Dokumentvorlage, Version vom 21.02.2019

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Trifaroten (Selgamis®)

Galderma Laboratorium GmbH

Modul 4 A

Patienten ab zwölf Jahren mit mittelschwerer Acne vulgaris im Gesicht und/oder am Rumpf

Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen

Stand: 13.08.2020

Inhaltsverzeichnis

Seite

Tabellenverzeichnis 4
Abbildungsverzeichnis 8
Abkürzungsverzeichnis 10
4 Modul 4 – allgemeine Informationen 12
4.1 Zusammenfassung der Inhalte von Modul 4 13
4.2 Methodik 21
4.2.1 Fragestellung 21
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung 22
4.2.3 Informationsbeschaffung 26
4.2.3.1 Studien des pharmazeutischen Unternehmers 26
4.2.3.2 Bibliografische Literaturrecherche 26
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken 27
4.2.3.4 Suche auf der Internetseite des G-BA 28
4.2.3.5 Selektion relevanter Studien 29
4.2.4 Bewertung der Aussagekraft der Nachweise 30
4.2.5 Informationssynthese und -analyse 31
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien 31
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien 32
4.2.5.3 Meta-Analysen 38
4.2.5.4 Sensitivitätsanalysen 40
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren 41
4.2.5.6 Indirekte Vergleiche 43
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen 46
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel 47
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel 47
4.3.1.1.1 Studien des pharmazeutischen Unternehmers 47
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche 49
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken 51
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA 51
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel 52
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 54
4.3.1.2.1 Studiendesign und Studienpopulationen 54
4.3.1.2.2 Verzerrungspotenzial auf Studienebene 62
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien 63
4.3.1.3.1 Therapieerfolgsrate im Gesicht gemäß IGA – RCT 67
4.3.1.3.2 Therapieerfolgsrate am Rumpf gemäß PGA – RCT 72

4.3.1.3.3 Änderung der Anzahl der Läsionen im Gesicht – RCT 77
4.3.1.3.4 Änderung der Anzahl der Läsionen am Rumpf – RCT..... 84
4.3.1.3.5 Selbsteinschätzung der Verbesserung der Akne – RCT..... 91
4.3.1.3.6 Gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI – RCT 96
4.3.1.3.7 Sicherheit – RCT..... 104
4.3.1.3.8 Lokale Verträglichkeit – RCT..... 117
4.3.1.3.9 Subgruppenanalysen – RCT..... 131
4.3.1.4 Liste der eingeschlossenen Studien - RCT 134
4.3.2 Weitere Unterlagen..... 135
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 135
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 135
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche..... 136
4.3.2.1.3 Ergebnisse aus indirekten Vergleichen 136
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT 136
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT 139
4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT... 139
4.3.2.2 Nicht randomisierte vergleichende Studien..... 139
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien 139
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien..... 140
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien 141
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien..... 141
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien..... 142
4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien 143
4.3.2.3 Weitere Untersuchungen..... 143
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen 143
4.3.2.3.2 Charakteristika der weiteren Untersuchungen 144
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen..... 144
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen 144
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen 145
4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen..... 145
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens..... 146
4.4.1 Beurteilung der Aussagekraft der Nachweise 146
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß..... 146
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht 154
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte 154
4.5.1 Begründung für die Vorlage indirekter Vergleiche..... 154
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen..... 154
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen 155
4.5.4 Verwendung von Surrogatendpunkten..... 155
4.6 Referenzliste..... 156

Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche 162
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken..... 166
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)..... 168
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)..... 169
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT 175
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten 205
Anhang 4-G : Ergänzende Daten..... 235
Auswertungen zum DLQI/CDLQI aus PERFECT-1 (18251) und PERFECT-2 (18252) . 236
Unerwünschte Ereignisse aus PERFECT-1 (18251) und PERFECT-2 (18252)..... 244
Subgruppenanalysen aus PERFECT-1 (18251) und PERFECT-2 (18252) 266

Tabellenverzeichnis

Seite

Tabelle 4-1: Zusammenfassung der Ergebnisse aus PERFECT-1 und -2 14
Tabelle 4-2: Ein- und Ausschlusskriterien für RCT mit Trifaroten und der zweckmäßigen Vergleichstherapie 23
Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel 48
Tabelle 4-4: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel 49
Tabelle 4-5: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel 51
Tabelle 4-6: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel 52
Tabelle 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel 53
Tabelle 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel 55
Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel 57
Tabelle 4-10: Charakterisierung der Studienpopulationen (demographische Charakteristika) – RCT mit dem zu bewertenden Arzneimittel 57
Tabelle 4-11: Charakterisierung der Studienpopulationen (Basischarakteristika der Erkrankung) – RCT mit dem zu bewertenden Arzneimittel 58
Tabelle 4-12: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel 62
Tabelle 4-13: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel 63
Tabelle 4-14: Operationalisierung des Endpunkts Therapieerfolgsrate im Gesicht gemäß IGA 67
Tabelle 4-15: Bewertung des Verzerrungspotenzials für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA in RCT mit dem zu bewertenden Arzneimittel 68
Tabelle 4-16: Ergebnisse zum Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA aus RCT (Woche 12, PERFECT-1 und -2) 69
Tabelle 4-17: Ergänzende Ergebnisse zum Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA (Woche 12 und 52, Studie SATISFY) 69
Tabelle 4-18: Operationalisierung des Endpunkts Therapieerfolgs am Rumpf gemäß PGA 72
Tabelle 4-19: Bewertung des Verzerrungspotenzials für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA in RCT mit dem zu bewertenden Arzneimittel 73
Tabelle 4-20: Ergebnisse zum Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA aus RCT (Woche 12, PERFECT-1 und -2) 74

Tabelle 4-21: Ergänzende Ergebnisse zum Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA (Woche 12 und 52, Studie SATISFY) 74
Tabelle 4-22: Operationalisierung der Änderung der Anzahl der Läsionen im Gesicht 77
Tabelle 4-23: Bewertung des Verzerrungspotenzials für den Endpunkt Änderung der Anzahl der Läsionen im Gesicht in RCT mit dem zu bewertenden Arzneimittel 78
Tabelle 4-24: Ergebnisse zur absoluten Änderung der Anzahl der Läsionen im Gesicht aus RCT (Woche 12, PERFECT-1 und -2) 79
Tabelle 4-25: Ergebnisse zur prozentualen Änderung der Anzahl der Läsionen im Gesicht aus RCT (Woche 12, PERFECT-1 und -2) 80
Tabelle 4-26: Operationalisierung der Änderung der Anzahl der Läsionen am Rumpf 84
Tabelle 4-27: Bewertung des Verzerrungspotenzials für den Endpunkt Änderung der Anzahl der Läsionen am Rumpf in RCT mit dem zu bewertenden Arzneimittel 85
Tabelle 4-28: Ergebnisse zur absoluten Änderung der Anzahl der Läsionen am Rumpf aus RCT (Woche 12, PERFECT-1 und -2) 86
Tabelle 4-29: Ergebnisse zur prozentualen Änderung der Anzahl der Läsionen am Rumpf aus RCT (Woche 12, PERFECT-1 und -2) 87
Tabelle 4-30: Operationalisierung der Selbsteinschätzung der Verbesserung der Akne 91
Tabelle 4-31: Bewertung des Verzerrungspotenzials für den Endpunkt Selbsteinschätzung der Verbesserung der Akne in RCT mit dem zu bewertenden Arzneimittel 92
Tabelle 4-32: Ergebnisse zum Endpunkt Selbsteinschätzung der Verbesserung der Akne aus RCT (Woche 12, PERFECT-1 und -2) 93
Tabelle 4-33: Ergänzende Ergebnisse zum Endpunkt Selbsteinschätzung der Verbesserung der Akne (Woche 12 und 52, Studie SATISFY) 94
Tabelle 4-34: Operationalisierung von gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI 96
Tabelle 4-35: Bewertung des Verzerrungspotenzials für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI in RCT mit dem zu bewertenden Arzneimittel 97
Tabelle 4-36: Ergebnisse zur mittleren Änderung des DLQI und CDLQI aus RCT (PERFECT-1 und -2) 98
Tabelle 4-37: Ergebnisse zum Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI aus RCT (Woche 12, PERFECT-1 und -2) 99
Tabelle 4-38: Ergänzende Ergebnisse zum Endpunkt Lebensqualität gemäß DLQI / CDLQI (Woche 12 und 52, Studie SATISFY) 100
Tabelle 4-39: Operationalisierung von Sicherheit 104
Tabelle 4-40: Bewertung des Verzerrungspotenzials für Endpunkte der Sicherheit in RCT mit dem zu bewertenden Arzneimittel 105
Tabelle 4-41: Ergebnisse zu unerwünschten Ereignissen aus RCT (12 Wochen, PERFECT-1 und -2) 107
Tabelle 4-42: Ergebnisse zu häufigen unerwünschten Ereignissen aus RCT (≥ 10 Ereignisse in einem Behandlungsarm, PERFECT-1 und -2) 108

Tabelle 4-43: Ergebnisse zu Therapieabbrüchen wegen unerwünschter Ereignisse aus RCT (PERFECT-1 und -2) ... 109
Tabelle 4-44: Ergebnisse zu UE von besonderem Interesse aus RCT (PERFECT-1 und -2) 110
Tabelle 4-45: Ergänzende Ergebnisse zu unerwünschten Ereignissen (12 und 52 Wochen, SATISFY) ... 111
Tabelle 4-46: Ergänzende Ergebnisse zu häufigen unerwünschten Ereignissen (≥ 10 Ereignisse, 52 Wochen, SATISFY) ... 112
Tabelle 4-47: Ergänzende Ergebnisse zu Therapieabbrüchen wegen unerwünschter Ereignisse (52 Wochen, SATISFY) ... 113
Tabelle 4-48: Ergänzende Ergebnisse zu UE von besonderem Interesse (52 Wochen, SATISFY) ... 113
Tabelle 4-49: Operationalisierung von lokaler Verträglichkeit ... 117
Tabelle 4-50: Bewertung des Verzerrungspotenzials für den Endpunkt Lokale Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel ... 118
Tabelle 4-51: Ergebnisse zur lokalen Verträglichkeit im Gesicht aus RCT (schlechtester Post-Baseline-Wert über 12 Wochen; PERFECT-1 und -2) ... 119
Tabelle 4-52: Ergebnisse zur lokalen Verträglichkeit am Rumpf aus RCT (schlechtester Post-Baseline-Wert über 12 Wochen; PERFECT-1 und -2) ... 120
Tabelle 4-53: Ergänzende Ergebnisse zur lokalen Verträglichkeit von Trifaroten im Gesicht und am Rumpf (schlechtester Post-Baseline-Wert über 52 Wochen; Studie SATISFY) ... 129
Tabelle 4-54: Matrix der durchgeführten Subgruppenanalysen ... 132
Tabelle 4-55: Ergebnisse des Interaktionsterms der Subgruppenanalysen je Endpunkt (Studien PERFECT-1 und -2) ... 133
Tabelle 4-56: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche ... 136
Tabelle 4-57: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden ... 137
Tabelle 4-58: Operationalisierung von <Endpunkt xxx> ... 137
Tabelle 4-59: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche ... 138
Tabelle 4-60: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche ... 138
Tabelle 4-61: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien ... 140
Tabelle 4-62: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien ... 141
Tabelle 4-63: Operationalisierung von <Endpunkt xxx> ... 141
Tabelle 4-64: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 142
Tabelle 4-65: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen ... 144

Tabelle 4-66: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen 144
Tabelle 4-67: Zusammenfassung der Ergebnisse aus PERFECT-1 und -2 148
Tabelle 4-68: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 154
Tabelle 4-69: Bibliografische Literaturrecherche in EMBASE – Suche nach RCT mit dem zu bewertenden Arzneimittel 163
Tabelle 4-70: Bibliografische Literaturrecherche in MEDLINE – Suche nach RCT mit dem zu bewertenden Arzneimittel 164
Tabelle 4-71: Bibliografische Literaturrecherche in der Cochrane Library – Suche nach RCT mit dem zu bewertenden Arzneimittel 165
Tabelle 4-72: Suchstrategie der Recherche in Studienregistern – Suche nach RCT mit dem zu bewertenden Arzneimittel 167
Tabelle 4-73: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – ClinicalTrials.gov 169
Tabelle 4-74: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – EU-CTR 170
Tabelle 4-75: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel - ICTRP 171
Tabelle 4-76 (Anhang): Studiendesign und -methodik für Studie PERFECT-1 (18251) 176
Tabelle 4-77 (Anhang): Studiendesign und -methodik für Studie PERFECT-2 (18252) 187
Tabelle 4-78 (Anhang): Studiendesign und -methodik für Studie SATISFY (18250) 198
Tabelle 4-79 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PERFECT-1 206
Tabelle 4-80 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PERFECT-2 217

Abbildungsverzeichnis

Seite

Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 50
Abbildung 4-2: Design der Studien PERFECT-1 und PERFECT-2 60
Abbildung 4-3: Meta-Analyse für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA aus RCT; Trifaroten vs. Vehikel 70
Abbildung 4-4: Meta-Analyse für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA aus RCT; Trifaroten vs. Vehikel 75
Abbildung 4-5: Meta-Analyse für den Endpunkt absolute Änderung der Anzahl der Läsionen im Gesicht aus RCT; Trifaroten vs. Vehikel 82
Abbildung 4-6: Meta-Analyse für den Endpunkt prozentuale Änderung der Anzahl der Läsionen im Gesicht aus RCT; Trifaroten vs. Vehikel 83
Abbildung 4-7: Meta-Analyse für den Endpunkt absolute Änderung der Anzahl der Läsionen am Rumpf aus RCT; Trifaroten vs. Vehikel 89
Abbildung 4-8: Meta-Analyse für den Endpunkt prozentuale Änderung der Anzahl der Läsionen am Rumpf aus RCT; Trifaroten vs. Vehikel 90
Abbildung 4-9: Meta-Analyse für den Endpunkt Selbsteinschätzung der Verbesserung der Akne aus RCT; Trifaroten vs. Vehikel 95
Abbildung 4-10: Meta-Analyse für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI aus RCT; Trifaroten vs. Vehikel 101
Abbildung 4-11: Meta-Analyse für den Endpunkt gesundheitsbezogene Lebensqualität gemäß CDLQI aus RCT; Trifaroten vs. Vehikel 102
Abbildung 4-12: Meta-Analyse für den Endpunkt jegliche UE aus RCT; Trifaroten vs. Vehikel 114
Abbildung 4-13: Meta-Analyse für den Endpunkt SUE aus RCT; Trifaroten vs. Vehikel 115
Abbildung 4-14: Meta-Analyse für den Endpunkt Therapieabbrüche wegen UE aus RCT; Trifaroten vs. Vehikel 116
Abbildung 4-15: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Rötung 121
Abbildung 4-16: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Schuppung 122
Abbildung 4-17: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Trockenheit 123
Abbildung 4-18: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Stechen/Brennen 124
Abbildung 4-19: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Rötung 125
Abbildung 4-20: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Schuppung 126
Abbildung 4-21: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Trockenheit 127
Abbildung 4-22: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Stechen/Brennen 128
Abbildung 4-23: Ergänzende Ergebnisse zur lokalen Verträglichkeit aus SATISFY im Zeitverlauf 130

Abbildung 4-24: Patientenfluss der Studie PERFECT-1 186
Abbildung 4-25: Patientenfluss der Studie PERFECT-2 197
Abbildung 4-26: Patientenfluss der Studie SATISFY 204

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ANCOVA	*Analysis of Covariance*
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
CDLQI	*Children´s Dermatology Life Quality Index*
CHMP	Ausschuss für Humanarzneimittel *(Committee for Medicinal Products for Human Use)*
CMH	Cochran-Mantel-Haenszel
CONSORT	*Consolidated Standards of Reporting Trials*
CTCAE	*Common Terminology Criteria for Adverse Events*
DIMDI	Deutsches Institut für Medizinische Dokumentation
DLQI	*Dermatology Life Quality Index*
EG	Europäische Gemeinschaft
EMA	*European Medicines Agency*
EUCTR	*European Union Clinical Trials Register*
FDA	*Food and Drug Administration*
ICH	*International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use*
ICTRP	*International Clinical Trials Registry Platform*
IGA	*Investigator´s Global Assessment*
IRT	*Interactive Response Technology*
ITT	*Intention to treat*
ITTT	*ITT on the trunk*
LOCF	*Last Observation Carried Forward*
LSM	*Least-Squares Mean*
LSMD	*Least-Squares Mean Difference*
MedDRA	*Medical Dictionary for Regulatory Activities*
MI	Multiple Imputation
MMRM	*Mixed effect Model Repeat Measurement*
MNAR	*Missing not at random*
MTC	*Mixed Treatment Comparison*
PGA	*Physician´s Global Assessment*
PIP	Pädiatrischer Prüfplan (*Paediatric Investigation Plan*)

PT	*Preferred Terms* nach MedDRA
RCT	*Randomized Controlled Trial*
SAF	*Safety*
SAFT	*Safety on the Trunk*
SAP	Statistischer Analyseplan
SD	Standardabweichung (*Standard Deviation*)
SE	Standardfehler (*Standard Error*)
SGB V	Fünftes Sozialgesetzbuch
SMQs	*Standardised MedDRA Queries*
SOC	Systemorganklasse (*System Organ Class*) nach MedDRA
SUE	Schwerwiegendes UE
TREND	*Transparent Reporting of Evaluations with Non-Randomized Design*
UE	Unerwünschtes Ereignis
WHO	*World Health Organization*

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

Die Fragestellung im vorliegenden Nutzendossier erstreckt sich auf den Zusatznutzen von Trifaroten im Vergleich zur zweckmäßigen Vergleichstherapie bei der Therapie der mittelschweren *Acne vulgaris* im Gesicht und/oder am Rumpf hinsichtlich patientenrelevanter Endpunkte im Rahmen von RCT.

Trifaroten ist zugelassen zur lokalen Therapie der *Acne vulgaris* im Gesicht und/oder am Rumpf bei Patienten ab zwölf Jahren, wenn viele Komedonen, Papeln und Pusteln vorhanden sind. Bei dieser Patientenpopulation handelt es sich um Patienten mit einer mittelschweren Akne. Trifaroten ist eine Creme mit der Wirkstärke 50 µg/g, die einmal täglich im Gesicht und/oder am Rumpf angewendet wird.

Der Gemeinsame Bundesausschuss (G-BA) sieht als zweckmäßige Vergleichstherapie für Trifaroten im vorliegenden Anwendungsgebiet eine topische Kombinationstherapie aus Adapalen und Benzoylperoxid oder eine topische Kombinationstherapie aus Clindamycin und Benzoylperoxid vor.

Datenquellen

Dazu sollen methodisch hochwertige, randomisierte, kontrollierte klinische Studien (RCT, Evidenzstufe 1b) eingeschlossen werden.

Relevante Studien wurden zum einen über eine bibliographische Literaturrecherche in den Literaturdatenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken, zum anderen über eine Suche in den Studienregistern *ClinicalTrials.gov*, *International Clinical Trials Registry Platform* (ICTRP) und *EU Clinical Trials Register* (EUCTR) identifiziert.

Ein-/Ausschlusskriterien für Studien

Es wurden Ein- und Ausschlusskriterien für Studien definiert, die für die Beantwortung der oben beschriebenen Fragestellung geeignet sind. Die Selektionskriterien beziehen sich auf die Patientenpopulation, die Intervention, die Vergleichstherapie, die Endpunkte, das Studiendesign sowie den Publikationstyp. Eine Übersicht der Ein- und Ausschlusskriterien findet sich in Tabelle 4-2. Bei der Auswahl von Studien wurde das jeweilige Einschlusskriterium als erfüllt angesehen, wenn es bei mindestens 80 % der in der Studie eingeschlossenen Patienten vorlag.

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

Zur Bewertung der Aussagekraft der Nachweise wurden Angaben zu Verzerrungsaspekten sowohl auf Studienebene als auch auf Endpunktebene mithilfe der Bewertungsbögen in Anhang 4-F dokumentiert. Als Quellen wurden die Studienberichte der dargestellten Studien herangezogen.

Entsprechend den Vorgaben der Modulvorlage wurde auf Studienebene und auf Endpunktebene das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft. Falls bereits die Einstufung des Verzerrungspotenzials auf Studienebene als „hoch“ erfolgte, wurde das Verzerrungspotenzial für den betreffenden Endpunkt ebenfalls als „hoch“ bewertet.

Das Design und die Methodik der eingeschlossenen Studien wurden ausführlich in Abschnitt 4.3.1.2 und in Anhang 4-E dokumentiert. Zur Bewertung des Nutzens von Trifaroten werden die RCT PERFECT-1 und PERFECT-2 herangezogen. Die Darstellung dieser Studien erfolgte nach den Anforderungen des CONSORT-Statements (Items 2b bis 14 und CONSORT-Flow-Chart). Ergänzend wird die nicht kontrollierte Studie SATISFY herangezogen und unter Berücksichtigung des TREND-Statements dargestellt.

Die Ergebnisse der patientenrelevanten Endpunkte der Studien PERFECT-1, PERFECT-2 und SATISFY sind in Abschnitt 4.3.1.3 detailliert dargestellt. Die Ergebnisse der Studien PERFECT-1 und -2 sind zudem meta-analytisch zusammengefasst.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die zulassungsbegründenden Phase-III-Studien PERFECT-1 und -2 zeigen den medizinischen Nutzen von Trifaroten im Vergleich zu einem Vehikel. Ergänzend wird die unkontrollierte Phase-III-Studie SATISFY dargestellt, um Ergebnisse zur Langzeitwirksamkeit und -sicherheit von Trifaroten zu zeigen.

In der folgenden Tabelle 4-1 sind die Ergebnisse aus den Studien PERFECT-1 und -2 zur Wirksamkeit und Sicherheit von Trifaroten im Vergleich zur Behandlung mit Vehikel sowie die Ergebnisse der Meta-Analyse zusammengefasst.

Tabelle 4-1: Zusammenfassung der Ergebnisse aus PERFECT-1 und -2

Endpunktkategorie **Endpunkt**	**Trifaroten vs. Vehikel**	
	PERFECT-1 **PERFECT-2** **Effektschätzer [95 %-KI] p-Wert**	**Meta-Analyse** **Effektschätzer [95 %-KI]**
Morbidität		
Therapieerfolgsrate im Gesicht (IGA 0/1)	RR: 1,53 [1,24; 1,87] p < 0,001 RR: 1,71 [1,45; 2,02] p < 0,001	RR: 1,64 [1,44; 1,86]
Therapieerfolgsrate am Rumpf (PGA 0/1)	RR: 1,41 [1,18; 1,69] p < 0,001 RR: 1,50 [1,29; 1,75] p < 0,001	RR: 1,46 [1,30; 1,64]
Anzahl der Läsionen im Gesicht:		
Absolute Änderung entzündlicher Läsionen	LSMD: -3,6 [-4,9; -2,2] p < 0,001 LSMD: -5,6 [-6,9; -4,3] p < 0,001	LSMD: -4,64 [-5,57; -3,70]

Endpunktkategorie Endpunkt	Trifaroten vs. Vehikel	
	PERFECT-1 PERFECT-2 Effektschätzer [95 %-KI] p-Wert	Meta-Analyse Effektschätzer [95 %-KI]
Prozentuale Änderung entzündlicher Läsionen	LSMD: -9,6 [-13,5; -5,7] p < 0,001 LSMD: -16,0 [-19,7; -12,4] p < 0,001	LSMD: -13,01 [-15,68; -10,35]
Absolute Änderung nicht entzündlicher Läsionen	LSMD: -7,1 [-9,4; -4,8] p < 0,001 LSMD: -8,5 [-10,3; -6,6] p < 0,001	LSMD: -7,95 [-9,39; -6,51]
Prozentuale Änderung nicht entzündlicher Läsionen	LSMD: -14,0 [-18,1; -9,8] p < 0,001 LSMD: -15,5 [-19,0; -12,0] p < 0,001	LSMD: -14,88 [-17,55; -12,20]
Anzahl der Läsionen am Rumpf:		
Absolute Änderung entzündlicher Läsionen	LSMD: -2,5 [-4,0; -1,1] p < 0,001 LSMD: -5,7 [-7,2; -4,2] p < 0,001	LSMD: -4,05 [-5,09; -3,00]
Prozentuale Änderung entzündlicher Läsionen	LSMD: -7,5 [-11,5; -3,5] p < 0,001 LSMD: -15,1 [-20,1; -10,2] p < 0,001	LSMD: -10,50 [-13,61; -7,39]
Absolute Änderung nicht entzündlicher Läsionen	LSMD: -4,1 [-6,6; -1,7] p < 0,001 LSMD: -5,0 [-6,8; -3,3] p < 0,001	LSMD: -4,70 [-6,12; -3,27]
Prozentuale Änderung nicht entzündlicher Läsionen	LSMD: -8,7 [-13,3; -4,1] p < 0,001 LSMD: -12,4 [-16,3; -8,5] p < 0,001	LSMD: -10,85 [-13,83; -7,88]
Selbsteinschätzung der Verbesserung der Akne (vollständige oder deutliche Verbesserung)	RR: 1,57 [1,30; 1,88] p < 0,001 RR: 1,56 [1,32; 1,83] p < 0,001	RR: 1,56 [1,38; 1,77]
Gesundheitsbezogene Lebensqualität		
DLQI 0/1	RR: 0,95 [0,79; 1,13] p = 0,549 RR: 0,90 [0,74; 1,09] p = 0,274	RR: 0,93 [0,81; 1,06]
CDLQI 0/1	RR: 0,84 [0,67; 1,04] p = 0,110 RR: 0,82 [0,66; 1,02] p = 0,072	RR: 0,83 [0,71; 0,97]
Sicherheit/Verträglichkeit		
Jegliche UE	RR: 1,63 [1,34; 1,97] p < 0,001 RR: 1,05 [0,84; 1,32] p = 0,665	- (1)

Endpunktkategorie Endpunkt	Trifaroten vs. Vehikel	
	PERFECT-1 PERFECT-2 Effektschätzer [95 %-KI] p-Wert	Meta-Analyse Effektschätzer [95 %-KI]
SUE	RR: 1,92 [0,35; 10,42] p = 0,687 RR: 0,50 [0,09; 2,75] p = 0,687	RR: 0,98 [0,29; 3,28]
Therapieabbrüche wegen UE	RR: 13,41 [1,77; 101,66] p < 0,001 RR: 10,10 [1,30; 78,65] p = 0,006	RR: 11,66 [2,76; 49,27]
Lokale Verträglichkeit	*nur deskriptiv dargestellt*	

(1) Aufgrund der vorliegenden Heterogenität ist die Bildung eines gemeinsamen Effektschätzers nicht sinnvoll.

CDLQI: *Children´s Dermatology Life Quality Index*; DLQI: *Dermatology Life Quality Index*; IGA: *Investigator´s Global Assessment*; KI: Konfidenzintervall; LSMD: *least squares mean difference*; PGA: *Physician´s Global Assessment*; RR: Relatives Risiko; SUE: schwerwiegende unerwünschte Ereignisse; UE: unerwünschte Ereignisse

Morbidität

Die Wirksamkeit von Trifaroten wird anhand verschiedener Morbiditätsendpunkte der Symptomverbesserung dargestellt. Als Endpunkte wurden die Therapieerfolgsrate im Gesicht gemäß *Investigator's Global Assessment* (IGA), die Therapieerfolgsrate am Rumpf gemäß *Physician´s Global Assessment* (PGA), die Reduktion der Anzahl der entzündlichen und nichtentzündlichen Läsionen im Gesicht und am Rumpf sowie die von den Studienteilnehmern selbst erfahrene Verbesserung des Hautbildes erhoben.

Die Ergebnisse der Phase-III-Studien PERFECT-1 und -2 belegen die Wirksamkeit der Behandlung mit Trifaroten. Für alle untersuchten Morbiditätsendpunkte zeigt sich in beiden Studien eine statistisch signifikante Überlegenheit von Trifaroten gegenüber Vehikel in der Verringerung der Krankheitssymptome im Gesicht und am Rumpf; diese bestätigt sich jeweils auch in der meta-analytischen Zusammenfassung der Studienergebnisse.

Signifikant mehr Patienten weisen bereits nach zwölf Wochen unter Trifaroten eine vollständig oder nahezu vollständig abgeheilte Haut im Gesicht (IGA 0/1) und am Rumpf (PGA 0/1) auf. Die Langzeitstudie SATISFY zeigt über weitere 52 Wochen, dass der Anteil der Patienten mit einem Therapieerfolg gemäß IGA bzw. PGA, das heißt einer vollständigen oder fast vollständigen Abheilung der Haut im Gesicht und am Rumpf, konstant steigt. Die Wirksamkeit der Behandlung mit Trifaroten ist somit langanhaltend und nimmt bei den Patienten, die nach zwölf Wochen noch nicht den maximalen Therapieeffekt erreicht haben, weiter zu.

Dem Ergebnis der Therapieerfolgsraten gemäß IGA bzw. PGA entsprechend verringert sich die Anzahl der entzündlichen und der nicht entzündlichen Läsionen im Gesicht und am Rumpf innerhalb der zwölfwöchigen Behandlungsphase signifikant stärker unter Trifaroten.

Diese Symptomfreiheit bzw. -reduktion wird auch durch den Patienten selbst so empfunden: Signifikant mehr Patienten unter Trifaroten schätzen die Verbesserung ihrer Akne nach zwölf Wochen Behandlung als vollständige oder deutliche Verbesserung ein.

Gesundheitsbezogene Lebensqualität

Die Ergebnisse der Studien PERFECT-1 und -2 zur Veränderung der gesundheitsbezogenen Lebensqualität gemäß DLQI zeigen nach zwölf Wochen Behandlung für die Patienten ab 17 Jahren keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen. Für die jüngeren Studienteilnehmer, die den CDLQI zur Erhebung der gesundheitsbezogenen Lebensqualität nutzten, ergibt die Meta-Analyse einen statistisch signifikanten Nachteil für Trifaroten hinsichtlich des Anteils der Patienten, die nach zwölf Wochen Behandlung einen CDLQI 0/1 erreichen. Die Einzelergebnisse der Studien zeigen keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen.

Die mittlere Beeinträchtigung der Lebensqualität bei den jüngeren Studienteilnehmern ist zu Beginn der Studien gering und eine starke Verbesserung infolge der Behandlung ist nicht zu erwarten. Weiterhin ist die Akne am Rumpf bei Kindern und Jugendlichen im frühen Alter wenig ausgeprägt. Patienten, bei denen die Akne auf das Gesicht beschränkt ist, zeigen eine geringere Beeinträchtigung der gesundheitsbezogenen Lebensqualität als Patienten mit großflächiger Akne im Gesicht und am Rumpf.

Kinder und Jugendliche nehmen Akne nicht als chronische Erkrankung der Haut mit dem Risiko der Entstehung bleibender Narben wahr, die man medikamentös behandeln kann und sollte. Die Behandlung der mittelschweren Akne mit einer aktiven Substanz wie Trifaroten ist im Gegensatz zur Behandlung mit Vehikel mit transienten und überwiegend milden oder moderaten Nebenwirkungen der Haut verbunden. Es ist anzunehmen, dass bei Patienten, die ihre Akne nicht als behandlungswürdige Hautveränderung wahrnehmen, das temporäre Auftreten von Nebenwirkungen die Bewertung der gesundheitsbezogenen Lebensqualität stärker beeinflusst, als das Empfinden der Verbesserung einer chronischen Erkrankung der Haut. Daraus könnte der beobachtete Nachteil für Trifaroten in dieser Subpopulation resultieren.

Sicherheit und Verträglichkeit

Die Auswertungen zu den dokumentierten unerwünschten Ereignissen (UE) belegen die Sicherheit von Trifaroten.

Hinsichtlich jeglicher UE ergibt sich in der Studie PERFECT-1 ein statisch signifikanter Nachteil für Trifaroten gegenüber Vehikel. In der Studie PERFECT-2 zeigt sich kein Unterschied in der Häufigkeit jeglicher UE. Die Art der auftretenden UE entspricht dem erwarteten Nebenwirkungsprofil. Die häufigsten UE sind Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“, dabei wurden insbesondere die PT „Reizung an der Applikationsstelle“ und „Pruritus an der Applikationsstelle“ dokumentiert. Diese Ereignisse treten unter Trifaroten häufiger auf als unter Vehikel. Weiterhin sind Ereignisse der SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ unter Trifaroten häufiger als unter Vehikel.

SUE sind sowohl unter Vehikel als auch unter Trifaroten sehr selten, und keines der schwerwiegenden Ereignisse wird als mit der Studienmedikation in Zusammenhang stehend eingestuft. Es besteht kein statistisch signifikanter Unterschied zwischen der Behandlung mit Trifaroten und Vehikel im Hinblick auf das Auftreten von SUE.

Therapieabbrüche aufgrund von UE treten häufiger unter Trifaroten auf, sind jedoch in beiden Studien selten. Es besteht ein statistisch signifikanter Nachteil für Trifaroten. Am häufigsten führen Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ und insbesondere das Ereignis „Reizung an der Applikationsstelle“ zu einem Abbruch der Behandlung mit Trifaroten.

Hautreaktionen bzw. Hautirritationen im Gesicht und am Rumpf treten häufiger bei Anwendung von Trifaroten im Vergleich zur Kontrollgruppe auf. Die Ereignisse sind jedoch in den meisten Fällen von milder oder moderater Schwere. Zudem sind die auftretenden Hautreaktionen in der überwiegenden Mehrheit der Fälle transient und klingen im Laufe der Behandlung schnell ab.

Die Daten der Langzeitstudie SATISFY bestätigen das gute Sicherheitsprofil von Trifaroten. Die Mehrheit der aufgetretenen UE wurde in den ersten drei Monaten der Behandlung dokumentiert. Insbesondere die Therapieabbrüche aufgrund von UE treten fast alle in den ersten drei Behandlungsmonaten auf. Bei langfristiger Anwendung von Trifaroten kommt es folglich nicht zu einer Zunahme der UE und nicht zu einem gehäuften Auftreten von UE, die einen Abbruch der Anwendung von Trifaroten erfordern. Somit ist die langfristige Behandlung mit Trifaroten und insbesondere die großflächige Anwendung am Rumpf über einen Zeitraum von 52 Wochen nachgewiesenermaßen gut verträglich und sicher.

Zusammenfassung der Aussagen zum therapeutisch relevanten Nutzen

Die Studien PERFECT-1 und -2 belegen mit hoher Aussagekraft die signifikante Überlegenheit von Trifaroten gegenüber einer Behandlung mit Vehikel im Hinblick auf die Verringerung der Krankheitslast durch eine ausgeprägte und klinisch relevante Verringerung der Symptome, die mit der Erkrankung *Acne vulgaris* einhergehen. Die Wirksamkeit der Behandlung mit Trifaroten ist über den Zeitraum von zwölf Wochen hinaus anhaltend bzw. nimmt bei den Patienten, die nach zwölf Wochen noch nicht den maximalen Therapieeffekt erreicht haben, weiter zu.

Die unter Trifaroten auftretenden UE entsprechen dem erwarteten Nebenwirkungsprofil. Die häufigsten UE sind Reaktionen der Haut, die überwiegend von milder bis moderater Schwere sind und nur transient auftreten. Bei der Anwendung von Trifaroten über 52 Wochen kommt es nicht zu einer relevanten Zunahme von Nebenwirkungen. Somit ist die Sicherheit und Verträglichkeit einer langfristigen Anwendung von Trifaroten und insbesondere einer großflächigen Anwendung am Rumpf bestätigt.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Die Studien PERFECT-1 und -2 zeigen die Wirksamkeit und Sicherheit von Trifaroten im Vergleich zu einem Vehikel. Das Ausmaß des Zusatznutzens im Vergleich zu der zweckmäßigen Vergleichstherapie ist daraus nicht abzuleiten und zu quantifizieren. Dennoch gibt es Anhaltspunkte für einen Zusatznutzen von Trifaroten gegenüber der zweckmäßigen Vergleichstherapie, der zum einen in den vorliegenden Daten und zum anderen in der zweckmäßigen Vergleichstherapie und den weiteren Therapieoptionen begründet ist.

Die Evidenz zur Wirksamkeit und Sicherheit von Trifaroten aus den beiden Phase-III-Zulassungsstudien belegt die klinische Wirksamkeit von Trifaroten hinsichtlich der Verringerung der Aknesymptomatik bei gleichzeitig guter Verträglichkeit. Insbesondere liefern die Studiendaten erstmalig hochwertige Evidenz für die Wirksamkeit eines topischen Retinoids bei der Behandlung der Akne am Rumpf. Die Daten zeigen die schnell einsetzende klinische Wirkung im Hinblick auf die Abheilung der Haut am Rumpf. Gleichzeitig belegen die Daten aus der Studie SATISFY die gute Verträglichkeit der Behandlung mit Trifaroten auch bei täglicher großflächiger Anwendung am Rumpf über einen Zeitraum von 52 Wochen.

Die Datenlage zur Behandlung der Akne am Rumpf ist für die bisher zugelassenen Therapieoptionen sehr begrenzt. Für keine der in Deutschland zugelassenen Therapien liegen Daten aus randomisierten klinischen Studien zur Behandlung der Akne am Rumpf vor. Die Zulassung der bisher verfügbaren Wirkstoffe ist jeweils auf Basis klinischer Studien an Patienten mit Akne im Gesicht erfolgt. Dies spiegelt sich auch in den mangelnden Leitlinien-Empfehlungen zur Behandlung der Akne am Rumpf wider. Trifaroten ist das einzige topische Retinoid, welches aufgrund seiner nachgewiesenen Wirksamkeit und Sicherheit explizit zur großflächigen Behandlung der Akne im Gesicht und am Rumpf zugelassen ist. Dies stellt einen Zusatznutzen gegenüber anderen Therapieoptionen inklusive der zweckmäßigen Vergleichstherapie dar.

Die Behandlung der Akne am Rumpf erfolgt bislang häufig mittels einer systemischen anstelle einer topischen Therapie, obwohl dies bei Vorliegen einer mittelschweren Akne als eine Übertherapie anzusehen ist. Dieser Therapieansatz stellt im Falle des systemischen Einsatzes von Isotretinoin und der meisten Antibiotika keine zulassungskonforme Anwendung dar und ist mit dem Risiko schwerwiegender therapiebedingter Nebenwirkungen bzw. der Entwicklung bakterieller Resistenzen verbunden. Durch die Verfügbarkeit eines neuen Wirkstoffes zur Behandlung der mittelschweren Akne kann – insbesondere bei der großflächigen Behandlung von Akne am Rumpf – potentiell der systemische Einsatz von Isotretinoin und Antibiotika reduziert werden.

Einen grundsätzlichen Zusatznutzen hat Trifaroten gegenüber der zweckmäßigen Vergleichstherapie Clindamycin und Benzoylperoxid in der Vermeidung einer topischen antibiotikahaltigen Therapie. Der Einsatz von topischen Antibiotika wie Clindamycin ist mit dem Risiko einer bakteriellen Resistenzentwicklung verbunden. Das potentielle Resistenzrisiko korreliert mit der Dauer der Anwendung sowie dem Umfang der behandelten Hautoberfläche.

Aus diesem Grund ist die Anwendung von Clindamycin und Benzoylperoxid laut Fachinformation auf zwölf Wochen zu beschränken und gemäß der Fachinformation sind die Richtlinien zur sachgemäßen Anwendung antimikrobieller Arzneimittel sowie verfügbare lokale Informationen zur Prävalenz von bakteriellen Resistenzen zu beachten. Auch die europäische S3-Leitlinie zur Behandlung der Akne weist explizit auf das potentielle Risiko von Antibiotika-Resistenzen hin. Folglich stellt die Verfügbarkeit einer wirksamen und gut verträglichen Therapiealternative zu einer antibiotikahaltigen Therapie per se einen Zusatznutzen dar.

Über die genannten Vorteile hinaus sind weitere positive Aspekte zu nennen: Durch den spezifischeren Wirkmechanismus von Trifaroten im Vergleich zu anderen Retinoiden ist eine Verringerung des Risikos für Nebenwirkungen zu erwarten. Des Weiteren hat Trifaroten im Gegensatz zu den Benzoylperoxid-haltigen Arzneimitteln der zweckmäßigen Vergleichstherapie keine bleichende Wirkung auf Haare und Textilien, was insbesondere für Patienten mit Akne am Rumpf relevant ist und potentiell zu einer besseren Therapieadhärenz führen kann.

Zusammenfassend besteht für Trifaroten ein **Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen**. Dieser basiert zum einen auf dem patientenrelevanten therapeutischen Nutzen, insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, bei Patienten mit mittelschwerer *Acne vulgaris* im Gesicht und/oder am Rumpf, welcher für die Anwendung am Rumpf erstmalig für ein Retinoid im Rahmen von randomisierten, kontrollierten klinischen Studien belegt wurde. Zum anderen ermöglicht Trifaroten als wirksame und gut verträgliche antibiotikafreie, topische Therapiealternative die Vermeidung einer topischen oder systemischen Antibiotika-Therapie sowie die Vermeidung der systemischen Anwendung von Isotretinoin.

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).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

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.

Die Fragestellung erstreckt sich auf den Zusatznutzen von Trifaroten im Vergleich zur zweckmäßigen Vergleichstherapie bei der Therapie der mittelschweren *Acne vulgaris* im Gesicht und/oder am Rumpf hinsichtlich patientenrelevanter Endpunkte im Rahmen von RCT.

Trifaroten ist zugelassen zur lokalen Therapie der *Acne vulgaris* im Gesicht und/oder am Rumpf bei Patienten ab zwölf Jahren, wenn viele Komedonen, Papeln und Pusteln vorhanden sind [1]. Bei dieser Patientenpopulation handelt es sich um Patienten mit einer mittelschweren Akne. Trifaroten ist eine Creme mit der Wirkstärke 50 µg/g, die einmal täglich im Gesicht und/oder am Rumpf angewendet wird.

Der G-BA sieht als zweckmäßige Vergleichstherapie für Trifaroten im vorliegenden Anwendungsgebiet eine topische Kombinationstherapie aus Adapalen und Benzoylperoxid oder eine topische Kombinationstherapie aus Clindamycin und Benzoylperoxid vor [2].

Die Beurteilung des Zusatznutzens geht aus den patientenrelevanten Endpunkten der Kategorien Morbidität, gesundheitsbezogene Lebensqualität und Sicherheit hervor. Die Endpunktkategorie Mortalität ist für das vorliegende Anwendungsgebiet nicht relevant. Es

sollen methodisch hochwertige, randomisierte, kontrollierte klinische Studien (RCT, Evidenzstufe 1b) eingeschlossen werden. Die RCT müssen mindestens einen patientenrelevanten Endpunkt enthalten.

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/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt 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. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Es wurden Ein- und Ausschlusskriterien zur Identifikation von Studien definiert, die für die Beantwortung der Fragestellung zum Zusatznutzen von Trifaroten geeignet sind. Die Selektionskriterien beziehen sich auf die Patientenpopulation, die Intervention, die Vergleichstherapie, die Endpunkte, das Studiendesign sowie den Publikationstyp. Eine Übersicht der Ein- und Ausschlusskriterien findet sich in Tabelle 4-2.

Bei der Auswahl von Studien wurde das jeweilige Einschlusskriterium als erfüllt angesehen, wenn es bei mindestens 80 % der in der Studie eingeschlossenen Patienten vorlag.

Tabelle 4-2: Ein- und Ausschlusskriterien für RCT mit Trifaroten und der zweckmäßigen Vergleichstherapie

		Einschlusskriterium (E)	**Ausschlusskriterium (A)**
1	Patienten-population	Patienten ≥ 12 Jahre mit mittelschwerer *Acne vulgaris* im Gesicht und/oder am Rumpf	Die Studienpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
2	Intervention	Trifaroten in der für Deutschland zugelassenen Dosierung und Anwendungsart	Die Intervention ist nicht Trifaroten in der für Deutschland zugelassenen Dosierung und Anwendungsart.
3	Vergleichs-therapie	Kombinationstherapie Adapalen und Benzoylperoxid oder Kombinationstherapie Clindamycin und Benzoylperoxid in der für Deutschland zugelassenen Dosierung und Anwendungsart	Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
4	Endpunkte	Mindestens ein patientenrelevanter Endpunkt wird berichtet.	Kein patientenrelevanter Endpunkt wird berichtet.
5	Studien-design	Die Studie ist eine randomisierte, kontrollierte Studie.	Die Studie ist keine randomisierte, kontrollierte Studie.
6	Studien-dauer	≥ 12 Wochen	< 12 Wochen
7	Ergebnis-darstellung	Eine Vollpublikation, ein Studienbericht oder ein ausführlicher Ergebnisbericht ist verfügbar.	Nur ein Review, Abstract oder Kongress-Poster ist verfügbar.

Erläuterungen zu den Ein- und Ausschlusskriterien

Patientenpopulation (Ein-/Ausschlusskriterium 1):

Gemäß zugelassenem Anwendungsgebiet von Trifaroten wurden Studien mit Jugendlichen ab zwölf Jahren oder Erwachsenen mit einer mittelschweren *Acne vulgaris* im Gesicht und/oder am Rumpf eingeschlossen [1].

Intervention (Ein-/Ausschlusskriterium 2):

Es wurden Studien eingeschlossen, in denen Trifaroten in der empfohlenen Dosierung und Art der Anwendung laut Fachinformation eingesetzt wurde. Trifaroten ist als Creme mit der Wirkstärke 50 µg/g zugelassen und wird einmal täglich im Gesicht und/oder am Rumpf aufgetragen [1].

Im Rahmen der klinischen Prüfungen wurde Trifaroten in unterschiedlichen Formulierungen untersucht. Dabei wurden verschiedene Cremes sowie Gele als Vehikel verwendet. Bei der Suche nach Studien für die Nutzenbewertung wurden ausschließlich Studien berücksichtigt, in denen Trifaroten in der zugelassenen Formulierung untersucht wurde.

Vergleichstherapie (Ein-/Ausschlusskriterium 3):

Gemäß den Vorgaben des G-BA hinsichtlich der zweckmäßigen Vergleichstherapie wurden als Komparatoren die topischen Kombinationstherapien Adapalen und Benzoylperoxid sowie

Clindamycin und Benzoylperoxid in den für Deutschland zugelassenen Dosierungen berücksichtigt [2].

Dabei ist zu beachten, dass in Deutschland zwei Arzneimittel mit den Wirkstoffen Adapalen und Benzoylperoxid [3, 4] sowie zwei Arzneimittel mit den Wirkstoffen Clindamycin und Benzoylperoxid [5, 6] zugelassen sind, die sich jeweils hinsichtlich der Wirkstärken unterscheiden. Bei der Suche nach Studien für die Nutzenbewertung wurden alle zugelassenen Wirkstärken berücksichtigt.

Endpunkte (Ein-/Ausschlusskriterium 4):

Es wurden Studien eingeschlossen, die mindestens einen patientenrelevanten Endpunkt berichten, da die Nutzenbewertung auf Basis von patientenrelevanten Endpunkten erfolgt. Nähere Angaben zu patientenrelevanten Endpunkten in der Indikation Akne finden sich im Abschnitt 4.2.5.2.

Studiendesign (Ein-/Ausschlusskriterium 5):

Es wurden ausschließlich randomisierte, kontrollierte, doppelblinde Studien eingeschlossen, da die Nutzenbewertung auf Basis der höchsten Evidenz erfolgt.

Studiendauer (Ein-/Ausschlusskriterium 6):

Die erforderliche Studiendauer sollte jeweils in Abhängigkeit von der Indikation, der Patientenpopulation, dem zu bewertenden Arzneimittel und der jeweiligen Fragestellung festgelegt werden.

Bei der *Acne vulgaris* handelt es sich um eine chronische Erkrankung. Bei chronischen Erkrankungen, die in der Regel eine längerfristige Behandlung erfordern, wird im Rahmen der frühen Nutzenbewertung im Allgemeinen eine Mindeststudiendauer von 24 Wochen als notwendig angesehen, um Aussagen zum Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie treffen zu können.

Klinische Studien zur Untersuchung topischer Arzneimittel zur Behandlung der Akne werden grundsätzlich über einen deutlich kürzeren Zeitraum, üblicherweise zwölf Wochen, durchgeführt, um die Wirksamkeit und Sicherheit der Initialtherapie zu belegen. Ebenso grundsätzlich erfolgen sodann Langzeitstudien, die eine Sicherheit und Wirksamkeit in der Dauertherapie zeigen.

In der Indikation Akne ist eine Studiendauer von zwölf Wochen angemessen, um die Wirksamkeit und Sicherheit einer topischen Therapie zu bewerten: Die europäischen Leitlinien empfehlen bei der Behandlung der Akne zunächst eine sogenannte Induktionstherapie, an welche sich nach Erreichen des Therapieziels eine Erhaltungstherapie anschließen sollte (siehe auch Modul 3A, Abschnitt 3.2.1) [7, 8]. Für Patienten mit einer mittelschweren Akne wird empfohlen, das Ansprechen auf die topische Behandlung acht Wochen nach Therapiebeginn zu prüfen. Bei Patienten, die nach dieser Zeit noch keine Besserung ihres Hautbildes zeigen,

sollten die Gründe für das Ausbleiben des Therapieerfolgs geprüft und die Behandlung gegebenenfalls angepasst werden [8]. Patienten, die nach acht Wochen ein Ansprechen auf die Therapie zeigen, sollten die Behandlung mindestens bis Woche 12 fortsetzen [8]. Sobald eine vollständige oder nahezu vollständige Abheilung der Haut erreicht ist, sollte von der Induktions- auf eine Erhaltungstherapie gewechselt werden. Eine Studie mit einer zwölfwöchigen Behandlungsdauer bildet somit den leitlinienkonformen Therapiealgorithmus für eine Induktionstherapie bzw. einen typischen, ggf. zu wiederholenden Behandlungszyklus in dieser Indikation ab.

Sowohl bei dem zu bewertenden Arzneimittel Trifaroten als auch bei den Fixkombinationen der zweckmäßigen Vergleichstherapie ist ein schneller Wirkeintritt zu beobachten. So zeigt sich bereits nach ein bis vier Wochen ein Ansprechen auf die Behandlung, welches durch eine signifikante Verringerung der Anzahl der Läsionen sichtbar wird [1, 3, 5, 6]. Auch wenn nach zwölf Wochen Behandlung noch nicht bei allen Patienten der maximale Therapieeffekt erreicht ist, so ist dennoch eine Vergleichbarkeit der Therapien zu diesem Zeitpunkt gegeben.

Ein weiterer Aspekt ist die Beschränkung der Anwendungsdauer der vom G-BA festgesetzten zweckmäßigen Vergleichstherapie Clindamycin und Benzoylperoxid. Die zugelassenen Fixkombinationen sollten gemäß Fachinformation nicht länger als zwölf Wochen angewendet werden, um Antibiotikaresistenzen zu vermeiden [5, 6]. Eine Studie mit der zweckmäßigen Vergleichstherapie Clindamycin und Benzoylperoxid könnte also in der aktiven Behandlungsphase nicht länger als zwölf Wochen sein.

Nicht zuletzt muss bei der Festlegung der Mindeststudiendauer auch die Patientenpopulation im vorliegenden Anwendungsgebiet berücksichtigt werden. *Acne vulgaris* ist eine Erkrankung, die vorwiegend in der Pubertät auftritt. Die Zielpopulation für Trifaroten umfasst folglich neben Erwachsenen auch Jugendliche ab zwölf Jahren. Die Durchführung von klinischen Studien mit Kindern und Jugendlichen muss deren besondere Schutzbedürftigkeit berücksichtigen. Entsprechende Einschränkungen bei der Durchführung von Studien mit pädiatrischen Populationen werden in der Leitlinie ICH-E11 der europäischen Zulassungsbehörde EMA (*European Medicines Agency*) genannt. So sollen unnötige klinische Prüfungen vermieden und die Belastung der pädiatrischen Studienteilnehmer durch klinische Erhebungen im Rahmen der Studie auf das notwendige Minimum reduziert werden [9, 10]. Die regulatorischen Rahmenbedingungen für Studien an pädiatrischen Populationen sind in der europäischen Richtlinie 2001/20/EG sowie der EU-Verordnung (EG) Nr. 1901/2006 geregelt [11, 12]. Im Einklang mit diesen Richtlinien wurde das Studienprogramm für Trifaroten im Rahmen eines pädiatrischen Prüfplans (PIP) mit dem Pädiatrieausschuss der EMA abgestimmt und als adäquat erachtet (PIP-Nummer EMEA-001492-PIP01-13-M01) [13].

Zusammenfassend ist eine Studiendauer von zwölf Wochen im Hinblick auf die Indikation, die Interventionen und die Patientenpopulation angemessen und ausreichend für eine Nutzenbewertung. Somit wurden im vorliegenden Dossier Studien mit einer Dauer von mindestens zwölf Wochen berücksichtigt.

Ergebnisdarstellung (Ein-/Ausschlusskriterium 7):

Es wurden Studien berücksichtigt, für die eine Vollpublikation, ein Studienbericht oder ein vergleichbarer ausführlicher Ergebnisbericht vorliegt, um die Kriterien des CONSORT-Statements zu erfüllen und eine Einschätzung der Studienergebnisse zu ermöglichen.

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 (inklusive „in-process & other non-indexed citations) 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 0 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 bibliographische Literaturrecherche wurde am 16.06.2020 über die Oberfläche Ovid in den Literaturdatenbanken MEDLINE und EMBASE durchgeführt, weiterhin wurde eine Recherche in den Cochrane-Datenbanken durchgeführt.

Für jede Datenbank wurde eine adaptierte Suchstrategie formuliert, die in Anhang 4-A detailliert beschrieben ist. Die Suche wurde jeweils in Blöcke zur Intervention und ggf. zum Studientyp gegliedert. Auf die Einschränkung der Suchkriterien über den Publikationszeitraum oder die Sprache wurde verzichtet.

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

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

Eine Suche in Studienregistern/ Studienergebnisdatenbanken 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 sowie Studienergebnisdatenbanken 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/ Studienergebnisdatenbanken clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (https://clinicaldata.ema.europa.eu) sowie dem Arzneimittel-Informationssystem (AMIS, https://www.pharmnet-bund.de/dynamic/de/arzneimittel-

informationssystem/index.html) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken 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 bibliographischen Literaturrecherche wurde am 17.06.2020 eine Suche nach Studien in den Studienregistern *ClinicalTrials.gov* (www.clinicaltrials.gov), *International Clinical Trials Registry Platform* (ICTRP, http://apps.who.int/trialsearch/AdvSearch.aspx) und *EU Clinical Trials Register* (http://www.clinicaltrialsregister.eu/ctr-search) durchgeführt.

Im Rahmen der Suche wurden abgeschlossene, laufende und abgebrochene Studien erfasst. Es wurde keine generelle Einschränkung der Suche vorgenommen. Die detaillierte Suchstrategie ist für jedes Studienregister getrennt in Anhang 4-B dargestellt.

Das Suchportal der EMA und die AMIS-Datenbank wurden nicht nach Studien durchsucht, da keine relevanten Studien durch die bibliografische Literaturrecherche oder die Registerrecherche identifiziert wurden.

4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu

Studienmethodik und –ergebnissen[1]. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA 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 auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

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

Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Die Suche auf der Internetseite des G-BA wurde am 17.06.2020 durchgeführt. Auf der Seite des G-BA wurde im Suchfeld nach dem zu bewertenden Arzneimittel Trifaroten gesucht. Eine Suche nach spezifischen Studien wurde nicht vorgenommen, da keine relevanten Studien über die bibliografische Literaturrecherche und die Studienregistersuche identifiziert wurden.

4.2.3.5 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 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 (siehe Abschnitt 4.2.2) durchgeführt. Die einzelnen Schritte zur Selektion der Studien wurden unabhängig von zwei Personen durchgeführt.

[1] Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staeck K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. BMJ 2015;350:h796

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):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

– 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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

– 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.

Von der oben angegebenen Methodik wurde nicht abgewichen.

Zur Bewertung der Aussagekraft der Nachweise wurden Angaben zu Verzerrungsaspekten sowohl auf Studienebene als auch auf Endpunktebene mithilfe der Bewertungsbögen in Anhang 4-F gemäß den obenstehenden Vorgaben dokumentiert. Als Quelle wurden die klinischen Studienberichte der eingeschlossenen Studien herangezogen.

Entsprechend den Vorgaben der Modulvorlage wurden zunächst auf Studienebene die Informationen zu den oben unter (A) genannten Aspekten bewertet und das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft.

Falls bereits die Einstufung des Verzerrungspotenzials auf Studienebene als „hoch“ erfolgte, wurde das Verzerrungspotenzial für den betreffenden Endpunkt ebenfalls als „hoch“ bewertet. Ansonsten wurden bei der Beurteilung des Verzerrungspotenzials auf Endpunktebene die oben unter (B) genannten Aspekte bewertet und das Verzerrungspotenzial als „niedrig“ oder „hoch“ eingestuft.

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)[2]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[3] bzw. STROBE-Statements[4] folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

[2] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[3] 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.
[4] 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.

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.

Das Design und die Methodik der eingeschlossenen Studien sind ausführlich in Abschnitt 4.3.1.2 und in Anhang 4-E dokumentiert. Zur Bewertung des Nutzens von Trifaroten werden die RCT PERFECT-1 und PERFECT-2 herangezogen. Somit erfolgt die Darstellung dieser Studien nach den Anforderungen des CONSORT-Statements (Items 2b bis 14 und CONSORT-Flow-Chart). Ergänzend wird die nicht kontrollierte Studie SATISFY herangezogen. Die Darstellung dieser Studie erfolgt unter Berücksichtigung des TREND-Statements.

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 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.

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

Patientencharakteristika

In den für den Nachweis des Nutzens relevanten RCT wurden demographische und andere Patientencharakteristika erhoben. Im vorliegenden Dossier werden folgende Charakteristika in Abschnitt 4.3.1.2.1 dargestellt:

- Alter
- Geschlecht
- Abstammung

- Hauttyp
- Krankheitsschwere im Gesicht gemäß *Investigator´s Global Assessment* (IGA)
- Krankheitsschwere am Rumpf gemäß *Physician´s Global Assessment* (PGA)
- Entzündliche Läsionen im Gesicht
- Entzündliche Läsionen am Rumpf
- Nicht entzündliche Läsionen im Gesicht
- Nicht entzündliche Läsionen am Rumpf

Endpunkte

In den für die Beschreibung des Nutzens von Trifaroten herangezogenen Studien wurden Endpunkte der Kategorien Morbidität, gesundheitsbezogene Lebensqualität und Sicherheit untersucht. Mortalität war in den herangezogenen Studien kein definierter Endpunkt, Todesfälle wurden aber grundsätzlich im Rahmen der unerwünschten Ereignisse erfasst. Im Dossier werden Ergebnisse bezüglich folgender patientenrelevanter Endpunkte dargestellt:

- Morbidität
    - Therapieerfolgsrate im Gesicht (Anteil der Patienten mit IGA 0/1)
    - Therapieerfolgsrate am Rumpf (Anteil der Patienten mit PGA 0/1)
    - Änderung der Anzahl der Läsionen im Gesicht
        - Absolute Änderung der Anzahl der entzündlichen Läsionen
        - Absolute Änderung der Anzahl der nicht entzündlichen Läsionen
        - Prozentuale Änderung der Anzahl der entzündlichen Läsionen
        - Prozentuale Änderung der Anzahl der nicht entzündlichen Läsionen
    - Änderung der Anzahl der Läsionen am Rumpf
        - Absolute Änderung der Anzahl der entzündlichen Läsionen
        - Absolute Änderung der Anzahl der nicht entzündlichen Läsionen
        - Prozentuale Änderung der Anzahl der entzündlichen Läsionen
        - Prozentuale Änderung der Anzahl der nicht entzündlichen Läsionen

- Selbsteinschätzung der Verbesserung der Akne (Anteil der Patienten mit vollständiger oder deutlicher Verbesserung)

- Gesundheitsbezogene Lebensqualität
  - *Dermatology Life Quality Index* (DLQI) (Anteil der Patienten mit DLQI 0/1)
  - *Children´s Dermatology Life Quality Index* (CDLQI) (Anteil der Patienten mit CDLQI 0/1)

- Sicherheit/Verträglichkeit
  - Sicherheit
    - Gesamtrate jeglicher unerwünschter Ereignisse (UE)
    - Gesamtrate schwerwiegender unerwünschter Ereignisse (SUE)
    - Gesamtrate der Therapieabbrüche wegen UE
    - Spezifische UE
  - Lokale Verträglichkeit
    - Lokale Verträglichkeit im Gesicht (Rötung, Schuppung, Trockenheit, Brennen)
    - Lokale Verträglichkeit am Rumpf (Rötung, Schuppung, Trockenheit, Brennen)

Im Folgenden wird die Patientenrelevanz der einzelnen Endpunkte beschrieben und die Validität der verwendeten Messinstrumente diskutiert.

Morbidität

Therapieerfolgsrate im Gesicht (IGA 0/1)

Das Therapieziel bei der Behandlung der Akne im Gesicht stellt eine vollständige oder nahezu vollständige Abheilung der Haut dar. Die Freiheit bzw. die Verringerung von Krankheitssymptomen ist ein therapierelevanter Nutzen eines Arzneimittels gemäß § 5 AM-NutzenV.

Die Bestimmung des Therapieerfolgs im Gesicht erfolgt mittels der IGA-Skala. Dabei handelt es sich um eine fünfstufige Skala von 0 (abgeheilt) bis 4 (schwer). Für die einzelnen Stufen wurden morphologische Kriterien für den jeweiligen Schweregrad definiert, um eine möglichst objektive Bewertung der Krankheitsschwere eines Patienten zu ermöglichen (Details siehe Abschnitt 4.3.1.3.1, Tabelle 4-14). Die IGA-Skala ist ein valides und etabliertes Instrument zur Bestimmung der globalen Krankheitsschwere [14, 15]. Die amerikanische Arzneimittelbehörde

FDA (*Food and Drug Administration*) empfiehlt explizit den Einsatz der IGA-Skala zur Bestimmung der Krankheitsschwere im Rahmen klinischer Studien für die Zulassung neuer Aknetherapien [16]. Auch die an der Zulassung von Trifaroten beteiligten europäischen Zulassungsbehörden erkennen die IGA-Skala als validiertes Instrument zur Bestimmung der Wirksamkeit einer Aknetherapie an [17].

Als klinisch relevantes Ergebnis im Rahmen von Zulassungsstudien wird von der FDA das Erreichen eines IGA von 0 (klare Haut ohne entzündliche oder nichtentzündliche Läsionen) oder 1 (vereinzelte Komedonen und wenige kleine Papeln) und eine Verbesserung um mindestens zwei Stufen angesehen [16]. Dies entspricht der im Rahmen der Studien PERFECT-1 und -2 angewandten Definition eines Therapieerfolgs im Gesicht. Auch die europäischen Zulassungsbehörden haben die Operationalisierung dieses Endpunkts als adäquat anerkannt [17].

Zusammenfassend stellt die IGA-Skala ein valides Instrument zur Bestimmung der Krankheitsschwere im Gesicht dar. Der Endpunkt Therapieerfolgsrate im Gesicht ist grundsätzlich als patientenrelevant anzusehen und die Operationalisierung als Anteil der Patienten mit IGA 0/1 und einer gleichzeitigen Verbesserung um mindestens zwei Stufen ist eine etablierte Endpunktdefinition im Rahmen klinischer Studien.

Therapieerfolgsrate am Rumpf (PGA 0/1)

Das Therapieziel bei der Behandlung der Akne am Rumpf stellt eine vollständige oder nahezu vollständige Abheilung der Haut dar. Die Freiheit von Krankheitssymptomen bzw. die Verringerung dieser ist ein therapierelevanter Nutzen eines Arzneimittels gemäß § 5 AM-NutzenV.

Die Bestimmung des Therapieerfolgs am Rumpf erfolgt mittels der PGA-Skala. Dabei handelt es sich um eine fünfstufige Skala von 0 (abgeheilt) bis 4 (schwer) mit definierten morphologischen Kriterien für den jeweiligen Schweregrad, welche den Kriterien der oben beschriebenen IGA-Skala entsprechen (Details siehe Abschnitt 4.3.1.3.2, Tabelle 4-18). Die an der Zulassung von Trifaroten beteiligten europäischen Zulassungsbehörden haben die Verwendung der PGA-Skala zur Bestimmung der Wirksamkeit am Rumpf sowie die Operationalisierung des Endpunkts Therapieerfolg am Rumpf als adäquat anerkannt [17]. Der Endpunkt Therapieerfolg am Rumpf war in den Studien PERFECT-1 und -2 analog zum Endpunkt Therapieerfolg im Gesicht als Erreichen eines PGA 0/1 und einer Verbesserung um mindestens zwei Stufen definiert.

Zusammenfassend entspricht die PGA-Skala der IGA-Skala und stellt ein valides Instrument zur Bestimmung der Krankheitsschwere am Rumpf dar. Der Endpunkt Therapieerfolgsrate am Rumpf ist als grundsätzlich patientenrelevant anzusehen und die Operationalisierung als Anteil der Patienten mit PGA 0/1 und einer gleichzeitigen Verbesserung um mindestens zwei Stufen ist eine adäquate Endpunktdefinition im Rahmen klinischer Studien.

Anzahl der Läsionen

Neben der globalen Bewertung der Krankheitsschwere dient die Anzahl der Läsionen in den betroffenen Hautarealen als objektives Maß für die Krankheitsschwere bzw. zur Beurteilung der Wirksamkeit einer Therapie.

Die FDA empfiehlt, im Rahmen klinischer Studien neben der globalen Krankheitsschwere auch die Zahl der Läsionen als Maß für das Ansprechen auf die Therapie zu bestimmen, da diese Endpunkte qualitative und quantitative Aspekte der Erkrankung abbilden und einander ergänzende Informationen zur Beurteilung der Wirksamkeit einer Therapie liefern [16]. Zudem empfiehlt die FDA-Leitlinie, sowohl die absolute Änderung als auch die relative Änderung der Anzahl der Läsionen auszuwerten, um Verzerrungen durch Patienten mit vielen oder wenigen Läsionen bei Studienbeginn zu vermeiden.

In den Studien PERFECT-1 und -2 wurden separat entzündliche und nicht entzündliche Läsionen im Gesicht und am Rumpf gezählt (Details zur Definition der einzelnen Läsionen siehe Abschnitt 4.3.1.3.3, Tabelle 4-22). Die Auswertung berücksichtigt jeweils die absolute und die prozentuale Änderung der Anzahl entzündlicher und nicht entzündlicher Läsionen im Gesicht und am Rumpf.

Zusammenfassend ist die Änderung der Anzahl der Läsionen als Maß für eine Verringerung der Krankheitssymptome als patientenrelevant anzusehen.

Selbsteinschätzung der Verbesserung der Akne

In einer Indikation wie der Akne, die mit einer starken Beeinträchtigung des psychischen Wohlbefindens des Patienten aufgrund der sichtbaren und spürbaren Ausprägungen der Erkrankung einhergeht, ist die Patientensicht zur Erfassung der Krankheitsschwere bzw. des Therapieerfolgs besonders relevant. Die selbst empfundene Verbesserung der Erkrankung und damit die Selbsteinschätzung der Verbesserung der Akne ist somit grundsätzlich als patientenrelevant anzusehen.

Die Selbsteinschätzung der Verbesserung der Akne wurde von den Patienten in den Studien PERFECT-1 und -2 über eine sechsstufige Skala von 0 (vollständige Verbesserung) bis 5 (Verschlechterung) erfasst. Für die Auswertung des Endpunktes im vorliegenden Dossier wurden Patienten mit einer vollständigen oder deutlichen Verbesserung der Akne, entsprechend den Stufen 0 und 1 der Skala, als Patienten mit einer relevanten Verbesserung betrachtet.

Die globale Selbsteinschätzung der Veränderung durch die Patienten wurde bereits in Verfahren der frühen Nutzenbewertung als patientenrelevant anerkannt [18-20]. In diesen Verfahren wurde die Veränderung jeweils mittels einer *Global-Impression-of-Change*-Skala erhoben, welche sich von der in den Studien PERFECT-1 und -2 verwendeten Skala lediglich darin unterscheidet, dass sie sieben anstelle von sechs Antwortmöglichkeiten umfasst.

Zusammenfassend ist die Selbsteinschätzung der Patienten bezüglich der Verbesserung ihrer Akne als patientenrelevant anzusehen. Die Erhebung über eine sechsstufige Skala und

Auswertung der Patienten mit einer relevanten Verbesserung stellt eine etablierte Operationalisierung dieses Endpunktes dar [21].

Gesundheitsbezogene Lebensqualität

DLQI / CDLQI

Die gesundheitsbezogene Lebensqualität der Patienten ist ein grundsätzlich patientenrelevanter Endpunkt, der die Auswirkungen einer Intervention direkt im aktuellen Befinden der Patienten widerspiegelt.

Der DLQI bzw. der CDLQI sind Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität bei erwachsenen bzw. pädiatrischen Patienten mit unterschiedlichen Hauterkrankungen. Der CDLQI wurde speziell für Kinder und Jugendliche bis einschließlich 16 Jahre entwickelt und validiert, während bei der Entwicklung des DLQI im Wesentlichen Erwachsene einbezogen wurden [22, 23]. Die Fragebögen umfassen zehn Fragen, die den Einfluss der Hauterkrankung auf die Lebensqualität abbilden sollen. Die zehn Fragen beziehen sich auf folgende sechs Domänen: Symptome und Befinden, tägliche Aktivitäten, Freizeit, Arbeit/Schule, persönliche Beziehungen und Auswirkungen der Behandlung auf den Alltag. Die Fragebögen unterscheiden sich im Wesentlichen darin, dass die Frage zur Sexualität (DLQI) durch eine Frage zum Schlaf (CDLQI) ersetzt wurde.

Der DLQI bzw. CDLQI kann Werte zwischen 0 und 30 annehmen, wobei niedrige Werte eine bessere Lebensqualität bedeuten [22, 23].

Um die DLQI-Werte besser interpretieren zu können, wurden mittels eines ankerbasierten Verfahrens Spannen des DLQI definiert und bestimmten Schweregraden der Beeinträchtigung der Lebensqualität zugeordnet [24]. Ein DLQI von 0 oder 1 gilt dabei als nicht beeinträchtigte Lebensqualität, während ein DLQI von 2 bis einschließlich 5 eine geringe Einschränkung und Werte von 6 bis einschließlich 10 eine moderate Einschränkung der Lebensqualität anzeigen. DLQI-Werte über 10 bedeuten eine große Beeinträchtigung der Lebensqualität. Für den CDLQI wurde folgende Kategorisierung der Schweregrade vorgeschlagen: Werte von 0 bis 1 für eine nicht beeinträchtigte Lebensqualität, 2 bis 6 für eine geringe Einschränkung, 7 bis 12 für eine moderate Einschränkung und Werte ab 13 für eine große Beeinträchtigung der Lebensqualität [25, 26].

Im vorliegenden Dossier wurde im Rahmen von Responderanalysen der Anteil der Patienten, die unter der Therapie einen DLQI von 0 oder 1 bzw. CDLQI von 0 oder 1 erreicht haben, ermittelt.

Zusammenfassend stellt der DLQI bzw. CDLQI ein valides und etabliertes Instrument zur Erhebung der gesundheitsbezogenen Lebensqualität bei Patienten mit Hauterkrankungen wie der Akne dar. Der DLQI wurde bereits in einer Reihe von Nutzenbewertungsverfahren als valide erachtet und für die Nutzenbewertung herangezogen (z. B. [27]). Auch die Auswertung des Anteils der Patienten mit einem DLQI 0/1 wurde in diesen Nutzenbewertungsverfahren als adäquat anerkannt.

Sicherheit/Verträglichkeit

Sicherheit

Die Erhebung von unerwünschten Ereignissen, die im zeitlichen Zusammenhang mit der Therapie auftreten, ist ein wesentlicher Bestandteil der Bewertung der Sicherheit eines Arzneimittels. Gemäß der AM-NutzenV ist der Nutzen eines Arzneimittels der patientenrelevante therapeutische Effekt unter anderem hinsichtlich der Verringerung von Nebenwirkungen. Die Verringerung von therapieassoziierten Nebenwirkungen ist somit als patientenrelevanter Aspekt eines zu bewertenden Arzneimittels anzusehen. Dennoch sind nicht alle untersuchten unerwünschten Ereignisse direkt patientenrelevant, da beispielsweise Änderungen von Laborparametern nicht per se für den Patienten spürbar und nicht zwingend symptomatisch sind. Als grundsätzlich patientenrelevant sind in der Regel SUE und schwere UE anzusehen, sowie Therapieabbrüche aufgrund von UE.

In den Studien PERFECT-1 und -2 waren zudem die folgenden UE als UE von besonderem Interesse präspezifiziert: klinisch relevante Laborwertänderungen, die als mit der Studienmedikation in Zusammenhang stehend erachtet wurden, kutane UE, die zum Therapieabbruch führten, und allergische Reaktionen, die als potentiell mit der Studienmedikation in Zusammenhang stehend erachtet wurden.

Lokale Verträglichkeit

Für die Bewertung eines topisch anzuwendenden Arzneimittels ist die lokale Verträglichkeit von besonderer Bedeutung. Das Auftreten von lokalen Unverträglichkeiten aufgrund der Anwendung eines Arzneimittels ist somit patientenrelevant, insbesondere, wenn es sich um schwere und langanhaltende Hautreaktionen handelt.

Hautirritationen gehören zu den häufigen Nebenwirkungen bei der topischen Anwendung von Retinoiden. Aus diesem Grund wurde die lokale Verträglichkeit in den Studien PERFECT-1 und -2 als separater Aspekt der UE anhand folgender vier Parameter erfasst: Rötung, Schuppung, Trockenheit und Brennen der Haut. Die Dokumentation dieser einzelnen Symptome erfolgte jeweils mittels einer vierstufigen Skala von 0 (kein Symptom) bis 3 (stärkstes Symptom) separat für das Gesicht und den Rumpf (Details siehe Abschnitt 4.3.1.3.8, Tabelle 4-49). Die Symptome Rötung, Schuppung und Trockenheit wurden dabei durch den Prüfarzt bewertet, das Symptom Brennen wurde vom Prüfarzt nach Angaben des Patienten eingeschätzt.

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 nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung[5] erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen 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[6] 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 geeigneter statistische Maße auf Vorliegen von Heterogenität[7, 5] erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0{,}05$), 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.

Die dargestellten Studien PERFECT-1 und -2 untersuchen die gleiche Fragestellung und sind in ihrem Design ist im Hinblick auf die Ein- und Ausschlusskriterien, die Behandlung der

[5] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. Cochrane Database Syst Rev 2015: 25-27.
[6] 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.
[7] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.

Patienten sowie die Erhebung der Endpunkte identisch. Die zu Studienbeginn erhobenen demografischen und klinischen Charakteristika der eingeschlossenen Patienten sind vergleichbar. Die Studien wurden im gleichen Zeitraum durchgeführt. Die Orte der Studiendurchführung unterscheiden sich zwischen den Studien, so wurde in PERFECT-1 die Mehrheit der Patienten in Nordamerika eingeschlossen, während in PERFECT-2 überwiegend europäische Patienten rekrutiert wurden. Aus Sicht der Zulassungsbehörden ist jedoch nicht zu erwarten, dass die Region einen relevanten Einfluss auf die Studienergebnisse hat [17]. Zusammenfassend sind die Ergebnisse der beiden Studien aus medizinischer und methodischer Sicht grundsätzlich geeignet für eine meta-analytische Zusammenfassung.

Da nur zwei Studien vorlagen, wurde für die Meta-Analyse ein Modell mit festem Effekt gewählt. Die Berechnung der gemeinsamen Effektschätzer erfolgte unter Gewichtung mittels der Inverse-Varianz-Methode.

Zur Beurteilung der Heterogenität wurde der p-Wert des statistischen Heterogenitätstests betrachtet. Bei einem p-Wert von unter 0,05 ist von einem Vorliegen einer bedeutsamen Heterogenität auszugehen. Lagen in den Studien gleichgerichtete und statistisch signifikante Effekte vor, wurde auch im Falle eines p-Wertes unter 0,05 eine meta-analytische Zusammenfassung vorgenommen. Waren heterogene Effekte nicht deutlich gleichgerichtet, wurden die Effekte der Einzelstudien betrachtet.

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.

In den Studien PERFECT-1 und -2 waren verschiedene Sensitivitätsanalysen präspezifiziert, um die Robustheit der Ergebnisse für die primären und sekundären Endpunkte zu prüfen. Zum einen wurden die Hauptanalysen auf Basis der *Per-Protocol*-Population anstelle der ITT-Population wiederholt, zum anderen wurden die Auswertungen auf Basis der ITT-Population mit unterschiedlichen alternativen Ersetzungsstrategien für fehlende Werte durchgeführt (Details siehe Angaben zu Sensitivitätsanalysen in Anhang 4-E).

Die in den Studienberichten dargestellten Ergebnisse der Sensitivitätsanalysen bestätigen die Robustheit der Effekte für die primären und sekundären Endpunkte. Da die

Sensitivitätsanalysen keine zusätzliche Information zur Bewertung des medizinischen Nutzens oder Zusatznutzens von Trifaroten liefern, wird auf eine Darstellung der Methodik und Ergebnisse im vorliegenden Modul verzichtet.

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 (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.9).

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 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. 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 Darstellung der Studien PERFECT-1 und -2 im vorliegenden Dossier dient nicht der Identifikation von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen, da der

Komparator in beiden Studien nicht der zweckmäßigen Vergleichstherapie entspricht. Die Studien werden zur Darstellung des medizinischen Nutzens herangezogen, so dass post-hoc-Subgruppenanalysen zusätzlich zu den a priori geplanten Auswertungen keinen weiteren Erkenntnisgewinn für die Nutzenbewertung bieten.

In den Studien PERFECT-1 und -2 waren für die drei ko-primären Endpunkte zur Wirksamkeit im Gesicht folgende Subgruppenanalysen a priori geplant:

- Alter (< 18 / ≥ 18 Jahre)
- Geschlecht (Männer / Frauen)
- Abstammung (weiß / nicht weiß)
- Land (PERFECT-1: USA / Kanada; PERFECT-2: USA / EU / Russland)
- Hauttyp (I – III / IV – VI)

Zusätzlich zu den a priori im statistischen Analyseplan (SAP) vorgesehenen Endpunkten wurden die Subgruppenanalysen auch für die drei sekundären Endpunkte zur Wirksamkeit am Rumpf durchgeführt. Abweichend von der ursprünglichen Planung im SAP wurden die Subgruppen nach Region bzw. Land nicht ausgewertet, da im Laufe der Rekrutierungsphase beider Studien die Anzahl der Länder mit teilnehmenden Studienzentren erhöht wurde, so dass die im SAP prädefinierten Subgruppenanalysen nach Land nicht wie geplant durchgeführt werden konnten.

Für die Darstellung des medizinischen Nutzens im vorliegenden Dossier werden lediglich die Subgruppenanalysen für die Faktoren Alter und Geschlecht als relevant erachtet. Die Faktoren Abstammung und Hauttyp sind für die Übertragbarkeit der Ergebnisse auf die Patienten in Deutschland nicht relevant. Die Studienpopulationen der beiden Studien bestehen zu mindestens 80 % jeweils aus Patienten mit weißer Abstammung und dem Hauttyp I bis III. Somit ist aufgrund der demographischen Merkmale der deutschen Bevölkerung aus den Subgruppenanalysen für die Faktoren Abstammung und Hauttyp kein relevanter Erkenntnisgewinn für Deutschland zu erwarten.

Eine Subgruppenanalyse nach Krankheitsschwere ist im vorliegenden Anwendungsgebiet nicht sinnvoll, da dieses ausschließlich Patienten mit mittelschwerer Akne umfasst.

Im vorliegenden Dossier werden somit die Ergebnisse für die Subgruppenanalysen nach Alter und Geschlecht für die Endpunkte Therapieerfolgsrate im Gesicht (IGA 0/1), Therapieerfolgsrate am Rumpf (PGA 0/1) sowie die Änderung der Anzahl der entzündlichen und nicht entzündlichen Läsionen im Gesicht bzw. am Rumpf dargestellt.

Für die Bewertung möglicher Effektmodifikatoren wird gemäß IQWiG-Methodik ein Interaktions-p-Wert < 0,05 als Beleg für eine Effektmