bis Woche 26)	Median [99 % KI]: **0,0 [-1,0, 1,0]** p=**0,2185**	**Zusatznutzen/größerer Schaden nicht belegt**
Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR „Symptome“-Scores (vom Basiswert bis Woche 26)	Median [99 % KI]: **0,0 [-0,4; 0,0]** p=**0,1700**	**Zusatznutzen/größerer Schaden nicht belegt**
Nebenwirkungen		
Unerwünschte Ereignisse	RR [95 % KI]: **1,01 [1,00; 1,03]** p=**0,1329**	**Zusatznutzen/größerer Schaden nicht belegt**
Schwere unerwünschte Ereignisse	RR [95 % KI]: **1,09 [0,96; 1,23]** p=**0,2124**	**Zusatznutzen/größerer Schaden nicht belegt**
Schwerwiegende unerwünschte Ereignisse	RR [95 % KI]: **0,93 [0,82; 1,06]** p=**0,2902**	**Zusatznutzen/größerer Schaden nicht belegt**
Therapieabbrüche wegen unerwünschter Ereignisse	RR [95 % KI]: **0,85 [0,73; 1,00]** p=**0,0554**	**Zusatznutzen/größerer Schaden nicht belegt**
Unerwünschte Ereignisse von besonderem Interesse		
Hämorrhagie	RR [95 % KI]: **0,99 [0,76; 1,30]** p=**0,9756**	**Zusatznutzen/größerer Schaden nicht belegt**
Hypotonie	RR [95 % KI]: **1,55 [0,92; 2,62]** p=**0,1025**	**Zusatznutzen/größerer Schaden nicht belegt**
Störung des Auges	RR [95 % KI]: **1,37 [0,96; 1,97]** p=**0,0968**	**Zusatznutzen/größerer Schaden nicht belegt**
Prostacyclin-assoziierte unerwünschte Ereignisse	RR [95 % KI]: **1,45 [1,35; 1,55]** p<**0,0001**	**Zusatznutzen/größerer Schaden nicht belegt**

[1]) Der Endpunkt umfasst die vom G-BA als wichtig erachteten PAH-Symptome Dyspnoe, Synkope, Ödeme, Schwindel und Müdigkeit

[2]) p-Wert aus Cochran-Mantel-Haenszel-Test

Nähere Informationen zur verwendeten statistischen Methodik finden sich in den jeweiligen Unterabschnitten von Abschnitt 4.3.1.3.

HR – Hazard Ratio, RR – Risk Ratio, OR – Odds Ratio, KI – Konfidenzintervall

Morbidität

Der Einsatz von Selexipag führt zu einer deutlichen und gegenüber der Vergleichstherapie bisher nicht erreichten Verbesserung des therapierelevanten Nutzens i. S. v. § 5 Abs. 7 Satz 2 der AM-NutzenV in der Behandlung von Patienten im Anwendungsgebiet von Selexipag, bei denen das *Risiko* einer Behandlung mit Iloprost zur Inhalation dessen *Nutzen* noch überwiegt.

Insbesondere führt Selexipag zu einer deutlichen und klinisch relevanten Verringerung des Risikos für das Auftreten eines Ereignisses des Endpunkts „Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis“ und damit zu einer für die Patienten spürbaren Linderung der Erkrankung bzw. einer Abschwächung schwerwiegender Symptome (jeweils im Sinne einer Vermeidung der Progression der Erkrankung).[41] Der für den primären Endpunkt beobachtete Unterschied zwischen den Behandlungsgruppen wird hierbei insbesondere durch eine Vermeidung von Hospitalisierungen, aber auch durch die Verhinderung des Fortschreitens der Erkrankung getrieben. Beide Komponenten für sich genommen würden ebenfalls bereits auf eine spürbare Linderung der Erkrankung (Fortschreiten der Erkrankung) bzw. eine Abschwächung schwerwiegender Symptome (indirekt erkennbar an einer geringeren Anzahl Hospitalisierungen) schließen lassen. Sensitivitätsanalysen, bei denen der primäre Endpunkt unter Ausschluss der vom G-BA kritisch bewerteten Komponenten „Fortschreiten der Erkrankung“ und „Initiierung einer Prostanoid- oder dauerhaften O_2-Therapie“ ausgewertet wurden, zeigen, dass die Ereignisse der beiden Komponenten nicht alleine ausschlaggebend für das Ergebnis des gesamten kombinierten Endpunkts sind. Auch nach Ausschluss beider Komponenten aus der Analyse verbleibt ein deutlicher und klinisch relevanter Effekt.

In der Auswertung der Zeit bis zur ersten Hospitalisierung aufgrund einer Verschlechterung der PAH oder bis zum Tod wegen PAH wurde zudem eine deutliche Verringerung des kombinierten Risikos für das Auftreten eines der beiden Ereignisse beobachtet. Der Unterschied zwischen den Behandlungsgruppen war hierbei überwiegend auf die Verringerung des Risikos einer PAH-bedingten Hospitalisierung zurückzuführen. Diese Beobachtung ist konsistent zur der Auswertung der Zeit bis zur ersten Hospitalisierung wegen PAH, in der ebenfalls ein verringertes Risiko für das Auftreten einer PAH-bedingten Hospitalisierung durch Selexipag beobachtet wurde. Insgesamt gesehen führt Selexipag demnach zu einer deutlichen und klinisch relevanten Verlängerung der Zeit bis zur stationären

[41] Die spürbare Linderung der Erkrankung bzw. die Abschwächung schwerwiegender Symptome im Sinne der AM-NutzenV ergibt sich hierbei durch die Vermeidung der Progression, d. h. im Vergleich zum schwereren Krankheitsstand des Patienten, wenn er nicht mit Selexipag behandelt werden würde.

Aufnahme wegen der PAH, was mit einer Abschwächung schwerwiegender Symptome gleichzusetzen ist.

Selexipag führt darüber hinaus zu einer signifikanten Verbesserung der 6-Minuten-Gehstrecke – einem Maß für die körperliche Leistungsfähigkeit des Patienten – und verringert das kombinierte Risiko des Auftretens der PAH-Symptome Dyspnoe, Synkope, Müdigkeit, Schwindel oder Ödeme. Folglich führt Selexipag zu einer relevanten Abschwächung schwerwiegender bzw. einer Verringerung von nicht-schwerwiegenden Symptomen der PAH.

Im Vergleich zu Placebo blieb die Symptomatik der PAH (bestimmt anhand der WHO-/ NYHA-Klasse) bei einem größeren Anteil der Patienten unter Selexipag stabil bzw. verbesserte sich sogar. Der Unterschied zwischen den Behandlungsgruppen war statistich nicht signifikant.

Insgesamt gesehen führt Selexipag zu einer Abschwächung schwerwiegender Symptome sowie zu einer für die Patienten spürbaren Linderung der Erkrankung (ggf. im Sinne einer Vermeidung einer Verschlechterung) und einer Verbesserung der körperlichen Leistungsfähigkeit. Gemäß § 5 Abs. 7 der AM-NutzenV ist der Zusatznutzen von Selexipag in der Endpunktkategorie Morbidität folglich als „beträchtlich“ anzunehmen.

Mortalität

Im Rahmen der GRIPHON-Studie konnte zwischen den Behandlungsgruppen kein wesentlicher Unterschied in Bezug auf die PAH-bedingte Mortalität und die Gesamtmortalität beobachtet werden. Eine leichte numerische Imbalance der PAH-bedingten Mortalität in der Auswertung bis zum individuellen Behandlungsende (EOT) durch Selexipag im Vergleich zu Placebo ist weitgehend auf eine Verzerrung durch informative Zensierung der Daten zurückzuführen. Überdies war der Unterschied statistisch nicht signifikant.

Auch die Analyse der Gesamtheit der vorhandenen Daten zur PAH-bedingten Mortalität, d. h. bis zum Studienschluss bestätigt, dass kein nachteiliger Effekt auf die Mortalität durch Selexipag vorliegt. In dieser Analyse wurde – trotz einer möglichen Verzerrung des Effekts zulasten von Selexipag durch den erlaubten Therapiewechsel von Placebo zu Selexipag nach Ende der Behandlungsphase – eine numerische Reduktion des Mortalitätsrisikos beobachtet, die jedoch nicht statistisch signifikant war.

Das Risiko eines Todesfalles jeglicher Ursache bis Studienschluss war in beiden Studienarmen vergleichbar, es ist aber aufgrund der Ergebnisse von Modellen, die die Effekte durch den erlaubten Therapiewechsel von Selexipag zu Placebo korrigieren, davon auszugehen, dass eine Verzerrung der Todesfälle zu Ungunsten von Selexipag vorliegt.

In der Gesamtbetrachtung sind insbesondere die Ergebnisse der Mortalitätsendpunkte als relevant zu erachten, bei denen der Erhebungszeitraum bis zum Studienschluss reichte und in deren Auswertung somit sämtliche im Studienzeitraum aufgetretenen Todesfälle berücksichtigt werden.

Ein Zusatznutzen/größerer Schaden von Selexipag in der Endpunktkategorie Mortalität ist daher nicht belegt.

Lebensqualität

Selexipag erweist sich als vergleichbar gegenüber Placebo sowohl in der Auswertung des „Symptome"-Scores des CAMPHOR-Fragebogens als auch in der „Kurzatmigkeit"-Subskala des gleichen Scores. Dies ist insofern bemerkenswert, als trotz der gewählten Titrationsstrategie, die eine Dosierung bis zum Auftreten von für den Patienten nicht mehr erträglichen prostacyclinspezifischen Nebenwirkungen vorsah, keine negative Beeinträchtigung der Lebensqualität beobachtet wurde.

Ein Zusatznutzen/größerer Schaden in der Endpunktkategorie Lebensqualität ist damit nicht belegt.

Nebenwirkungen

Die Behandlung mit Selexipag erweist sich insgesamt als ähnlich sicher und verträglich wie die Behandlung mit Placebo und führt zu einer im Vergleich zu Placebo ausgeglichenen Gesamtrate der unerwünschten Ereignisse. Das Risiko für das Auftreten schwerwiegender unerwünschter Ereignisse und für Therapieabbrüche aufgrund unerwünschter Ereignisse wurde tendenziell verringert.

Unter Selexipag traten insbesondere im Verlauf der Titrationsphase häufiger Prostacyclin-assoziierte unerwünschte Ereignisse auf als unter Placebo. Aufgrund der Titrationsstrategie, in der das Auftreten entsprechender UE *de facto* Voraussetzung zur Bestimmung der patientenindividuell optimalen Dosis war, und aufgrund des Wirkmechanismus war ein solcher Unterschied zulasten von Selexipag zu erwarten. Bei den Prostacyclin-assoziierten UE handelt es sich zudem vorwiegend um Ereignisse milder oder moderater Intensität, die mit üblichen Therapiemethoden i. d. R. gut zu beherrschen waren. Sie betreffen zudem vorwiegend die Verträglichkeit von Selexipag und nicht dessen Sicherheit.

Für die Endpunktkategorie Nebenwirkungen ist ein Zusatznutzen/größerer Schaden somit nicht belegt.

Fazit

Die für die Teilpopulation a vorgelegte Studie GRIPHON besitzt ein niedriges Verzerrungspotential auf Studien- und Endpunktebene und somit eine hohe Aussagekraft. Es haben sich darüber hinaus keine Hinweise ergeben, die gegen eine Übertragbarkeit auf den deutschen Versorgungskontext sprechen. Aus GRIPHON lässt sich somit ein Hinweis auf einen Zusatznutzen ableiten.

Die Ergebnisse aus GRIPHON zeigen, dass Selexipag das erste den Prostacyclin-Signalweg adressierende Arzneimittel in der PAH ist, durch dessen Anwendung in einer kontrollierten, ereignisgetriebenen Langzeitstudie eine deutliche und klinisch relevante Reduktion patientenrelevanter Parameter der Erkrankung erreicht werden konnte. In der Endpunktkategorie Morbidität wurde hierbei ein beträchtlicher Zusatznutzen gegenüber

Placebo nachgewiesen, in den übrigen Endpunktkategorien Mortalität, Lebensqualität und Nebenwirkungen ist ein Zusatznutzen/größerer Schaden nicht belegt.

In der Gesamtschau der einzelnen Endpunktkategorien ergibt sich eine bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens, die insbesondere auf der Abschwächung schwerwiegender Symptome, einer für den Patienten spürbaren Linderung der Erkrankung und einer Verbesserung der körperlichen Leistungsfähigkeit beruht. Gemäß § 5 Abs. 7 der AM-NutzenV ist das Ausmaß des Zusatznutzens von Selexipag in der bewerteten Teilpopulation a mit „beträchtlich“ anzunehmen.

Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der Nutzen einer Behandlung mit Iloprost deren Risiko überwiegt (Teilpopulation b)

Wie bereits im vorherigen Abschnitt 4.4.1 beschrieben, konnte für die Teilpopulation b keine relevante RCT identifiziert werden, die eine direkte Bewertung des Zusatznutzens von Selexipag gegenüber einer patientenindividuellen Anwendung von Iloprost zur Inhalation ermöglicht. Ebenso war auch ein indirekter Vergleich nicht möglich, da die zur Verfügung stehenden Iloprost-Studien mit einer Studiendauer von 12 Wochen zu kurz für eine Aussage zum Zusatznutzen in einer chronisch progredienten Erkrankung wie der PAH waren.

Auch wenn damit keine RCT-basierte klinische Evidenz zur Bestimmung des Ausmaßes des Zusatznutzens vorliegt, so lassen sich dennoch verschiedene qualitative Vorteile von Selexipag gegenüber der patientenindividuellen Anwendung von Iloprost zur Inhalation feststellen:

1. Iloprost ist in Deutschland ausschließlich zur inhalativen Anwendung zugelassen. Aufgrund seiner vergleichweise kurzen Halbwertszeit ist es notwendig, dass Iloprost 6- bis 9-mal täglich inhaliert wird, wobei jede Inhalation mit einem Zeitaufwand von 4 bis 10 Minuten verbunden ist [4]. Hinzu kommt die Zeit, die für die Vorbereitung des Inhalators und den anschließenden Reinigungsvorgang aufgewendet werden muss, sodass sich die Gesamtzeit für die Anwendung von Iloprost auf 15-30 Minuten je Inhalation bzw. auf 1,5-4,5 Stunden (bei 6-9 Inhalationen) je Behandlungstag summiert.

   Die inhalative Anwendung von Iloprost ist zudem nicht komplikationslos. Aufgrund der Ablagerung des Wirkstoff-Aerosols im Inhalator kann es zu größeren und möglicherweise variablen Unterschieden zwischen vorbereiteter und tatsächlich inhalierter Dosierung kommen [5]. Weiterhin haben auch die Tröpfchengröße und die Aerosolkonzentration einen bedeutenden Einfluss auf den Ablagerungsort in den Atemwegen und damit auch auf die Aufnahme in die pulmonalen Gefäße [5].

   Es ist davon auszugehen, dass die Komplikationen und Probleme der inhalativen Anwendung von Iloprost, insbesondere der große Zeitaufwand, der für die Behandlung aufgewendet werden muss, einen starken negativen Einfluss auf die Lebensqualität der Patienten haben und auch zur Beeinträchtigung der Therapieadhärenz der Patienten führen könnten. Durch seine einfache und

komplikationsfreie Anwendung als zweimal täglich oral einzunehmende Tablette besitzt Selexipag die genannten Einschränkungen von inhalativem Iloprost nicht, es ist daher auch nicht von einem nachteiligen Effekt auf Lebensqualität und Therapieadhärenz auszugehen. Dies wird auch durch die Ergebnisse der GRIPHON-Studie bestätigt, in der trotz der gewählten Titrationsstrategie, die *de facto* das Auftreten von unerwünschen Ereignissen voraussetzte, keine negative Beeinträchtigung der Lebensqualität beobachtet wurde.

2. Im Gegensatz zu Selexipag ist das Prostacyclin-Analogon Iloprost kein selektiver IP-Rezeptor-Agonist, sondern bindet (mit unterschiedlicher Affinität) auch andere Rezeptoren der Prostanoid-Rezeptor-Familie (vgl. Modul 2). Neben dem IP-Rezeptor aktiviert Iloprost beispielsweise auch die Prostanoid-Rezeptoren EP_1 und EP_3, mit zum Teil vergleichbarer Potenz. Dies kann aber beispielsweise dazu führen, dass der IP-Rezeptor-vermittelte vasodilatative Effekt durch die vasokonstriktive Aktivität des EP_3-Rezeptors negativ beeinflusst und sogar aufgehoben wird [91].

   Weiterhin kann die fehlende Selektivität für den IP-Rezeptor aber auch ungewollte Nebenwirkungen auslösen bzw. deren Schwere und/oder Dauer negativ beeinflussen. Die Aktivierung der EP_1- und EP_3-Rezeptoren durch nicht-selektive Prostacyclin-Analoga kann u. a. eine Kontraktion der glatten Muskulatur im Magen auslösen [92], was zu Abdominalschmerzen oder Bauchkrämpfen – häufigen Nebenwirkungen der Prostanoidtherapie – führen kann. Versuche, die mit dem Magenfundus von Ratten durchgeführt wurden, konnten in diesem Zusammenhang zeigen, dass Selexipag und ACT-333679 im Gegensatz zu Iloprost nicht zu einer Kontraktion der Magenmuskulatur führen [92]. Durch die gleichzeitige Behandlung mit einem EP_3-Antagonisten konnten die Iloprost-induzierten Kontraktionen signifikant reduziert werden.

   Die Selektivität Selexipags für den IP-Rezeptor ermöglicht es, Patienten mit einer wirksamen Dosierung zu behandeln, ohne dass die Behandlung durch Tolerabilitätsprobleme eingeschränkt würde.

3. Das vorrangige Ziel der Therapie einer chronisch progressiven Erkrankung wie der PAH ist grundsätzlich immer der Langzeittherapieerfolg. Um diesen sicherzustellen, ist es im Erkrankungsbild der PAH nicht ausreichend, die Wirksamkeit eines Wirkstoffs anhand der Veränderung der 6-MWD – einem Endpunkt, der nachweislich keine Aussagekraft für die Langzeitprognose der PAH besitzt (vgl. Abschnitt 4.2.5.2) – und nur über einen vergleichsweise kurzen Zeitraum von weniger als 6 Monaten zu bestimmen. Vielmehr ist es notwendig, die Wirksamkeit des Wirkstoffs anhand eines klinischen und patientenrelevanten Endpunkts zu bestimmen, der neben der Mortalität die Zeit bis zur klinischen Verschlechterung verlässlich abbildet und gleichzeitig über einen Zeitraum von mehr als 6 Monaten untersucht wird. Dementsprechend unterscheidet die EMA in der PAH auch zwischen Zulassungen zur kurzfristigen Verbesserung der körperlichen Leistungsfähigkeit und Zulassungen zur Langzeitbehandlung [16].

Iloprost ist in Europa ausschließlich für die kurzfristige Verbesserung der körperlichen Leistungsfähigkeit und der Symptomatik zugelassen [4]. Für Selexipag hatte die EMA hingegen explizit eine Zulassung für die Langzeittherapie der PAH erteilt [1]. Durch die Überprüfung der Wirksamkeit in der Zulassung ist demnach sichergestellt, dass Selexipag zum Erreichen eines Langzeittherapieerfolgs geeignet ist, während in Bezug auf die Iloprost-Behandlung unklar ist, ob ein über die kurzfristige Verbesserung der körperlichen Leistungsfähigkeit hinausgehender Effekt in der Langzeittherapie erreicht wird.

Die dargestellten Argumente zeigen deutlich, dass ein Zusatznutzen von Selexipag gegenüber der zweckmäßigen Vergleichstherapie in Teilpopulation b grundsätzlich vorhanden ist. Da keine direkt vergleichenden RCT vorliegen und auch indirekte Vergleiche nicht möglich sind, lässt die wissenschaftliche Datengrundlage derzeit eine Quantifizierung des Zusatznutzens nicht zu. Gemäß § 5 Abs. 7 der AM-NutzenV ist das Ausmaß des Zusatznutzens von Selexipag in der bewerteten Teilpopulation b daher mit „nicht quantifizierbar“ anzunehmen.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-128: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	**Ausmaß des Zusatznutzens**
Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten das *Risiko* einer Behandlung mit Iloprost deren *Nutzen* noch überwiegt	Hinweis auf einen beträchtlichen Zusatznutzen
Patienten im Anwendungsgebiet von Selexipag, bei denen aufgrund des Erkrankungszustandes des Patienten der *Nutzen* einer Behandlung mit Iloprost deren *Risiko* überwiegt	Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005[42],

[42] *Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.*

Molenberghs 2010[43]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006[44]) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006[45]) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

– Patientenpopulation
– Intervention
– Kontrolle
– Datenherkunft
– verwendete Methodik
– entsprechende Ergebnisse (abhängig von der Methode)
– Untersuchungen zur Robustheit
– ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht zutreffend.

---

[43] *Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.*
[44] *Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.*
[45] *Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.*

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Tabelle 4-129: Liste der eingeschlossenen Studien

Kurzbezeichnung	Zitat	Titel
GRIPHON (AC-065A302)	[76-80]	Prostacyclin (PGI_2) receptor agonist in pulmonary arterial hypertension - A multicenter, double-blind, placebo-controlled Phase 3 study assessing the efficacy and safety of selexipag (ACT-293987) on morbidity and mortality in patients with pulmonary arterial hypertension.

4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Actelion Pharmaceuticals Deutschland GmbH (2016): Uptravi® Filmtabletten. Fachinformation. Stand: Mai 2016.
2. Gemeinsamer Bundesausschuss (2016): Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV - Beratungsanforderung 2015-B-141. Wirkstoff: Macitentan.
3. Gemeinsamer Bundesausschuss (2015): Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV - Beratungsanforderung 2015-B-021. Wirkstoff: Selexipag.
4. Bayer Vital GmbH (2003): Ventavis® 10 Mikrogramm/ml Lösung für einen Vernebler; Fachinformation. Stand: November 2014 [Zugriff: 10.03.2016]. URL: http://www.fachinfo.de.
5. Olschewski H, Rohde B, Behr J, Ewert R, Gessler T, Ghofrani HA, et al. (2003): Pharmacodynamics and pharmacokinetics of inhaled iloprost, aerosolized by three different devices, in severe pulmonary hypertension. Chest; 124(4):1294-304.
6. OMT Med GmbH & Co. KG (2012): REMODULIN 10 mg/ml Infusionslösung zur subkutanen Anwendung; Fachinformation. Stand: August 2014 [Zugriff: 10.03.2016]. URL: http://www.fachinfo.de.
7. Panmedica (2011): Epoprostenol-Rotexmedica 1,5 mg; Fachinformation. Stand: Juli 2015 [Zugriff: 10.03.2016]. URL: http://www.fachinfo.de.
8. Humbert M, Sitbon O, Simonneau G (2004): Treatment of pulmonary arterial hypertension. The New England journal of medicine; 351(14):1425-36.
9. Mubarak KK (2010): A review of prostaglandin analogs in the management of patients with pulmonary arterial hypertension. Respiratory medicine; 104(1):9-21.
10. Kallen AJ, Lederman E, Balaji A, Trevino I, Petersen EE, Shoulson R, et al. (2008): Bloodstream infections in patients given treatment with intravenous prostanoids. Infection control and hospital epidemiology; 29(4):342-9.

11. Simonneau G, Barst RJ, Galie N, Naeije R, Rich S, Bourge RC, et al. (2002): Continuous subcutaneous infusion of treprostinil, a prostacyclin analogue, in patients with pulmonary arterial hypertension: a double-blind, randomized, placebo-controlled trial. American journal of respiratory and critical care medicine; 165(6):800-4.
12. Galiè N, Hoeper MM, Humbert M, Torbicki A, Vachiery JL, Barbera JA, et al. (2009): Guidelines for the diagnosis and treatment of pulmonary hypertension: the Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT). European heart journal; 30(20):2493-537.
13. Ghofrani HA, Distler O, Gerhardt F, Gorenflo M, Grunig E, Haefeli WE, et al. (2011): Treatment of pulmonary arterial hypertension (PAH): updated Recommendations of the Cologne Consensus Conference 2011. International journal of cardiology; 154 Suppl 1:S20-33.
14. Gemeinsamer Bundesausschuss (2013): Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV - Beratungsanforderung 2012-B-057. Wirkstoff: Macitentan.
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2015): Allgemeine Methoden Version 4.2. [Zugriff: 22.04.2015]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-2.pdf.
16. European Medicines Agency - Committee for Medicinal Products for Human Use (2009): Guideline on the clinical investigation of medicinal products for the treatment of pulmonary arterial hypertension [Zugriff: 12.09.2012]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/12/WC500016686.pdf.
17. Galiè N, Corris PA, Frost A, Girgis RE, Granton J, Jing ZC, et al. (2013): Updated Treatment Algorithm of Pulmonary Arterial Hypertension. Journal of the American College of Cardiology; 62(25, Supplement):D60-D72.
18. Galiè N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, et al. (2016): 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS)Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). European heart journal; 37(1):67-119.
19. Coghlan JG, Pope J, Denton CP (2010): Assessment of endpoints in pulmonary arterial hypertension associated with connective tissue disease. Current opinion in pulmonary medicine; 16 Suppl 1:S27-34.
20. Peacock A, Keogh A, Humbert M (2010): Endpoints in pulmonary arterial hypertension: the role of clinical worsening. Current opinion in pulmonary medicine; 16 Suppl 1:S1-9.
21. Ventetuolo CE, Benza RL, Peacock AJ, Zamanian RT, Badesch DB, Kawut SM (2008): Surrogate and combined end points in pulmonary arterial hypertension. Proceedings of the American Thoracic Society; 5(5):617-22.
22. Galiè N, Simonneau G, Barst RJ, Badesch D, Rubin L (2010): Clinical worsening in trials of pulmonary arterial hypertension: results and implications. Current opinion in pulmonary medicine; 16 Suppl 1:S11-9.
23. McLaughlin VV, Archer SL, Badesch DB, Barst RJ, Farber HW, Lindner JR, et al. (2009): ACCF/AHA 2009 expert consensus document on pulmonary hypertension: a

report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents and the American Heart Association: developed in collaboration with the American College of Chest Physicians, American Thoracic Society, Inc., and the Pulmonary Hypertension Association. Circulation; 119(16):2250-94.

24. Delcroix M, Channick R, Galiè N, Ghofani H-A, Jansa P, Le Brun F-O, et al. (2013): Is 6-minute walk distance (6MWD) associated with long-term outcomes in pulmonary arterial hypertension (PAH)? Results from SERAPHIN [Präsentiert auf dem European Respiratory Society Annual Congress, Barcelona, Spanien, 7–11.09.2013] - VERTRAULICH.
25. Gemeinsamer Bundesausschuss (2014): Mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses - hier: Wirkstoff Riociguat. [Zugriff: 12.12.2014]. URL: https://www.g-ba.de/downloads/91-1031-110/2014-09-08_Wortprotokoll_end_Riociguat.pdf.
26. Chakinala MM, Barst R (2013): From short-term benefits to long-term outcomes: the evolution of clinical trials in pulmonary arterial hypertension. Pulmonary circulation; 3(3):507-22.
27. Gomberg-Maitland M, Bull TM, Saggar R, Barst RJ, Elgazayerly A, Fleming TR, et al. (2013): New Trial Designs and Potential Therapies for Pulmonary Artery Hypertension. Journal of the American College of Cardiology; 62(25, Supplement):D82-D91.
28. McLaughlin VV, Badesch DB, Delcroix M, Fleming TR, Gaine SP, Galie N, et al. (2009): End points and clinical trial design in pulmonary arterial hypertension. J Am Coll Cardiol; 54(1 Suppl):S97-107.
29. McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, et al. (2014): Angiotensin-neprilysin inhibition versus enalapril in heart failure. The New England journal of medicine; 371(11):993-1004.
30. Halpern SD, Doyle R, Kawut SM (2008): The ethics of randomized clinical trials in pulmonary arterial hypertension. Proceedings of the American Thoracic Society; 5(5):631-5.
31. Barst RJ, Langleben D, Badesch D, Frost A, Lawrence EC, Shapiro S, et al. (2006): Treatment of pulmonary arterial hypertension with the selective endothelin-A receptor antagonist sitaxsentan. J Am Coll Cardiol; 47(10):2049-56.
32. Actelion Pharmaceuticals Ltd (2015): GRIPHON-Studie - Analyse der Zeit bis zum Tod von Monat 6 bis zum Studienschluss: Landmark-Analyse (Selexipag- und Placebo-Gruppe wurden zusammengefasst) der Kaplan-Meier-Schätzer nach Auftreten oder Nicht-Auftreten eines CEC-bestätigten Morbiditätsereignisses als erstes Ereignis vor Monat 6. Post-hoc-Analyse. VERTRAULICH.
33. Hoeper MM, Oudiz RJ, Peacock A, Tapson VF, Haworth SG, Frost AE, et al. (2004): End points and clinical trial designs in pulmonary arterial hypertension: clinical and regulatory perspectives. J Am Coll Cardiol; 43(12 Suppl S):48S-55S.
34. Ghofrani HA, Wilkins MW, Rich S (2008): Uncertainties in the diagnosis and treatment of pulmonary arterial hypertension. Circulation; 118(11):1195-201.
35. Batal O, Khatib OF, Dweik RA, Hammel JP, McCarthy K, Minai OA (2012): Comparison of baseline predictors of prognosis in pulmonary arterial hypertension in patients surviving </=2 years and those surviving >/=5 years after baseline right-sided cardiac catheterization. The American journal of cardiology; 109(10):1514-20.

36. Benza RL, Miller DP, Gomberg-Maitland M, Frantz RP, Foreman AJ, Coffey CS, et al. (2010): Predicting survival in pulmonary arterial hypertension: insights from the Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL). Circulation; 122(2):164-72.
37. Humbert M, Sitbon O, Chaouat A, Bertocchi M, Habib G, Gressin V, et al. (2010): Survival in patients with idiopathic, familial, and anorexigen-associated pulmonary arterial hypertension in the modern management era. Circulation; 122(2):156-63.
38. Lee WT, Ling Y, Sheares KK, Pepke-Zaba J, Peacock AJ, Johnson MK (2012): Predicting survival in pulmonary arterial hypertension in the UK. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 40(3):604-11.
39. Nickel N, Golpon H, Greer M, Knudsen L, Olsson K, Westerkamp V, et al. (2012): The prognostic impact of follow-up assessments in patients with idiopathic pulmonary arterial hypertension. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 39(3):589-96.
40. Taguchi H, Kataoka M, Yanagisawa R, Kawakami T, Tamura Y, Fukuda K, et al. (2012): Platelet level as a new prognostic factor for idiopathic pulmonary arterial hypertension in the era of combination therapy. Circulation journal : official journal of the Japanese Circulation Society; 76(6):1494-500.
41. Wensel R, Francis DP, Meyer FJ, Opitz CF, Bruch L, Halank M, et al. (2012): Incremental prognostic value of cardiopulmonary exercise testing and resting haemodynamics in pulmonary arterial hypertension. International journal of cardiology;
42. Sitbon O, Humbert M, Nunes H, Parent F, Garcia G, Herve P, et al. (2002): Long-term intravenous epoprostenol infusion in primary pulmonary hypertension: prognostic factors and survival. J Am Coll Cardiol; 40(4):780-8.
43. Farber HW (2012): Validation of the 6-minute walk in patients with pulmonary arterial hypertension: trying to fit a square PEG into a round hole? Circulation; 126(3):258-60.
44. Macchia A, Marchioli R, Tognoni G, Scarano M, Marfisi R, Tavazzi L, et al. (2010): Systematic review of trials using vasodilators in pulmonary arterial hypertension: why a new approach is needed. American heart journal; 159(2):245-57.
45. Fritz JS, Blair C, Oudiz RJ, Dufton C, Olschewski H, Despain D, et al. (2013): Baseline and follow-up 6-min walk distance and brain natriuretic peptide predict 2-year mortality in pulmonary arterial hypertension. Chest; 143(2):315-23.
46. Gabler NB, French B, Strom BL, Palevsky HI, Taichman DB, Kawut SM, et al. (2012): Validation of 6-minute walk distance as a surrogate end point in pulmonary arterial hypertension trials. Circulation; 126(3):349-56.
47. Provencher S, Sitbon O, Humbert M, Cabrol S, Jais X, Simonneau G (2006): Long-term outcome with first-line bosentan therapy in idiopathic pulmonary arterial hypertension. European heart journal; 27(5):589-95.
48. Gemeinsamer Bundesausschuss (2014): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Macitentan. [Zugriff: 25.11.2014]. URL: https://www.g-ba.de/downloads/40-268-2888/2014-07-17_AM-RL-XII_Macitentan_2014-02-01-D-096_TrG.pdf.

49. Gemeinsamer Bundesausschuss (2014): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Riociguat. [Zugriff: 04.05.2015]. URL: https://www.g-ba.de/downloads/40-268-2978/2014-10-16_AM-RL-XII_Riociguat_2014-05-01-D-103_TrG.pdf.
50. Pulido T, Adzerikho I, Channick RN, Delcroix M, Galiè N, Ghofrani H-A, et al. (2013): Macitentan and Morbidity and Mortality in Pulmonary Arterial Hypertension. New England Journal of Medicine; 369(9):809-18.
51. Ghofrani HA, Galie N, Grimminger F, Grunig E, Humbert M, Jing ZC, et al. (2013): Riociguat for the treatment of pulmonary arterial hypertension. The New England journal of medicine; 369(4):330-40.
52. Galiè N, Brundage BH, Ghofrani HA, Oudiz RJ, Simonneau G, Safdar Z, et al. (2009): Tadalafil therapy for pulmonary arterial hypertension. Circulation; 119(22):2894-903.
53. Galie N, Ghofrani HA, Torbicki A, Barst RJ, Rubin LJ, Badesch D, et al. (2005): Sildenafil citrate therapy for pulmonary arterial hypertension. The New England journal of medicine; 353(20):2148-57.
54. Rubin LJ, American College of Chest P (2004): Diagnosis and management of pulmonary arterial hypertension: ACCP evidence-based clinical practice guidelines. Chest; 126(1 Suppl):7S-10S.
55. Hunt SA, Baker DW, Chin MH, Cinquegrani MP, Feldman AM, Francis GS, et al. (2001): ACC/AHA Guidelines for the Evaluation and Management of Chronic Heart Failure in the Adult: Executive Summary A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1995 Guidelines for the Evaluation and Management of Heart Failure): Developed in Collaboration With the International Society for Heart and Lung Transplantation; Endorsed by the Heart Failure Society of America. Circulation; 104(24):2996-3007.
56. Benza RL, Gomberg-Maitland M, Naeije R, Arneson CP, Lang IM (2011): Prognostic factors associated with increased survival in patients with pulmonary arterial hypertension treated with subcutaneous treprostinil in randomized, placebo-controlled trials. The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation; 30(9):982-9.
57. D'Alonzo GE, Barst RJ, Ayres SM, Bergofsky EH, Brundage BH, Detre KM, et al. (1991): Survival in patients with primary pulmonary hypertension. Results from a national prospective registry. Annals of internal medicine; 115(5):343-9.
58. McLaughlin VV, Shillington A, Rich S (2002): Survival in primary pulmonary hypertension: the impact of epoprostenol therapy. Circulation; 106(12):1477-82.
59. McLaughlin VV, Sitbon O, Badesch DB, Barst RJ, Black C, Galie N, et al. (2005): Survival with first-line bosentan in patients with primary pulmonary hypertension. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 25(2):244-9.
60. Swiston JR, Johnson SR, Granton JT (2010): Factors that prognosticate mortality in idiopathic pulmonary arterial hypertension: a systematic review of the literature. Respiratory medicine; 104(11):1588-607.

61. Klose H, Opitz C, Bremer H, Ewert R, Bonderman D, Rosenkranz S, et al. (2014): Gezielte Therapie der pulmonal arteriellen Hypertonie (PAH). Deutsche medizinische Wochenschrift (1946); 139 Suppl 4:S142-50.
62. Taichman DB, Ornelas J, Chung L, Klinger JR, Lewis S, Mandel J, et al. (2014): Pharmacologic therapy for pulmonary arterial hypertension in adults: CHEST guideline and expert panel report. Chest; 146(2):449-75.
63. Gemeinsamer Bundesausschuss (2014): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO - Wirkstoff: Macitentan. [Zugriff: 26.06.2015]. URL: https://www.g-ba.de/downloads/92-975-449/Nutzenbewertung%20G-BA.pdf.
64. Rubenfire M, Lippo G, Bodini BD, Blasi F, Allegra L, Bossone E (2009): Evaluating health-related quality of life, work ability, and disability in pulmonary arterial hypertension: an unmet need. Chest; 136(2):597-603.
65. Chen H, Taichman DB, Doyle RL (2008): Health-related quality of life and patient-reported outcomes in pulmonary arterial hypertension. Proceedings of the American Thoracic Society; 5(5):623-30.
66. Cenedese E, Speich R, Dorschner L, Ulrich S, Maggiorini M, Jenni R, et al. (2006): Measurement of quality of life in pulmonary hypertension and its significance. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 28(4):808-15.
67. Zlupko M, Harhay MO, Gallop R, Shin J, Archer-Chicko C, Patel R, et al. (2008): Evaluation of disease-specific health-related quality of life in patients with pulmonary arterial hypertension. Respiratory medicine; 102(10):1431-8.
68. McKenna SP, Doughty N, Meads DM, Doward LC, Pepke-Zaba J (2006): The Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR): a measure of health-related quality of life and quality of life for patients with pulmonary hypertension. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation; 15(1):103-15.
69. Gomberg-Maitland M, Thenappan T, Rizvi K, Chandra S, Meads DM, McKenna SP (2008): United States validation of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR). The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation; 27(1):124-30.
70. Galen Research (2015): Measures Database. [Zugriff: 26.06.2015]. URL: http://www.galen-research.com/measures-database/.
71. Cima K, Twiss J, Speich R, McKenna SP, Grunig E, Kahler CM, et al. (2012): The German adaptation of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR). Health and quality of life outcomes; 10:110.
72. Actelion Pharmaceuticals Ltd (2014): Common Technical Document - Abschnitt - 2.7.4 - Summary of Clinical Safety for Selexipag (ACT-293987) in Pulmonary Arterial Hypertension. VERTRAULICH.
73. European Medicines Agency (2016): European Public Assessment Report - Uptravi. International non-propietary name: Selexipag. Procedure No. EMEA/H/C/003774/0000. [Zugriff: 30.05.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/003774/WC500207175.pdf.
74. European Medicines Agency (2014): Guideline on the investigation of subgroups in confirmatory clinical trials (Entwurf). [Zugriff: 29.10.2015]. URL:

http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500160523.pdf.

75. Medizinischer Dienst der Spitzenverbände der Krankenkassen e.V. (2004): G-2 Gutachten: Aussagekraft von Subgruppenanalysen. [Zugriff: 29.10.2015]. URL: https://www.mds-ev.de/fileadmin/dokumente/Publikationen/GKV/Begutachtungsgrundlagen_GKV_von_WM/subgruppen-gutachten.pdf.
76. Actelion Pharmaceuticals Ltd (2009): AC-065A302/GRIPHON - Selexipag (ACT-293987) in Pulmonary Arterial Hypertension (NCT01106014). Stand des Eintrags: 16.03.2016. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT01106014
77. Actelion Pharmaceuticals Ltd (2010): AC-065A302/GRIPHON - A multicenter, double-blind, placebo-controlled Phase 3 study assessing the safety and efficacy of selexipag on morbidity and mortality in patients with pulmonary arterial hypertension (2009-014490-41). [Zugriff: 09.05.2016]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-014490-41
78. Actelion Pharmaceuticals Ltd (2014): AC-065A302/GRIPHON: Prostacyclin (PGI 2) receptor agonist in pulmonary arterial hypertension - A multicenter, double-blind, placebo-controlled Phase 3 study assessing the efficacy and safety of selexipag (ACT-293987) on morbidity and mortality in patients with pulmonary arterial hypertension (PAH) (including data from AC-065A303/GRIPHON OL up to 10 March 2014). Studienbericht. VERTRAULICH.
79. Actelion Pharmaceuticals Deutschland GmbH (2016): GRIPHON-Studie - Zusätzliche für das Nutzendossier durchgeführte Post-Hoc-Analysen. VERTRAULICH.
80. Sitbon O, Channick R, Chin KM, Frey A, Gaine S, Galie N, et al. (2015): Selexipag for the Treatment of Pulmonary Arterial Hypertension. The New England journal of medicine; 373(26):2522-33.
81. Actelion Pharmaceuticals Deutschland GmbH (2015): GRIPHON-Studie - Auswertung der PAH-spezifischen Begleitmedikation zum Zeitpunkt des Studieneinschlusses. VERTRAULICH.
82. Huscher D, Hoeper M, Pittrow D (2015): Prävalenz und Inzidenz der PAH - Auswertung der COMPERA-Daten vom 01.11.2015.
83. Higgins J, Green S (2011): Cochrane Handbook for Systematic Reviews of Interventions, Section 9.2.6 - Effect measures for time-to-event (survival) outcomes. [Zugriff: 05.09.2013]. URL: http://handbook.cochrane.org/.
84. Statfinn (2015): Implementation of the Marginal Structural Cox Proportional Hazards Model (MSM) using the Inverse Probability of Censoring Weights for Probability of Switching Treatment and Censoring. An application to the Clinical Investigations to Evaluate Efficacy of Selexipag study. VERTRAULICH.
85. Statfinn (2015): Implementation of the Rank Preserving Structural Accelerated Failure Time (RPSFT) Model in the Clinical Investigations to Evaluate Efficacy of Selexipag study. VERTRAULICH.
86. Peacock AJ, Naeije R, Galie N, Rubin L (2009): End-points and clinical trial design in pulmonary arterial hypertension: have we made progress? The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 34(1):231-42.

87. Rubin LJ, Badesch DB, Barst RJ, Galie N, Black CM, Keogh A, et al. (2002): Bosentan therapy for pulmonary arterial hypertension. The New England journal of medicine; 346(12):896-903.
88. Galiè N, Rubin LJ, Hoeper MM, Jansa P, Al-Hiti H, Meyer GMB, et al. (2008): Treatment of patients with mildly symptomatic pulmonary arterial hypertension with bosentan (EARLY study): a double-blind, randomised controlled trial. The Lancet; 371(9630):2093-100.
89. Howard LS, Ferrari P, Mehta S (2014): Physicians' and patients' expectations of therapies for pulmonary arterial hypertension: where do they meet? Eur Respir Rev; 23(134):458-68.
90. Actelion Pharmaceuticals Deutschland GmbH (2016): Berechnung von Wahrscheinlichkeiten falsch positiver Ergebnisse im Zusammenhang mit Subgruppenanalysen.
91. Jones RL, Qian YM, Wise H, Wong HN, Lam WL, Chan HW, et al. (1997): Relaxant actions of nonprostanoid prostacyclin mimetics on human pulmonary artery. Journal of cardiovascular pharmacology; 29(4):525-35.
92. Morrison K, Ernst R, Hess P, Studer R, Clozel M (2010): Selexipag: a selective prostacyclin receptor agonist that does not affect rat gastric function. The Journal of pharmacology and experimental therapeutics; 335(1):249-55.
93. Wong SS, Wilczynski NL, Haynes RB (2006): Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. Journal of the Medical Library Association : JMLA; 94(4):451-5.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50") und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[46]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

[46] Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)" sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-130 (Anhang): Direkter Vergleich - Dokumentation der Recherche in MEDLINE

Datenbankname	Ovid Medline® In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present Ovid Medline® Daily Update May 03, 2016
Suchoberfläche	Ovid
Datum der Suche	04.05.2016
Zeitsegment	Keine zeitliche Einschränkung
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [93]

#	Suchbegriffe	Ergebnis
1	randomized controlled trial.pt.	415.006
2	randomized.mp.	642.918
3	placebo.mp.	175.036
4	1 OR 2 OR 3	696.042
5	selexipag*.mp.	32
6	(ns 304 or ns-304 or ns304).mp.	7
7	(act 293987 or act-293987 or act293987).mp.	1
8	uptravi.mp.	3
9	5 OR 6 OR 7 OR 8	35
10	4 AND 9	12

Tabelle 4-131 (Anhang): Direkter Vergleich - Dokumentation der Recherche in EMBASE

Datenbankname	EMBASE 1974 to 2016 May 03
Suchoberfläche	Ovid
Datum der Suche	04.05.2016
Zeitsegment	Keine zeitliche Einschränkung
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [93]

#	Suchbegriffe	Ergebnis
1	(random* or double-blind*).tw.	1.126.320
2	placebo*.mp.	373.932
3	1 OR 2	1.310.575
4	exp selexipag/	145
5	selexipag*.mp.	151
6	(ns 304 OR ns-304 OR ns304).mp.	21
7	(act 293987 OR act-293987 OR act293987).mp.	13
8	uptravi.mp.	6
9	4 OR 5 OR 6 OR 7 OR 8	168
10	3 AND 9	72

Tabelle 4-132 (Anhang): Direkter Vergleich - Dokumentation der Recherche bei „The Cochrane Library“

Datenbankname	The Cochrane Library
Suchoberfläche	The Cochrane Library
Datum der Suche	04.05.2016
Zeitsegment	Keine zeitliche Einschränkung
Suchfilter	Nur klinische Studien [Trials] in 5

#	Suchbegriffe	Ergebnis
1	selexipag*	13
2	"ns 304" OR ns-304 OR ns304	0
3	"act 293987" OR act-293987 OR act293987	3
4	uptravi	0
5	1 OR 2 OR 3 OR 4	15
6	Limit: In Trials	15

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. clinicaltrials.gov), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-133 (Anhang): Dokumentation der Recherche in clinicaltrials.gov

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	09.05.2016
Suchstrategie	selexipag OR uptravi[47]
Treffer	9 Studien (davon 1 Studie relevant)

Tabelle 4-134 (Anhang): Dokumentation der Recherche in EU-CTR

Studienregister	EU-CTR
Internetadresse	http:\\www.clinicaltrialsregister.eu
Datum der Suche	09.05.2016
Suchstrategie	selexipag* OR ns-304 OR ns304 OR (ns 304) OR act-293987 OR act293987 OR (act 293987) OR uptravi
Treffer	6 Studien (davon 1 Studie relevant)

[47] Die Suchbegriffe „ns-304“, „ns304“, „ns 304“, „act-293987“, „act293987“ und „act 293987“ sind bereits durch den Suchbegriff „selexipag“ abgedeckt.

Tabelle 4-135 (Anhang): Dokumentation der Recherche in WHO ICTRP

Studienregister	WHO ICTRP
Internetadresse	http://apps.who.int/trialsearch
Datum der Suche	09.05.2016
Suchstrategie	selexipag OR ns-304 OR ns304 OR ns 304 OR act-293987 OR act293987 OR act 293987 OR uptravi
Treffer	63 Treffer für 19 Studien (davon 2 Studien relevant; beide Einträge sind jedoch Verweise auf Clinicaltrials.gov- bzw. EU-CTR-Einträge)

Tabelle 4-136 (Anhang): Dokumentation der Recherche in PharmNet.Bund

Studienregister	PharmNet.Bund
Internetadresse	http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm
Datum der Suche	09.05.2016
Suchstrategie	selexipag (in "active substance") ODER ns-304 (in "active substance") ODER ns304 (in "active substance") ODER ns 304 (in "active substance") ODER act-293987 (in "active substance") ODER act293987 (in "active substance") ODER act 293987 (in "active substance") ODER uptravi (in "product name/code") ODER uptravi (in "title")
Treffer	5 Studien (davon 1 Studie relevant; der Eintrag ist jedoch ein Verweis auf einen EU-CTR-Eintrag)

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-137 (Anhang): Liste der im Volltext gesichteten und ausgeschlossenen Dokumente (bibliografische Literaturrecherche)

Zitat	Titel	Ausschluss-grund
[1]	Kovacs G, Olschewski H (2012): [Pulmonary hypertension]. (Pulmonale Hypertonie).Dtsch Med Wochenschr; 137(40):2026-8.	A5
[2]	Simonneau G, Torbicki A, Hoeper MM, Delcroix M, Karlocai K, Galie N, et al. (2012): Selexipag: an oral, selective prostacyclin receptor agonist for the treatment of pulmonary arterial hypertension. Eur Respir J; 40(4):874-80.	A6
[3]	Asaki T, Kuwano K, Morrison K, Gatfield J, Hamamoto T, Clozel M (2015): Selexipag: An Oral and Selective IP Prostacyclin Receptor Agonist for the Treatment of Pulmonary Arterial Hypertension. J Med Chem; 58(18):7128-37.	A5

A1 - Population nicht wie definiert
A2 - Prüfintervention nicht wie definiert
A3 - Vergleichsintervention nicht wie definiert
A4 - kein patientenrelevanter Endpunkt berichtet
A5 - keine Studie im RCT-Design
A6 - Studiendauer kürzer als 6 Monate
A7 - keine Volltextpublikation verfügbar
A8 - Publikation in anderer Sprache als deutsch oder englisch
D - Dublette

Referenzliste Anhang 4-C1

1. Kovacs G, Olschewski H (2012): [Pulmonary hypertension]. (Pulmonale Hypertonie).Dtsch Med Wochenschr; 137(40):2026-8.
2. Simonneau G, Torbicki A, Hoeper MM, Delcroix M, Karlocai K, Galie N, et al. (2012): Selexipag: an oral, selective prostacyclin receptor agonist for the treatment of pulmonary arterial hypertension. Eur Respir J; 40(4):874-80.
3. Asaki T, Kuwano K, Morrison K, Gatfield J, Hamamoto T, Clozel M (2015): Selexipag: An Oral and Selective IP Prostacyclin Receptor Agonist for the Treatment of Pulmonary Arterial Hypertension. J Med Chem; 58(18):7128-37.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)

Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-138 (Anhang): Suche nach RCT - Liste der ausgeschlossenen Studien (Suche in Studienregistern)

Zitat	Studientitel	Registernummer	Ausschluss-grund
Clinicaltrials.gov			
[1]	Study in Healthy Male Subjects to Demonstrate Bioequivalence of 1600 µg Selexipag Administered as Eight Tablets of 200 µg or as Single Tablet of 1600 µg	NCT02206295	A1
[2]	Study of Selexipag and Its Metabolite ACT-333679 on Cardiac Repolarization in Healthy Male and Female Subjects	NCT02206204	A1
[3]	Study of ACT-293987 (NS-304) in Pulmonary Arterial Hypertension (PAH)	NCT00993408	A5
[4]	Raynaud's Phenomenon Secondary to Systemic Sclerosis	NCT02260557	A1
[5]	Study to Assess the Tolerability and the Safety of the Transition From Inhaled Treprostinil to Oral Selexipag in Patients With Pulmonary Arterial Hypertension	NCT02471183	A5
[6]	ACT-293987 in Pulmonary Arterial Hypertension	NCT01112306	A5
[7]	The Efficacy and Safety of Initial Triple Versus Initial Dual Oral Combination Therapy in Patients With Newly Diagnosed Pulmonary Arterial Hypertension	NCT02558231	A7[1]
[16]	Comparison of Two Dose Strengths of Selexipag in Healthy Adults	NCT02745860	A1
EU-CTR			
[8]	A multi-centre, multinational, open- label study to evaluate the long-term safety, tolerability and efficacy of selexipag / ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over (open-extension study to NS-304/-02)	2007-006453-12	A5
[9]	A multi-centre, multinational, open-label single-dose acute hemodynamic study followed by a multi-centre, multinational, randomized, double-blind, parallel-group, placebo controlled study to assess the safety, tolerability, pharmacokinetics and preliminary efficacy (proof-of-concept) of ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over	2007-003328-39	A6
[10]	Long-term single-arm open-label study, to assess the safety and tolerability of selexipag in patients with pulmonary arterial hypertension	2009-014992-31	A5
[11]	A multi-center, double-blind, randomized, placebo-controlled, parallel group, exploratory Phase 2 study to assess efficacy and safety of selexipag in adult subjects with Raynaud's Phenomenon secondary to Systemic Sclerosis	2014-000865-34	A1
[17]	The efficacy and safety of initial triple versus initial dual oral combination therapy in patients with newly diagnosed pulmonary arterial hypertension: A multi-center, double-blind, placebo-controlled, Phase 3b study	2015-003438-28	A7

WHO ICTRP			
[12]	An exploratory study to assess the efficacy and safety of NS-304 in patients with chronic thromboembolic pulmonary hypertension (CTEPH)	JPRN-JapicCTI-111667	A1
[8]	A multi-centre, multinational, open- label study to evaluate the long-term safety, tolerability and efficacy of ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over (open-extension study to NS-304/-02)	EUCTR2007-006453-12-HU	D/A5
[9]	A multi-centre, multinational, open-label single-dose acute hemodynamic study followed by a multi-centre, multinational, randomized, double-blind, parallel-group, placebo controlled study to assess the safety, tolerability, pharmacokinetics and preliminary efficacy (proof-of-concept) of ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over	EUCTR2007-003328-39-BE	D/A6
[13]	The clinical study to assess the efficacy, safety and pharmacokinetics of NS-304 in patients with pulmonary arterial hypertension (PAH)	JPRN-JapicCTI-111532	A5
[11]	A multi-center, double-blind, randomized, placebo-controlled, parallel group, exploratory Phase 2 study to assess efficacy and safety of selexipag in adult subjects with Raynaud's Phenomenon secondary to Systemic Sclerosis	EUCTR2014-000865-34-DE	D/A1
[14]	The open-label extension study of NS-304 in patients with chronic thromboembolic pulmonary hypertension (CTEPH) who completed the exploratory evaluation study	JPRN-JapicCTI-111666	A1
[15]	The open-label extension study of NS-304 in patients with chronic thromboembolic pulmonary hypertension (CTEPH)	JPRN-JapicCTI-132159	A1
[3]	A Multi-centre, Multinational, Open-label, Single-dose Acute Hemodynamic Study Followed by Multi-centre, Multinational, Randomized, Double-blind, Parallel-group, Placebo Controlled Study to Assess the Safety, Tolerability, Pharmacokinetics and Preliminary Efficacy (Proof-of-concept) of ACT-293987 (NS-304) in the Treatment of Pulmonary Arterial Hypertension in Subjects Aged 18 Years and Over	NCT00993408	D/A6
-	A multicenter, double-blind, placebo-controlled Phase 3 study assessing the safety and efficacy of selexipag on morbidity and mortality in patients with pulmonary arterial hypertension - GRIPHON	EUCTR2009-014490-41-DK	D
[2]	Single-center, Double-blind, Randomized, Placebo- and Positive-controlled, Parallel-group With Nested Cross-over, Multiple-dose, Up-titration Study of the Effects of Selexipag and Its Metabolite ACT-333679 on Cardiac Repolarization in Healthy Male and Female Subjects	NCT02206204	D/A1
[1]	A Single-center, Open-label, Randomized, Two-period, Two-treatment, Crossover Study in Healthy Male Subjects to Demonstrate Bioequivalence of 1600 µg Selexipag Administered as Eight Tablets of 200 µg (Reference Drug) or as Single Tablet of 1600 µg (Test Drug)	NCT02206295	D/A1

-	A Multicenter, Double-blind, Placebo-controlled Phase 3 Study Assessing the Safety and Efficacy of Selexipag on Morbidity and Mortality in Patients With Pulmonary Arterial Hypertension	NCT01106014	D
[6]	Long-term Single-arm Open-label Study, to Assess the Safety and Tolerability of Selexipag (ACT-293987) in Patients With Pulmonary Arterial Hypertension	NCT01112306	D/A5
[10]	Long-term single-arm open-label study, to assess the safety and tolerability of selexipag in patients with pulmonary arterial hypertension - GRIPHON OL	EUCTR2009-014992-31-BE	D/A5
[4]	A Multi-center, Double-blind, Randomized, Placebo-controlled, Parallel Group, Exploratory Phase 2 Study to Assess Efficacy and Safety of Selexipag in Adult Subjects With Raynaud's Phenomenon Secondary to Systemic Sclerosis	NCT02260557	D/A1
[5]	Multicenter, Open-label, Single-group Study to Assess the Tolerability and the Safety of the Transition From Inhaled Treprostinil to Oral Selexipag in Adult Patients With Pulmonary Arterial Hypertension	NCT02471183	D/A5
[7]	AC-065A308 - The Efficacy and Safety of Initial Triple Versus Initial Dual Oral Combination Therapy in Patients With Newly Diagnosed Pulmonary Arterial Hypertension: A Multi-center, Double-blind, Placebo-controlled, Phase 3b Study	NCT02558231	D/A7
[16]	Comparison of Two Dose Strengths of Selexipag in Healthy Adults	NCT02745860	D/A1
[17]	The efficacy and safety of initial triple versus initial dual oral combination therapy in patients with newly diagnosed pulmonary arterial hypertension: A multi-center, double-blind, placebo-controlled, Phase 3b study	2015-003438-28	D/A7
PharmNet.Bund			
[11]	A multi-center, double-blind, randomized, placebo-controlled, parallel group, exploratory Phase 2 study to assess efficacy and safety of selexipag in adult subjects with Raynaud's Phenomenon secondary to Systemic Sclerosis	2014-000865-34	D/A1
[10]	Long-term single-arm open-label study, to assess the safety and tolerability of selexipag in patients with pulmonary arterial hypertension	2009-014992-31	D/A5
[8]	A multi-centre, multinational, open- label study to evaluate the long-term safety, tolerability and efficacy of selexipag / ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over (open-extension study to NS-304/-02)	2007-006453-12	D/A5

-	A multicenter, double-blind, placebo-controlled Phase 3 study assessingthe safety and efficacy of selexipag on morbidity and mortality in patientswith pulmonary arterial hypertension	2009-014490-41	D
[17]	The efficacy and safety of initial triple versus initial dual oral combination therapy in patients with newly diagnosed pulmonary arterial hypertension: A multi-center, double-blind, placebo-controlled, Phase 3b study	2015-003438-28	D/A7

[1]) Studie noch nicht abgeschlossen, es liegen noch keine Studienergebnisse vor.
A1 - Population nicht wie definiert
A2 - Prüfintervention nicht wie definiert
A3 - Vergleichsintervention nicht wie definiert
A4 - keine Studie im RCT-Design
A5 - kein patientenrelevanter Endpunkt berichtet
A6 - keine Volltextpublikation verfügbar
A7 - Publikation in anderer Sprache als deutsch oder englisch
D - Dublette
Bei Studien ohne Zitatangabe handelt es sich um Dubletten relevanter Studien.

Referenzliste Anhang 4-D1

1. Actelion (2012): AC-065-108 - Study in Healthy Male Subjects to Demonstrate Bioequivalence of 1600 µg Selexipag Administered as Eight Tablets of 200 µg or as Single Tablet of 1600 µg (NCT02206295). Stand des Eintrags: 30.07.2014. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT02206295
2. Actelion (2012): AC-065-106 - Study of Selexipag and Its Metabolite ACT-333679 on Cardiac Repolarization in Healthy Male and Female Subjects (NCT02206204). Stand des Eintrags: 30.07.2014. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT02206204
3. Actelion (2008): NS-304/-02 - Study of ACT-293987 (NS-304) in Pulmonary Arterial Hypertension (PAH) (NCT00993408). Stand des Eintrags: 09.10.2009. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT00993408
4. Actelion (2014): AC-065C202 - Raynaud's Phenomenon Secondary to Systemic Sclerosis (NCT02260557). Stand des Eintrags: 21.05.2015. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT02260557
5. Actelion (2015): AC-065A304 - Study to Assess the Tolerability and the Safety of the Transition From Inhaled Treprostinil to Oral Selexipag in Patients With Pulmonary Arterial Hypertension (NCT02471183). Stand des Eintrags: 14.12.2015. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT02471183
6. Actelion (2009): AC-065A303 - ACT-293987 in Pulmonary Arterial Hypertension (NCT01112306). Stand des Eintrags: 13.01.2016. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT01112306
7. Actelion (2016): AC-065A308 - The Efficacy and Safety of Initial Triple Versus Initial Dual Oral Combination Therapy in Patients With Newly Diagnosed Pulmonary Arterial Hypertension (NCT02558231). Stand des Eintrags: 23.09.2015. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT02558231
8. Actelion Pharmaceuticals Ltd (2008): NS-304-03 - A multi-centre, multinational, open- label study to evaluate the long-term safety, tolerability and efficacy of

selexipag / ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over (open-extension study to NS-304/-02) (2007-006453-12). [Zugriff: 09.05.2016]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2007-006453-12

9. Actelion Pharmaceuticals Ltd (2008): NS-304-02 - A multi-centre, multinational, open-label single-dose acute hemodynamic study followed by a multi-centre, multinational, randomized, double-blind, parallel-group, placebo controlled study to assess the safety, tolerability, pharmacokinetics and preliminary efficacy (proof-of-concept) of ACT-293987 (NS-304) in the treatment of pulmonary arterial hypertension in subjects aged 18 years and over (2007-003328-39). [Zugriff: 09.05.2016]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2007-003328-39
10. Actelion Pharmaceuticals L (2010): AC-065A303 - Long-term single-arm open-label study, to assess the safety and tolerability of selexipag in patients with pulmonary arterial hypertension (2009-014992-31). [Zugriff: 09.05.2016]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-014992-31
11. Actelion Pharmaceuticals L (2014): AC-065C202 - A multi-center, double-blind, randomized, placebo-controlled, parallel group, exploratory Phase 2 study to assess efficacy and safety of selexipag in adult subjects with Raynaud's Phenomenon secondary to Systemic Sclerosis (2014-000865-34). [Zugriff: 09.05.2016]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-000865-34
12. Nippon Shinyaku Co. Ltd., Actelion Pharmaceuticals Japan Ltd. (2011): An exploratory study to assess the efficacy and safety of NS-304 in patients with chronic thromboembolic pulmonary hypertension (CTEPH) (JPRN-JapicCTI-111667). Stand des Eintrags: 07.12.2015. [Zugriff: 09.05.2016]. URL: http://www.clinicaltrials.jp/user/showCteDetailE.jsp?japicId=JapicCTI-111667
13. Nippon Shinyaku Co. Ltd., Actelion Pharmaceuticals Japan Ltd. (2011): The clinical study to assess the efficacy, safety and pharmacokinetics of NS-304 in patients with pulmonary arterial hypertension (PAH) (JPRN-JapicCTI-111532). Stand des Eintrags: 07.12.2015. [Zugriff: 09.05.2016]. URL: http://www.clinicaltrials.jp/user/showCteDetailE.jsp?japicId=JapicCTI-111532
14. Nippon Shinyaku Co. Ltd., Actelion Pharmaceuticals Japan Ltd. (2011): The open-label extension study of NS-304 in patients with chronic thromboembolic pulmonary hypertension (CTEPH) who completed the exploratory evaluation study (JPRN-JapicCTI-111666). Stand des Eintrags: 07.12.2015. [Zugriff: 09.05.2016]. URL: http://www.clinicaltrials.jp/user/showCteDetailE.jsp?japicId=JapicCTI-111666
15. Nippon Shinyaku Co. Ltd. (2013): The open-label extension study of NS-304 in patients with chronic thromboembolic pulmonary hypertension (CTEPH) (JPRN-JapicCTI-132159). Stand des Eintrags: 07.12.2015. [Zugriff: 09.05.2016]. URL: http://www.clinicaltrials.jp/user/showCteDetailE.jsp?japicId=JapicCTI-132159
16. Actelion (2016): AC-065-112 - Comparison of Two Dose Strengths of Selexipag in Healthy Adults (NCT02745860). Stand des Eintrags: 20.04.2016. [Zugriff: 09.05.2016]. URL: https://ClinicalTrials.gov/show/NCT02745860
17. Actelion Pharmaceuticals L (2016): AC-065A308/TRITON - The efficacy and safety of initial triple versus initial dual oral combination therapy in patients with newly diagnosed pulmonary arterial hypertension: A multi-center, double-blind, placebo-

controlled, Phase 3b study (2015-003438-28). [Zugriff: 09.05.2016].
URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003438-28

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-139 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-139 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-139 (Anhang): Studiendesign und -methodik für Studie AC-065A302 – GRIPHON

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primäres Ziel der Studie war der Nachweis der Wirksamkeit von Selexipag in Bezug auf die Zeit bis zum ersten Morbiditäts- und/oder Mortalitätsereignis (M/M-Ereignis) bei Patienten mit PAH. Sekundäre Ziele waren: • die Untersuchung der Wirksamkeit in Bezug auf die körperliche Leistungsfähigkeit und auf andere sekundäre und exploratorische Wirksamkeitsendpunkte bei Patienten mit PAH. • die Untersuchung von Sicherheit und Verträglichkeit von Selexipag bei Patienten mit PAH.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Bei GRIPHON handelt es sich um eine multizentrische, randomisierte, doppelblinde, placebokontrollierte und ereignisgetriebene Phase-3-Studie im Parallelgruppendesign zum Vergleich der Wirkung (Wirksamkeit, Sicherheit, Verträglichkeit, Pharmakokinetik, Pharmakodynamik) von Selexipag (zweimal täglich verabreicht in individuellen Dosierungen von 200-1.600 µg nach Auftitration) gegenüber Placebo bei Patienten mit symptomatischer PAH. Die Studie bestand aus den folgenden Phasen: • Screening-Phase: Screening-Untersuchungen wurden bis höchstens 28 Tage vor der Basiswert-Untersuchung durchgeführt. • Behandlungsphase: vom Tag der Basiswertuntersuchung (Untersuchung 1) bis zur Untersuchung am Ende der Studie (EOS-Untersuchung). • Die Patienten wurden instruiert am Abend des ersten Untersuchungstermins (Untersuchung 1) mit der Einnahme der ersten Studienmedikation (200 µg Selexipag oder Placebo) zu beginnen. • Die Studienmedikation sollte in wöchentlichen Schritten um zweimal täglich 200 µg so lange gesteigert werden, bis die individuell maximal tolerierte Dosis des jeweiligen Patienten oder die Maximaldosis von 1.600 µg in Woche 12 erreicht wurde. • Die Behandlungsphase wurde mit der EOS-Untersuchung zum Zeitpunkt der Bekanntgabe des Studienendes (d. h. sobald die Gesamtzahl von 331 CEC-bestätigten M/M-Ereignisse, deren Beginn innerhalb von 7 Tagen nach der letzten Einnahme der Studienmedikation lag, auftraten) abgeschlossen. Die EOS-Untersuchung sollte innerhalb von 4 Wochen nach Bekanntgabe des Studienendes durchgeführt werden. Bei Patienten, bei denen ein CEC-bestätigtes M/M-Ereignis auftrat oder die die Behandlung mit der Studienmedikation vor dem Studienende abgebrochen hatten, wurde die EOS-Untersuchung nach dem Auftreten des Morbiditätsereignisses

Item[a]	Charakteristikum	Studieninformation
		oder nach dem vorzeitigen Behandlungsabbruch durchgeführt. • Beobachtungsphase nach der Behandlungsphase (*post-treatment observation period*, PTOP): Patienten, die (unabhängig vom Auftreten eines M/M-Ereignisses) die Behandlung vor Bekanntgabe des Studienendes beendeten, hatten die Möglichkeit an einer Beobachtungsphase nach der Behandlungsphase teilzunehmen, um zusätzliche klinische Daten zu erfassen. Diese Phase war definiert als der Zeitraum nach der EOS-Untersuchung aufgrund des vorzeitigen Behandlungsabbruches bis zur Abschlussbehandlung der PTOP im Anschluss an die Bekanntgabe des Studienendes durch Actelion. • Sicherheitsnachbeobachtung nach der Behandlungsphase zur Erkennung von SUE: bei den Patienten wurde jegliches Auftreten von SUE bis 30 Tage nach der letzten Einnahme der Studienmedikation oder bis zur ersten Anwendung von Selexipag in der Studie AC-065A303 (je nachdem, was zuerst auftrat) dokumentiert. • Zum Zeitpunkt des Studienendes wurden sämtliche Patienten (mit Ausnahme derer, die ihre Zustimmung zu allen Studienkomponenten zurückgezogen hatten) kontaktiert, um ihren Vitalstatus zu erheben (telefonische Nachverfolgung des Überlebens). • Behandlungsextensionsphase: Patienten, die zum Zeitpunkt des Studienendes noch ihre Medikation aus der Studie AC-065A302 erhalten hatten, wurde eine Behandlungsextensions-phase angeboten. Die Extensionsphase war geplant über einen zeitlichen Verlauf von bis zu 3 Monaten von der EOS-Untersuchung bis zur Untersuchung am Ende der Extensions-phase anschließend an die Entblindung der AC-065A302-Datenbank. Während dieser Phase erhielten die Patienten weiterhin die AC-065A302 Studienmedikation in doppel-verblindeter Weise. Patienten, die nach Auftreten eines Morbiditätsereignisses die EOS-Untersuchung absolviert hatten, waren anschließend berechtigt an der Extensionsstudie AC-065A303 teilzunehmen, sofern diese von der zuständigen nationalen Gesundheitsbehörde im jeweiligen Land genehmigt worden war.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Insgesamt wurden im Verlauf der Studie 6 globale Änderungen des Studienprotokolls vorgenommen. Die folgende Auflistung fasst die relevanten Änderungen der Studienmethodik, die mit diesen 6 Amendments eingeführt wurden, zusammen. <u>Global Amendment 1 vom 11.03.2010</u> • Das Amendment wurde aufgrund der Empfehlungen der FDA aus einem Special Protocol Assessment durchgeführt. • Da die Zeit bis zur klinischen Verschlechterung im Vergleich zur Veränderung der 6-MWD gegenüber dem Basiswert als klinisch relevanterer Endpunkt für PAH-Patienten angesehen wurde der ko-primäre Endpunkt „Zeit bis zur ersten klinischen Verschlechterung“ (ko-primärer Endpunkt der identischen Studien AC-065A301 (nicht initiiert) und AC-065A302) als primärer Endpunkt beibehalten. Patienten ohne Ereignis der klinischen Verschlechterung sollten in der Analyse der Zeit bis

Item[a]	Charakteristikum	Studieninformation

		zur klinischen Verschlechterung zudem 7 Tage nach Beendigung der Behandlung zensiert werden.
		• Der vormals primäre Endpunkt „Veränderung der 6-MWD vom Basiswert bis Woche 16“ wurde als sekundärer Endpunkt klassifiziert und sollte nun zu Woche 26 statt zu Woche 16 analysiert werden. Ebenso sollten auch die anderen sekundären Endpunkte zu Woche 26 statt zu Woche 16 untersucht werden. Die Imputationsmethoden wurden entsprechend angepasst.
		• Als Konsequenz der oben genannten Modifikationen wurden die beiden Studien AC-065A301 und AC-065A302 zu einer Studie zusammengelegt.
		• Für den primären Endpunkt „Zeit bis zur ersten klinischen Verschlechterung bis 7 Tage nach der letzten Einnahme der Studienmedikation“ mussten nun statt 203 Ereignissen 202 Ereignisse auftreten, um mit einer Power von 90 % eine Hazard Ratio von 0,5729 bei einem zweiseitigen Signifikanzlevel von 0,01 mithilfe des Log-Rang-Tests zu beobachten. Bei einer Steigerung der Stichprobengröße von 544 auf 670 Patienten wurde davon ausgegangen, dass 202 Ereignisse maximal in einer Zeit von 3,5 Jahren beobachtet werden könnten.
		• Nach Feststellung der maximal tolerierten Dosis des Patienten in Woche 12, sollte diese für die kommenden 14 Wochen, d. h. bis Woche 26 (Analysezeitpunkt der sekundären Wirksamkeitsparameter) stabil gehalten werden.
		• Die Veränderung der 6-MWD, des Borg-Dyspnoe-Index sowie die ausbleibende Verschlechterung der WHO-/NYHA-Klasse vom Basiswert bis Woche 16 wurden in die Liste explorativer Endpunkte aufgenommen.
		• Der Zeitpunkt, bis zu dem UE oder SUE aufgenommen werden sollte, wurde von 3 Tagen nach Beendigung der Behandlung mit der Studienmedikation auf 7 Tage verlängert. Als Zeitpunkt der letzten Einnahme der Studienmedikation galt das Datum der tatsächlichen letzten Einnahme, so wie sie im CRF vermerkt war.
		• Das Einschlusskriterium zum 6-MWT wurde dahin gehend abgeändert, dass Patienten nur aufgrund des Screening-Tests und nicht auf der Testwiederholung zum Zeitpunkt der Basiswerterhebung ausgeschlossen werden konnten. Zudem sollten alle 6-MWT-Erhebungen (außer zum Screening und zum Zeitpunkt der Basiswerterhebung) im Talspiegel erhoben werden.
		• Die folgenden Ein-/Ausschlusskriterien wurden näher erläutert oder abgeändert (nur relevante Änderungen der Methodik sind dargestellt):
		  ○ Um das Risiko zu minimieren, dass Patienten mit anderen PH-Formen als PAH eingeschlossen wurden, wurde eine obere Altersgrenze von 75 Jahren definiert und die Untergrenze des pulmonalen Gefäßwiderstandes in Ruhe von 240 auf 400 dyn s cm^{-5} heraufgesetzt.
		  ○ Da in vorangegangen Phase-1- und -2a-Studien kein Effekt auf den Blutdruck festgestellt wurde, wurde

Item[a]	Charakteristikum	Studieninformation
		der Ausschluss hypotonischer Patienten (systemischer systolischer Blutdruck <100 mmHg zum Zeitpunkt des Screenings) als klinisch nicht mehr relevant betrachtet und daher nicht mehr vorgenommen. ○ In Übereinstimmung mit der Definition wurde das Ausschlusskriterium bezüglich moderater/schwerer obstruktiver Lungenerkrankungen von FEV[1]/FVC <70 % oder FEV_1 <65 % des prognostizierten Werts nach Bronchodilator-Anwendung zu FEV_1/FVC <70 % und FEV_1 <65 % nach Bronchodilator-Anwendung. ○ Das Ausschlusskriterium bezüglich der restriktiven Lungenerkrankungen wurde dahin gehend präzisiert, dass Patienten mit restriktiver Lungenerkrankung ausgeschlossen wurden. Die Obergrenze der gesamten Lungenkapazität wurde von <60 % des prognostizierten Wertes auf <70 % angehoben. ○ Als zusätzliches Ausschlusskriterium wurde aufgenommen, dass Patienten in einem Reha-Programm auf Grundlage von körperlichem Training von der Studie ausgeschlossen sein sollten. • Im Zusammenhang mit erlaubter PAH-Begleittherapie wurde der Zeitraum, in der die Dosis stabil gehalten werden sollte von „ein Monat vor der Basiswertuntersuchung bis Woche 16“ zu „drei Monate vor der Basiswertuntersuchung bis Woche 26“ geändert. <u>Global Amendment 2 vom 20.12.2010</u> • Das Einschlusskriterium bezüglich Sonneneinstrahlung wurde entfernt, da die Ergebnisse einer Phase-1-Studie (AC-065-102) nicht auf einen relevantes phototoxisches Potential von Selexipag in Dosen von zweimal täglich 800 bis 1.200 µg nach UV-A- oder UV-B-Bestrahlung gesunder männlicher Probanden hinwiesen. <u>Global Amendment 3 vom 11.05.2011</u> • Die maximal tolerierte Dosis sollte als kategorielle und nicht als kontinuierliche Variable analysiert werden. <u>Global Amendment 4 vom 10.08.2011</u> • Global Amendment 4 wurde nach Kommentierung durch die FDA nicht an weitere Gesundheitsbehörden gesendet. • Auf Grundlage der Kommentare der FDA wurde Global Amendment 5 erstellt. <u>Global Amendment 5 vom 14.12.2011</u> Es wurden die folgenden (relevanten) Änderungen der geplanten statistischen Auswertung vorgenommen: • Um mögliche Bedenken auszuschließen, dass die im Global Amendment 4 geplanten Änderungen als informiert angesehen werden könnten, sollten sämtliche CEC-bestätigten Ereignisse der klinischen Verschlechterung (Morbiditäts- oder Mortalitätsereignisse), die vor dem 16.08.2011 aufgetreten waren, nicht in die Analyse des primären Endpunkts einfließen. • Die ursprünglich angenommen Hazard Ratio von 0,5729 für

Item[a]	Charakteristikum	Studieninformation
		den primären Wirksamkeitsendpunkt basierte auf Daten aus früheren PAH-Monotherapie-Studien mit Patienten in den WHO-/NYHA-Klassen III und IV, also nicht mit Patienten in den Klassen II und III. Da der Großteil der Patienten in der Studie nicht therapienaiv war, sondern bereits PAH-spezifische Begleittherapie erhielt und zudem überwiegend in den WHO-/NYHA-Klassen II und III war, wurde die angenommene Hazard Ratio zu 0,650 geändert, um mit einer 90 % Power einen geringeren, aber immer noch klinisch relevanten Behandlungseffekt zu erreichen. • Um die erhöhte angenommene Hazard Ratio von 0,65 in der Analyse des primären Endpunkts widerzuspiegeln ohne gleichzeitig die Protokollvoraussetzungen für die Typ-I- und Typ-II-Fehlerraten ändern zu müssen, wurde die benötigte Anzahl von Ereignissen des primären Endpunkts auf 332 erhöht. • Um die Anzahl von 332 Ereignissen des primären Endpunkts in 3,5 Jahren erreichen zu können, wurde die Patientenzahl von 670 auf 1.150 Patienten (von 335 auf 575 Patienten je Gruppe) angepasst. Der Zeitraum bis zum Studienende, in der Patienten die Studienmedikation erhalten sollten, wurde nun mit 1,2-4,3 Jahren angenommen. • Aufgrund der beträchtlichen Vergrößerung der Stichprobengröße wurde die Möglichkeit eines vorzeigten Studienendes wegen sehr guter oder unzureichender Wirksamkeit in das Studiendesign eingeführt. Es wurde geplant eine Interim-Analyse durch eine unabhängiges Data Monitoring Committee nach 202 Ereignissen des primären Endpunkts (entsprechend etwa 61 % der benötigten Ereignisse) durchführen zu lassen. • Entsprechend der Haybittle-Peto-Einstellregel wurde der Typ-I-Fehler der Interimanalyse als 0,0001 (einseitig) festgelegt. Die Grenzwerte für sehr gute oder unzureichende Wirksamkeit (Z-Scores und geschätzte Hazard Ratios) wurde im Amendment festgelegt. • Zur Verbesserung von Präzision und Power sowie zur Verringerung von Verzerrung, wurde die Durchführung eines Log-Rang-Tests, stratifiziert nach WHO-/NYHA-Klasse I/II und III/IV zum Zeitpunkt der Basiswerterhebung als Sensitivitätsanalyse hinzugefügt. • Die alternative Hypothese (H1) wurde näher erläutert und der Typ-II-Fehler wurde von β=0,100 (Power=90 %) zu β<0,10 (Power von mindestens 90 %) geändert. • Die Hauptsensitivitätsanalyse des primären Endpunkts sollte mit 378 Ereignissen (inklusive der zensierten Ereignisse bis zum 16.08.2011) durchgeführt werden. • Die sekundären Endpunkte sollten zu einem einseitigen Signifikanzniveau von 0,005 durchgeführt werden (entspricht einem zweiseitigen Niveau von 0,01) Weitere Änderungen, die mit dem Amemdment eingeführt wurden: • Die Sicherheitspopulation wurde von der Liste zu analysierender Populationen gestrichen. Stattdessen sollte die All-treated-Population für die Analyse der Sicherheit

Item[a]	Charakteristikum	Studieninformation
		verwendet werden. Global Amendment 6 vom 23.01.2013 • Die Bezeichnung des primären Endpunkts wurde von „Zeit bis zum ersten Ereignis der klinischen Verschlechterung“ zu „Zeit bis zum ersten Morbiditäts- oder Mortalitätsereignis“ geändert. • Um weitere klinische Daten bis zum Studienende bei Patienten sammeln zu können, die die Behandlung mit der Studienmedikation vorzeitig abgebrochen hatten, wurde eine Beobachtungsphase nach der Behandlungsphase eingeführt. Diese wurde als der Zeitraum zwischen EOS-Untersuchung (nicht eingeschlossen) und der Abschlussuntersuchung der Nachbeobachtungsphase (eingeschlossen) nach Bekanntgabe des Studienendes durch den Sponsor definiert. Während der Nachbeobachtungsphase sollten Untersuchungstermine vor Ort und Telefonanrufe alternierend im Abstand von 3 Monaten durchgeführt werden. • Folgende Änderungen am Abschnitt zur statistischen Analyse wurden vorgenommen (nur relevante Änderungen sind dargestellt): o Die Wahrscheinlichkeit des einseitigen Typ-I-Fehlers der Interim-Analyse wurde zu 0,00005 korrigiert. o Es wurden weitere unterstützende Analysen für den primären Endpunkt hinzugefügt. o Für die statistische Analyse sekundärer Endpunkte mit ordinalen Variablen wurde die Verwendung einer nicht-parametrischen ANCOVA mit den Basiswerten als Kovariablen anstelle eines Wilcoxon-Rangsummen-Tests festgelegt.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien • Unterschriebene Einverständniserklärung vor Beginn jeglicher studienbezogener Prozeduren • Männliche und weibliche Patienten im Alter von 18 bis 75 Jahren mit symptomatischer PAH o Frauen im gebährfähigen Alter mussten eine zuverlässige Verhütungsmethode (Versagensquote niedriger als 1 % pro Jahr) vom Zeitpunkt des Screenings bis 1 Monat nach Behandlungsende mit der Studienmedikation anwenden. o Frauen wurde dann als im gebährfähigen Alter angesehen, wenn sie eins der folgenden Kriterien erfüllten: ▪ Frühere bilaterale Salpingekto-Oophorektomie oder Hysterektomie ▪ Vorzeitige Ovarialinsuffizienz, durch einen spezialisierten Gynäkologen bestätigt ▪ Prä-pubertär, XY-Genotyp, Turner-Syndrom, Mayer-Rokitansky-Küster-Hauser-Syndrom ▪ Alter >50 Jahre mit Amenorrhoe für mindestens 24 aufeinanderfolgende Monate

Item[a]	Charakteristikum	Studieninformation

		vor dem Screening
		• Patienten mit pulmonaler Hypertonie (PH), die einer der folgenden Subgruppen von Gruppe 1 des aktualisierten klinischen Klassifikationssystem aus Dana Point 2008 zuzuordnen ist:
		  ○ Idiopathische PAH
		  ○ Angeborene PAH
		  ○ Durch Medikamente oder Toxine verursachte PAH
		  ○ Assoziierte PAH in Verbindung mit:
		  ▪ Bindegewebserkrankungen
		  ▪ Angeborenen systemisch-pulmonalen Shunts (mindestens 1 Jahr nach chirurgischem Verschluss)
		  ▪ Einer HIV-Infektion
		• Dokumentierte hämodynamische Diagnose der PAH mittels Rechtsherzkatheteruntersuchung zu einem beliebigen Zeitpunkt vor dem Screening, die die folgenden Kriterien erfüllt:
		  ○ Mittlerer pulmonal arterieller Druck (mPAP) in Ruhe von ≥25 mmHg
		  und
		  ○ Pulmonaler Gefäßwiderstand (PVR) in Ruhe von ≥400 dyn s cm^{-5}
		  und
		  ○ Pulmonaler Verschlussdruck (PCWP) oder linksventrikulärer enddiastolischer Druck (LVEDP) ≤15 mmHg
		• 6-MWD zwischen 50 und 450 m (einschließlich) zum Zeitpunkt des Screening (bis zu 2 Wochen vor der Basiswert-Untersuchung und an einem anderen Tag als die Basiswert-Untersuchung)
		
		<u>Ausschlusskriterien</u>
		• Patienten mit PH, die den Gruppen 2-5 oder den Subgruppen von Gruppe 1, die nicht durch die Einschlusskriterien abgedeckt sind, des aktualisierten klinischen Klassifikationssystem aus Dana Point 2008 zuzuordnen ist
		• Patienten, die innerhalb eines Monats vor der Basiswertuntersuchung mit Prostacyclin (Epoprostenol) oder Prostacyclin-Analoga (Treprostinil, Iloprost, Beraprost) behandelt worden waren, oder bei denen während der Studie eine solche Behandlung geplant war
		  ○ Die einmalige Applikation von intravenösen oder inhalierten Prostacyclin-Analoga während des Vasoreagibilitätstests im Rahmen der Rechtsherz-katheteruntersuchung war erlaubt.
		• Patienten mit moderater oder schwerer obstruktiver Lungenerkrankung (FEV_1/FVC <70 % und FEV_1 <65 % des prognostizierten Wertes nach Bronchodilatorbehandlung
		• Patienten mit moderater oder schwerer restriktiver

Item[a]	Charakteristikum	Studieninformation
		Lungenerkrankung (gesamte Lungenkapazität <70 % des prognostizierten Wertes • Patienten mit moderater oder schwerer Leberfunktionsstörung (Child-Pugh B and C, sofern Child-Pugh-Status bekannt) • Patienten mit bekannter Fehlfunktion des linken Ventrikels (Auswurffraktion <45 %) • Patienten mit schwerer Nierenfunktionsstörung (geschätzte Kreatinin-Clearance <30 ml/min oder Serum-Kreatinin >2,5 mg/dl) • Patienten mit BMI <18,5 kg/m^2 • Patienten die innerhalb eines Monats vor der Basiswert-Untersuchung mit einem beliebigen Arzneimittel in der klinischen Erprobung behandelt wurden. • Akute oder chronische Störung (mit Ausnahme von Dyspnoe), die die Fähigkeit entsprechend der Studienerfordernisse zu handeln – im Speziellen in Bezug auf den 6-MWT –, einschränken (z. B. Angina pectoris, Hinken, muskuloskeletale Störungen, Notwendigkeit für Gehhilfen) • Kürzlich absolvierte (das Programm sollte mindestens 8 Wochen vor der Screening-Untersuchung abgeschlossen sein) oder geplante Teilnahme an einem kardio-pulmonalen Reha-Programm, das auf körperlichem Training basiert • Psychotische Störung, Suchtstörung oder jegliche Störung, die die Fähigkeit eine informierte Einverständniserklärung abzugeben oder entsprechend der Studienerfordernisse zu handeln, einschränkt • Lebenserwartung von weniger als 12 Monaten • Stillende oder schwangere (positiver Serum-Schwangerschaftstest vor der Randomisierung) Frauen oder solche Frauen, die während der Studie planen schwanger zu werden. • Bekannte Überempfindlichkeit gegen einen der Inhaltsstoffe der Studienmedikation Änderungen der Ein- und Ausschlusskriterien im Verlauf der Studie sind unter Punkt 3b beschrieben.
4b	Studienorganisation und Ort der Studiendurchführung	Die Studie wurde in 181 Zentren in 39 Ländern durchgeführt. Geographische Regionen: Asien, Australien, Europa, Lateinamerika und Nordamerika
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	*Intervention* Selexipag *Darreichungsform* Filmtabletten mit 200 µg Selexipag zur oralen Aufnahme *Anwendung* Patienten erhielten am Abend des ersten Tages die erste Dosierung Selexipag (200 µg). Beginnend von Tag 2 sollte der Patient das Arzneimittel zweimal täglich und im Abstand von etwa 12 Stunden einnehmen. Sofern die erste Dosierung (200 µg zweimal täglich) gut vertragen wurde, sollte der Patienten die Dosierung in wöchentlichen Schritten von 200 µg zweimal täglich, so lange

Item[a]	Charakteristikum	Studieninformation
		steigern, bis die maximal tolerierte Dosis (oder eine Maximaldosierung von 1.600 µg zweimal täglich) erreicht wurde. Der Prüfarzt konnte sich bei jedem Titrationsschritt gegen eine weitere Auftitration entscheiden, sofern nach seiner medizinischen Einschätzung Auftreten und Schwere der typischen Nebeneffekte einer Behandlung mit Agonisten des IP-Rezeptors wie z. B. Kopfschmerzen, Durchfall, Kieferschmerz, Myalgie, Erröten oder Übelkeit für den Patienten nicht mehr tolerierbar war. In solchen Fällen sollte die Dosierung des Arzneimittels um 200 µg zweimal täglich reduziert und als individuelle maximal tolerierte Dosis des Patienten beibehalten werden. Um die Wahrscheinlichkeit des Auftretens typischer pharmakologischer Nebeneffekte der IP-Rezeptorantagonisten-Behandlung zum Zeitpunkt der Auftitration zu vermeiden, sollten die Patienten die erste gesteigerte Dosis am Abend einnehmen. Die individuell höchste tolerierte Dosis des Patienten sollte von Woche 12 bis Woche 26 stabil gehalten werden. Nach Woche 26 waren weitere Steigerungen der Dosis durch den Prüfarzt in Schritten von 200 µg zweimal täglich bis zur Maximaldosierung von 1.600 µg zweimal während der Untersuchungen erlaubt. Reduktionen der Dosierung konnten in diesem Zeitraum bei Verträglichkeitsproblemen jederzeit durchgeführt werden. Die Patienten wurden instruiert vom Morgen der zweiten Untersuchung an und nachfolgend an jedem weiteren Untersuchungstermin die morgendliche Dosis erst nach Durchführung des 6-MWT (sofern vorgesehen), des CAMPHOR (sofern vorgesehen), der EKG-Untersuchung oder Blutprobenentnahme einzunehmen. *Intervention* Placebo *Darreichungsform* Filmtabletten zur oralen Aufnahme, identisch in Aussehen und Geschmack zu Selexipag *Anwendung* Das Vorgehen zur Anwendung der Placebo-Tabletten entspricht dem Vorgehen für Selexipag.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<u>Primärer Endpunkt</u> Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis bis zu 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase von AC-065A302 (d. h. EOT + 7 Tage). Als M/M-Ereignisse wurden berücksichtigt: • Tod (jeglicher Ursache) oder • Hospitalisierung aufgrund einer Verschlechterung der PAH (basiert auf vordefinierten Kriterien) oder • Verschlechterung der PAH, die eine Lungentransplantation oder eine atriale Ballon-Septostomie notwendig macht

Item[a]	Charakteristikum	Studieninformation

		oder
		• Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH
		oder
		• Fortschreiten der Erkrankung (Patienten, die zum Zeitpunkt der Basiswerterhebung in den WHO-/NYHA-Klassen II oder III waren), bestätigt durch:
		  ○ Verschlechterung der 6-MWD im Vergleich zum Basiswert (≥15 %, bestätigt durch 2 Tests an zwei Tagen innerhalb von 2 Wochen)
		  und
		  ○ Verschlechterung der WHO-/NYHA-Klasse
		oder
		• Fortschreiten der Erkrankung (Patienten, die zum Zeitpunkt der Basiswerterhebung in den WHO-/NYHA-Klassen III oder IV waren), bestätigt durch:
		  ○ Verschlechterung der 6-MWD im Vergleich zum Basiswert (≥15 %, bestätigt durch 2 Tests an zwei Tagen innerhalb von 2 Wochen)
		  und
		  ○ Notwendigkeit neuer PAH-spezifischer Medikation
		Patienten, die zum Zeitpunkt der Basiswerterhebung in der WHO-/NYHA-Klasse III waren, waren für beide Definitionen des Fortschreitens der Erkrankung qualifiziert.
		Sämtliche M/M-Ereignisse wurden durch ein unabhängiges Kontrollgremium, dem Critical Event Committee (CEC), in verblindeter Weise begutachtet und bewertet.
		
		<u>Sekundäre Wirksamkeitsendpunkte</u>
		• Absolute Veränderung der 6-MWD in Woche 26 im Vergleich zum Basiswert, gemessen zum Zeitpunkt des Talspiegels der Studienmedikation („*at trough*“).
		  Der 6-MWT wurde dann als zum Zeitpunkt des Talspiegels gemessen betrachtet, wenn der Zeitpunkt der letzten Einnahme der Studienmedikation entweder am vorherigen Tag war oder aber am gleichen Tag, die Einnahme aber mindestens 12 Stunden zurücklag. Wenn der 6-MWT für Woche 26 dem 6-MWT entsprach, der für ein M/M-Ereignis durchgeführt wurde, so wurde er grundsätzlich als „zum Talspiegel“ angesehen.
		• Ausbleibende Verschlechterung der WHO-/NYHA-Klasse in Woche 26 im Vergleich zum Basiswert
		• Zeit bis zum CEC-bestätigten Tod oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH bis zu 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase von AC-065A302 (EOT + 7 Tage)
		• Zeit bis zum Tod (Studienschluss)
		• Absolute Veränderung der „Kurzatmigkeit“-Subskala des

Item[a]	Charakteristikum	Studieninformation
		CAMPHOR-„Symptome“-Scores in Woche 26 im Vergleich zum Basiswert • Absolute Veränderung des CAMPHOR-„Symptome“-Scores in Woche 26 im Vergleich zum Basiswert <u>Explorative Wirksamkeitsendpunkte</u> • Zeit bis zum ersten CEC-bestätigten M/M-Ereignis (unter Ausschluss von „Fortschreiten der Erkrankung“) bis Studienschluss • Zeit bis zum ersten CEC-bestätigten M/M-Ereignis (unter Ausschluss von „Fortschreiten der Erkrankung“) bis EOT + 7 Tage) • Zeit bis zum ersten CEC-bestätigten M/M-Ereignis (unter Ausschluss von „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“) bis Studienschluss • Zeit bis zum ersten CEC-bestätigten M/M-Ereignis (unter Ausschluss von „Fortschreiten der Erkrankung“ und „Initiierung einer parenteralen Prostanoidtherapie oder einer dauerhaften Sauerstofftherapie aufgrund einer Verschlechterung der PAH“) bis EOT + 7 Tage • Zeit bis zum CEC-bestätigten Tod wegen PAH bis EOT + 7 Tage • Zeit bis zum CEC-bestätigten Tod oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH bis Studienschluss • Zeit bis zum ersten CEC-bestätigten M/M-Ereignis bis Studienschluss • Zeit bis zum CEC-bestätigten Tod wegen PAH bis Studienschluss • Absolute Veränderung der 6-MWD, gemessen zum Talspiegel, alle geplanten Untersuchungen (außer Woche 26) im Vergleich zum Basiswert • Absolute Veränderung der 6-MWD, gemessen zum Talspiegel, EOS im Vergleich zum Basiswert • Ausbleibende Verschlechterung der WHO-/NYHA-Klasse, alle geplanten Untersuchungen (außer Woche 26) im Vergleich zum Basiswert • Ausbleibende Verschlechterung der WHO-/NYHA-Klasse, EOS im Vergleich zum Basiswert • Borg-Dyspnoe-Index zum Zeitpunkt geplanter Untersuchungen • Borg-Dyspnoe-Index zum Zeitpunkt der EOS-Untersuchung • Absolute Veränderung des Plasma-Wertes für NT-pro-BNP zum Zeitpunkt geplanter Untersuchungen im Vergleich zum Basiswert • Absolute Veränderung des Plasma-Wertes für NT-pro-BNP zum Zeitpunkt der EOS-Untersuchung im Vergleich zum Basiswert • Absolute Veränderung der „Kurzatmigkeit“-Subskala des

Item[a]	Charakteristikum	Studieninformation
		CAMPHOR-„Symptome“-Scores, alle geplanten Untersuchungen (außer Woche 26) im Vergleich zum Basiswert • Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR-„Symptome“-Scores, EOS im Vergleich zum Basiswert • Absolute Veränderung des CAMPHOR-„Symptome“-Scores, alle geplanten Untersuchungen (außer Woche 26) im Vergleich zum Basiswert • Absolute Veränderung des CAMPHOR-„Symptome“-Scores, EOS im Vergleich zum Basiswert <u>Pharmakoökonomische Endpunkte</u> • Anzahl PAH-bezogener Hospitalisierungen pro Jahr bis zur EOS-Untersuchung • Anzahl an Tagen mit stationärem Aufenthalt aufgrund PAH-bezogener Gründe bis zur EOS-Untersuchung • Anzahl jeglicher Ursache pro Jahr bis zur EOS-Untersuchung • Anzahl an Tagen mit stationärem Aufenthalt jeglicher Ursache bis zur EOS-Untersuchung <u>Sicherheitsendpunkte</u> • Unerwünschte Ereignisse (UE) • Schwerwiegende unerwünschte Ereignisse (SUE) • Therapieabbrüche wegen unerwünschter Ereignisse • Unerwünschte Ereignisse von besonderem Interesse (Hämorrhagie, Störungen des Auges und Hypotonie) • Prostacyclin-assoziierte unerwünschte Ereignisse • EKG • Labortests • Vitalzeichen und Körpergewicht • Fundoskopie
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	siehe Abschnitt 3b
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Aufgrund von Änderungen durch die Global Amendments 1 (Zusammenlegung der Studien AC-065A301 (nicht initiiert) und AC-065A302), 5 (Veränderung der erwarteten Hazard Ratio) und 6 (Korrektur eines Rundungsfehlers) wurde die Fallzahlplanung im Verlauf der Studie mehrfach angepasst. Dargestellt wird nur die letztlich relevante letzte Fallzahlplanung. Die vorherigen Fallzahlplanungen sind in Abschnitt 9.8.12 des Studienberichts beschrieben. <u>Fallzahlplanung nach Global Amendment 6</u> Die Fallzahlplanung wurde mit Version 5.3 von East™ durchgeführt. Das Gruppen-sequentielle Design wurde folgendermaßen

Item[a]	Charakteristikum	Studieninformation
		zusammengefasst: • Die einseitige Typ-I-Fehlerwahrscheinlichkeit wurde auf α=0,005 festgelegt • Die maximal benötigte Information wurde auf 331 CEC-bestätigte M/M-Ereignisse festgelegt • Die einseitige Typ-I-Fehlerwahrscheinlichkeit für die Interim-Analyse (nach Erreichen von 202 Ereignissen geplant) wurde auf 0,00005 festgelegt Es wurden zur Berechnung von Stichprobengröße und Power die folgenden Annahmen getroffen: • Konstante Hazard-Rate von 0,2231 pro Jahr für Placebo • Konstante Hazard-Rate von 0,145 pro Jahr für Selexipag (entspricht einer Hazard Ratio von 0,65) • Konstante Hazard-Rate von 0,051 pro Jahr für Drop-Outs (Beendigung der Studie ohne ein M/M-Ereignis) für beide Behandlungsgruppen Dies führte zu einer Gesamtzahl von 331 CEC-bestätigten M/M-Ereignissen, um mit einer Power von 90 % die Nullhypothese ablehnen zu können. Durch die Rekrutierung von 1.150 Patienten, die im 1:1 Verhältnis auf den aktiven Wirkstoff oder Placebo randomisiert werden, über eine Zeitraum von 3,1 Jahren (150 Patienten im ersten Jahr und 520 Patienten pro Jahr in den darauffolgenden Jahren) sollte die benötigte Anzahl von Ereignissen in einem Zeitraum von maximal 4,3 Jahren erreicht werden. Unter diesen Bedingungen beträgt die Wahrscheinlichkeit die Nullhypothese während der Interim-Analyse (korrekt) abzulehnen etwa 19 %, während die Wahrscheinlichkeit sie zu diesem Zeitpunkt (fälschlicherweise) abzulehnen weniger als 0,2 % beträgt. Sollte die tatsächliche Hazard Ratio 1 betragen, so beträgt die Wahrscheinlichkeit die Nullhypothese während der Interim-Analyse (korrekt) abzulehnen etwa 50 %.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurde eine Interim-Analyse nach Auftreten von 202 Ereignissen des primären Endpunkts durchgeführt. Die Analyse wurde vom Statistical Analysis Center (Frontier Science) durchgeführt, die ansonsten nicht in die Studie involviert waren. Die Ergebnisse der Interim-Analyse wurden durch das Data Monitoring Committee begutachtet. Die Entscheidung, ob die Studie vorzeitig beendet oder fortgeführt werden sollte, wurde anhand der Z-Statistik des Log-Rang-Tests getroffen. Es wurden dabei folgende Grenzwerte festgelegt: • Vorzeitiges Studienende wegen sehr guter Wirksamkeit: Z-Score ≥3,8906 • Vorzeitiges Studienende wegen unzureichender Wirksamkeit: Z-Score ≤0 Unabhängig vom Über- bzw. Unterschreiten der jeweiligen Entscheidungsgrenzen konnte das Data Monitoring Committee jedoch auch eine Weiterführung der Studie bis zum geplanten Ende empfehlen. Ebenso konnte sich auch Actelion unabhängig von der Entscheidung des DMC für eine Weiterführung entscheiden.

Item[a]	Charakteristikum	Studieninformation
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Zuteilung auf die Behandlungsgruppen erfolgte durch ein zentralisiertes Randomisierungssystem mithilfe von Interactive Voice Response (IVR) oder Interactive Web Response (IWR). Die Randomisierung wurde durch ein unabhängiges Auftragsunternehmen (ICON Clinical Research, USA) durchgeführt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden in einem 1:1-Verhältnis auf die beiden Behandlungsgruppen randomisiert. Jedem Patienten wurde dabei eine eindeutige Randomisierungsnummer zugeteilt. Es fand eine Stratifizierung nach Zentren statt. Als Blockgröße wurde 4 verwendet.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Randomisierungsliste wurde bis zur Entblindung für die abschließende Daten-Analyse strikt vertraulich behandelt und war nur berechtigten Personen zugänglich, die nicht in die Durchführung und Analyse der Studie involviert waren. Eine versiegelte Randomisierungsliste wurde durch Actelion Quality Management an einem sicheren Ort aufbewahrt. Eine zweite Randomisierungsliste wurde über ICON dem Statistiker des unabhängigen Statistical Analysis Center und der SwissBioQuant AG (zur Bestimmung der Plasmakonzentration von Selexipag und ACT-333679).
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Zuteilung auf die Behandlungsgruppen erfolgte durch ein zentralisiertes Randomisierungssystem mithilfe von Interactive Voice Response (IVR) oder Interactive Web Response (IWR). Die Randomisierung wurde durch ein unabhängiges Auftragsunternehmen (ICON Clinical Research, USA) durchgeführt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar. Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Selexipag und Placebo waren nicht unterscheidbar, alle Studienmedikation-Kits wurden in gleicher Weise verpackt.
12	Statistische Methoden	
12a	Statistische Methoden zur	Hypothesen

Item[a]	Charakteristikum	Studieninformation
	Bewertung der primären und sekundären Zielkriterien	Die Hypothesen wurden hinsichtlich der Überlebensfunktion S(t) formuliert, d. h. die Wahrscheinlichkeit der Zeit bis zum ersten CEC-bestätigten Morbiditäts- oder Mortalitätsereignis betrug ≥t für einen Tag t (relative zum Zeitpunkt der Randomisierung). Die einseitigen Hypothesen wurden hinsichtlich der Überlebens-funktionen $S_{M/M,Selexipag}(t)$ und $S_{M/M,Placebo}(t)$ ausgedrückt. $H_{0,M/M}$: $S_{M/M,Selexipag}(t) \leq S_{M/M,Placebo}(t)$ für alle $t \geq 0$ gegenüber $H_{A,M/M}$: $S_{M/M,Selexipag}(t) \geq S_{M/M,Placebo(}t)$ für alle $t \geq 0$ und $S_{M/M,Selexipag}(t) > S_{M/M,Placebo}(t)$ für einige $t > 0$. <u>Primärer Endpunkt</u> Die primäre Wirksamkeitsanalyse wurde auf Grundlage der FAS-Population mit einem einseitigen, nicht-stratifizierten Log-Rang-Test im Group-Sequential-Design. Die Eigenschaften des Group-Sequential-Designs waren: • Die einseitige Gesamtwahrscheinlichkeit des Typ-I-Fehlers wurde auf α=0,0005 festgelegt. • Die maximal notwendige Information wurde mit 331 CEC-bestätigten M/M-Ereignissen spezifiziert. • Die einseitige Wahrscheinlichkeit des Typ-I-Fehlers der Interim-Analyse wurde auf 0,00005 festgelegt. • Für die Interim-Analyse wurde mit 202 CEC-bestätigten M/M-Ereignissen geplant (entspricht einem Informationsanteil von *f*=202/331=0,61) • Der Behandlungseffekt wurde mit einem Proportional-Hazard-Modell bestimmt. <u>Sekundäre Wirksamkeitsendpunkte</u> Die sekundären Wirksamkeitsendpunkte wurden in folgender konditionaler Abfolge getestet: • Absolute Veränderung der 6-MWD in Woche 26 im Vergleich zum Basiswert, gemessen zum Talspiegel • Abwesenheit der Verschlechterung der WHO-/NYHA-Klasse in Woche 26 im Vergleich zum Basiswert • Zeit bis zum CEC-bestätigten Tod oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH bis zu 7 Tage nach der letzten Einnahme der Studienmedikation in der Behandlungsphase von AC-065A302 (EOT + 7 Tage) • Zeit bis zum Tod (Studienschluss) • Absolute Veränderung der „Kurzatmigkeit“-Subskala des CAMPHOR-„Symptome“-Scores in Woche 26 im Vergleich zum Basiswert • Absolute Veränderung des CAMPHOR-„Symptome“-Scores in Woche 26 im Vergleich zum Basiswert Die Nullhypothese konnten für einen sekundären Wirksamkeitsendpunkt nur dann abgelehnt werden, wenn auch bereits für alle vorangegangenen Wirksamkeitsendpunkte (einschließlich des primären Endpunkts) die Nullhypothese

Item[a]	Charakteristikum	Studieninformation
		abgelehnt worden war. Das einseitige Signifikanzniveau wurde für alle sekundären Endpunkte als 0,005 festgelegt; zusätzlich wurden die zweiseitigen 99 %-Konfidenzintervalle dargestellt. <u>Sicherheitsendpunkte</u> Die Sicherheitsendpunkte wurden deskriptiv analysiert. Das Auftreten von (potentiell wiederkehrenden) UE von besonderem Interesse wurde mithilfe des Andersen-Gill-Modells analysiert.

12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Es wurden folgende Subgruppenanalysen durchgeführt:
		• PAH-spezifische Begleittherapie zum Zeitpunkt der Basiswerterhebung: keine PAH-spezifische Medikation, ERA-Monotherapie, PDE-5-I-Monotherapie und Kombination aus ERA und PDE-5-I
		• WHO-/NYHA-Klasse zum Zeitpunkt der Basiswerterhebung: I/II und III/IV
		• Geschlecht: männlich, weiblich
		• Rasse/Ethnizität: Kaukasisch/Hispanisch, Asiatisch, Schwarz und Andere
		• Alter zum Zeitpunkt des Screenings: <65 Jahre und ≥65 Jahre
		• PAH-Ätiologie zum Zeitpunkt der Basiswerterhebung:
		  o Idiopathische, angeborene, durch Medikamente oder Toxine verursachte PAH
		  o PAH assoziiert mit Bindegewebserkrankungen
		  o PAH assoziert mit angeborenem systemisch-pulmonalen Shunt (mindestens 1 Jahr nach chirurgischem Verschluss)
		  o Geographische Region
		Die Subgruppenanalysen wurden in der FAS-Population für den primären und sekundäre Endpunkte anhand der folgenden Methodik durchgeführt:
		• Separate Analyse von jeder der jeweiligen Subgruppen mit den gleichen statistischen Methoden:
		  o Log-Rang-Test und Proportional-Hazard-Modell für Endpunkte mit Ereigniszeitanalysen
		  o (nicht-parametrische) ANCOVA für kontinuierliche Endpunkte
		• Modifizierte statistische Modelle, die die Berücksichtigung von Behandlung-Subgruppe-Interaktionen ermöglichen:
		  o Proportional-Hazard-Modell mit Studienbehandlung, Subgruppe und Behandlung-Subgruppe-Interaktions-Termen für Ereigniszeitanalysen.
		Je Subgruppe wurden die Punktschätzer des Behandlungseffekts und zugehörige zweiseitige 99 %-Konfidenzintervalle in Form eines Forest-Plots dargestellt.

Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	

Item[a]	Charakteristikum	Studieninformation
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) 1.156 Patienten wurden randomisiert, 574 auf Selexipag und 582 auf Placebo b) 1.152 Patienten wurden behandelt; 574 Patienten davon erhielten Selexipag und 578 Patienten Placebo c) Alle randomisierten Patienten wurden entsprechend ihrer Randomisierung in der Analyse des primären Endpunkts ausgewertet (vgl. Punkt a).
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Patientenfluss (Abbildung 4-40)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Erster Patient, erste Aufnahme: 30.12.2009 Letzter Patient, letzte Untersuchung: 27.04.2014
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

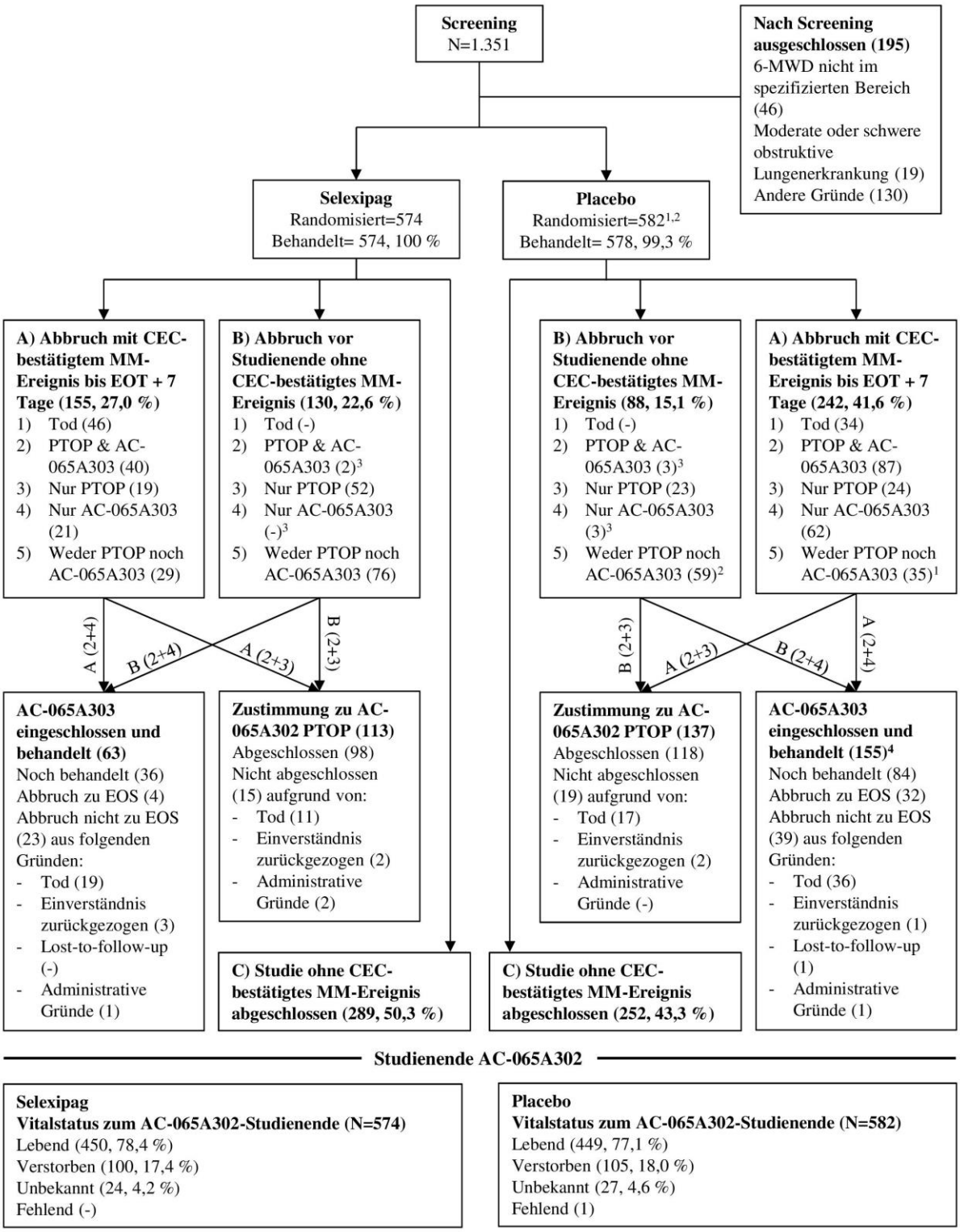


Abbildung 4-40: Patientenfluss in Studie AC-065A302 – GRIPHON

[1]) Ein Patient hatte ein CEC-bestätigtes M/M-Ereignis einen Tag nach Randomisierung zur Placebo-Gruppe und hatte daher keine Studienmedikation erhalten.

[2]) 2 Patienten, die auf Placebo randomisiert wurden, erhielten keine Studienmedikation, da sie vor ihre Zustimmung zu sämtlichen Studienkomponenten zurückgezogen hatten. Ein Patient in der Placebo-Gruppe hatte aus administrativen Gründen keine Studienmedikation erhalten.

[3]) Der Patient wurde aufgrund eines zunächst vom CEC-bestätigten Morbiditätsereignis in die Extensionsstudie AC-065A303 aufgenommen. Während des Wiederbegutachtungsprozesses wurde das Ereignis nachträglich jedoch als nicht durch das CEC bestätigt eingestuft.

[4]) Patienten wurden auf Placebo in GRIPHON (AC-065A302) randomisiert und in AC-065A303 mit Selexipag behandelt.

6-MWD - 6-Minuten-Gehdistanz, CEC - Critical Event Committee, M/M-Ereignis - Morbiditäts- oder Mortalitätsereignis, EOT - Ende der Behandlung (*end-of-treatment*), EOS - Ende der Studie (*end-of-study*), PTOP - Beobachtungsphase nach der Behandlungsphase (*post-treatment observation phase*)

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.

- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C …) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Tabelle 4-140 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie AC-065A302 – GRIPHON

Studie: AC-065A302 – GRIPHON

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht einschließlich Appendices [78]	A
Für das Nutzendossier durchgeführte Post-hoc-Analysen [79]	B
Registereintrag EU-CTR [77]	C
Registereintrag Clinicaltrials.gov [76]	D
Publikation [80]	E

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☒ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Laut Studientitel und Beschreibung im Studienbericht handelt es sich um eine randomisierte, kontrollierte Studie. (A)

1.
<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

2.

<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Erzeugung der Randomisierungssequenz wurde in GRIPHON adäquat durchgeführt, die Gruppenzuteilung fand verdeckt statt. Bei GRIPHON handelt es sich um eine doppelt verblindete Studie, d. h. sowohl Behandler als auch Patient waren verblindet. Die Ereignisse des primären Endpunkts wurden zudem durch ein verblindetes Critical Event Committee dahin gehend begutachtet, ob ein direkter/unmittelbarer Zusammenhang zur PAH besteht.
Hinweise auf eine ergebnisabhängige Berichterstattung sind nicht zu beobachten, und es liegen auch keine sonstigen endpunktübergreifenden Aspekte vor, die zu einer Verzerrung der Ergebnisse führen könnten.
Das Verzerrungspotential in der dargestellten Studie ist somit als niedrig anzusehen.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Zeit bis zum ersten CEC-bestätigten kombinierten Morbiditäts- oder Mortalitätsereignis

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC

darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Zeit bis zum CEC-bestätigten Tod wegen PAH oder bis zur ersten CEC-bestätigten Hospitalisierung aufgrund einer Verschlechterung der PAH

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.

Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Zeit bis zum Tod (Studienschluss)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht

unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patienten im Placebo-Arm, bei denen ein nicht-tödliches Ereignis des primären Endpunkts aufgetreten war, konnten die doppelblinde Behandlung mit der Studienmedikation beenden (EOT) und anschließend mit Selexipag oder einem anderen PAH-spezifischen Arzneimittel weiter behandelt werden. Aufgrund dieses erlaubten Therapiewechsels stellen die Ergebnisse bis zum Ende nicht mehr den "reinen" Placebo-Effekt dar. Es ist jedoch höchstens mit einer Verzerrung der Studienergebnisse zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch den erlaubten Therapiewechsel im Placebo-Arm ein bedeutendes Verzerrungspotential. Dies führt tendenziell zu einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Zeit bis zum CEC-bestätigten Tod wegen PAH (EOT)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Bei PAH-Patienten tritt in aller Regel erst ein Morbiditätsereignis auf bevor sie dann im weiteren Verlauf der Krankheit versterben. In der Studiensituation von GRIPHON bedeutet dies, dass Todesfälle, die sich nach Auftreten eines Morbiditätsereignisses ergeben und dadurch die randomisierte Behandlungsphase für den Patienten beenden, in Auswertungen bis EOT nicht erfasst werden. Diese informative Zensierung führt letztlich zu einer Unterschätzung des Sterberisikos in beiden Behandlungsgruppen und kann damit zu einer substantiellen Verzerrung der Ergebnisse führen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips

hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Zeit bis zum CEC-bestätigten Tod wegen PAH (Studienschluss)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patienten im Placebo-Arm, bei denen ein nicht-tödliches Ereignis des primären Endpunkts aufgetreten war, konnten die doppelblinde Behandlung mit der Studienmedikation beenden (EOT) und anschließend mit Selexipag oder einem anderen PAH-spezifischen Arzneimittel weiter behandelt werden. Aufgrund dieses erlaubten Therapiewechsels stellen die Ergebnisse bis zum Ende nicht mehr den "reinen" Placebo-Effekt dar. Es ist jedoch höchstens mit einer Verzerrung der Studienergebnisse zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch den erlaubten Therapiewechsel im Placebo-Arm ein bedeutendes Verzerrungspotential. Dies führt tendenziell zu einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Symptome der PAH

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt die Auswertung eines Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer prostacyclinspezifischer Neben-wirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchsten mit einer Verzerrung zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch die Auswertung als naive Raten in Verbindung mit der ungleichen medianen Behandlungsdauer in den Studienarmen ein bedeutendes Verzerrungspotential. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm ist tendenziell jedoch von einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag auszugehen.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Absolute Veränderung der 6-MWD

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Ausbleibende Verschlechterung der WHO-/NYHA-Klasse

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Zeit bis zur ersten CEC-bestätigten Hospitalisierung wegen PAH

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die

Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Absolute Veränderung des Borg-Dyspnoe-Index

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht

unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☐ **ja** ☐ **unklar** ☒ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Alle Patienten wurden innerhalb der Behandlungsgruppen ausgewertet, auf die sie randomisiert worden waren. Es wurde keine Ersetzung fehlender Werte vorgenommen, sondern es wurden nur diejenigen Patienten in die Auswertung miteinbezogen, für die Werte zum Zeitpunkt der Basiswerterhebung und zu Woche 26 vorlagen. Insgesamt wurden in der Selexipag-Gruppe die Daten von 80 % der Patienten und in der Placebo-Gruppe von 76 % der Patienten in die Analyse miteinbezogen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet.
Da es sich beim Borg-Dyspnoe-Index um einen explorativen Endpunkt handelte, war eine Ersetzung fehlender Werte vorgenommen nicht *a priori* geplant und daher auch nicht durchgeführt. Es wurden ausschließlich diejenigen Patienten in die Analyse miteinbezogen, für die Werte zum Zeitpunkt der Basiswerterhebung und in Woche 26 vorlagen. Das ITT-Prinzip ist damit nicht vollständig umgesetzt.
Ansonsten konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen.

Endpunkt: Absolute Veränderung des CAMPHOR „Symptome "-Scores

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Absolute Veränderung der „Kurzatmigkeit "-Subskala des CAMPHOR „Symptome"-Scores

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☒ **ja** ☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Weiterhin konnten keine Hinweise für eine ergebnisabhängige Berichterstattung des Endpunktes oder sonstige das Verzerrungspotential beeinflussende Aspekte beobachtet werden.
Das Verzerrungspotential ist für diesen Endpunkt somit als niedrig anzusehen.

Endpunkt: Unerwünschte Ereignisse

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt die Auswertung eines Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer prostacyclinspezifischer Neben-wirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchsten mit einer Verzerrung

zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch die Auswertung als naive Raten in Verbindung mit der ungleichen medianen Behandlungsdauer in den Studienarmen ein bedeutendes Verzerrungspotential. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm ist tendenziell jedoch von einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag auszugehen.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Schwerwiegende unerwünschte Ereignisse

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt die Auswertung eines Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer prostacyclinspezifischer Neben-wirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchsten mit einer Verzerrung zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch die Auswertung als naive Raten in Verbindung mit der ungleichen medianen Behandlungsdauer in den Studienarmen ein bedeutendes Verzerrungspotential. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm ist tendenziell jedoch von einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag auszugehen.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Therapieabbrüche wegen unerwünschter Ereignisse

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC

darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt die Auswertung eines Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer prostacyclinspezifischer Neben-wirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchsten mit einer Verzerrung zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch die Auswertung als naive Raten in Verbindung mit der ungleichen medianen Behandlungsdauer in den Studienarmen ein bedeutendes Verzerrungspotential. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm ist tendenziell jedoch von einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag auszugehen.

Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Unerwünschte Ereignisse von besonderem Interesse

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. **Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt die Auswertung eines Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer

prostacyclinspezifischer Neben-wirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchsten mit einer Verzerrung zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch die Auswertung als naive Raten in Verbindung mit der ungleichen medianen Behandlungsdauer in den Studienarmen ein bedeutendes Verzerrungspotential. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm ist tendenziell jedoch von einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag auszugehen.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Endpunkt: Prostacyclin-assoziierte unerwünschte Ereignisse

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie wurde doppelblind durchgeführt. Sowohl Prüfärzte als auch Studienmitarbeiter, Patienten, Monitore, der Sponsor und Mitarbeiter von beteiligten Auftragsunternehmen (mit Ausnahme von SwissBioQuant AG, die zur Bestimmung der Plasmakonzentrationen von Selexipag und ACT-333679 entblindet werden mussten) hatten keine Kenntnis der zugeteilten Studienmedikation. Selexipag- und Placebo-Tabletten waren nicht unterscheidbar.
Ein Critical Event Committee bewertete zudem alle Mortalitäts- und Morbiditätsereignisse in verblindeter Art und urteilte (zustimmend oder ablehnend), ob diese Ereignisse in die Analyse des primären Endpunktes aufgenommen werden sollten. Zusätzlich urteilte das CEC darüber, ob Todesfälle ursächlich durch die PAH ausgelöst wurden.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

☐ **ja** ☒ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Gemäß dem Cochrane Handbook for Systematic Reviews of Interventions setzt die Auswertung eines Endpunkts als naive Raten, d. h. in Form von Risk Ratios, Odds Ratios und absoluter Risikodifferenz samt zugehörigen Konfidenzintervallen voraus, dass alle Patienten zu einem fixen Zeitpunkt ausgewertet werden. In die Auswertung der Sicherheitsendpunkte wurden jedoch sämtliche UE eingeschlossen, die vom Zeitpunkt der Basiswerterhebung bis zum Ende der Behandlung plus 7 Tage aufgetreten waren, d. h. es fand eben nicht eine Auswertung zu einem festgelegten und für alle Patienten gleichen Zeitpunkt statt. Eine Auswertung der unerwünschten Ereignisse in Form einer Ereigniszeitanalyse ist aufgrund des gewählten Studiendesigns jedoch ebenfalls nicht sinnvoll. Durch das gewählte Anwendungsschema von Selexipag, das die Titration der Dosierung bis zum (vorübergehenden) Auftreten nicht mehr tolerierbarer prostacyclinspezifischer Neben-wirkungen vorsieht, ist das Risiko für das Auftreten eines unerwünschten Ereignisses in Titrationsphase und Erhaltungsphase nicht gleich. Damit sind auch die Voraussetzungen für die Durchführung der Ereigniszeitanalyse nicht gegeben. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm (70,7 Wochen im Vergleich zu 63,7 Wochen in der Placebo-Gruppe) ist jedoch höchsten mit einer Verzerrung zuungunsten von Selexipag zu rechnen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

☐ **niedrig** ☒ **hoch**

Begründung für die Einstufung:

Die Auswertung des Endpunktes erfolgte verblindet, eine adäquate Umsetzung des ITT-Prinzips hatte stattgefunden. Hinweise für eine ergebnisabhängige Berichterstattung des Endpunkts wurden nicht beobachtet, jedoch besteht durch die Auswertung als naive Raten in Verbindung mit der ungleichen medianen Behandlungsdauer in den Studienarmen ein bedeutendes Verzerrungspotential. Aufgrund der längeren medianen Behandlungsdauer im Selexipag-Arm ist tendenziell jedoch von einer Unterschätzung des tatsächlichen Therapieeffekts von Selexipag auszugehen.
Das Verzerrungspotential ist für diesen Endpunkt somit als hoch anzusehen

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

☐ **ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

☐ **nein**: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.
→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

☐ **ja**: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

☐ **unklar**: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

☐ **nein**: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

☐ **ja**: Die Gruppen wurden zeitlich parallel verfolgt.

☐ **unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

☐ **nein**: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

☐ **unklar**: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

☐ **nein**: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja**: Eines der folgenden Merkmale trifft zu:
- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

☐ **unklar**: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

☐ **nein**: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja:** Die Patienten waren verblindet.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

☐ **ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:
- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter
- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:
- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
  Zulässige Gründe sind:
  - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
  - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
  - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.
Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.
- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
  - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
  - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
  - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
  - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

☐ **ja**

☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig**: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch**: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. **Verblindung der Endpunkterheber**

   Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
   In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

   ☐ **ja:** Der Endpunkt wurde verblindet erhoben.

   ☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

   ☐ **nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

   Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. **Adäquate Umsetzung des ITT-Prinzips**

   Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT(„intention to treat")-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set"). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

   ☐ **ja:** Eines der folgenden Merkmale trifft zu:
   - Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

☐ **unklar**: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

☐ **nein**: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

☐ **nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.
- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

☐ **ja**

☐ **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

☐ **niedrig**: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

☐ **hoch**: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung: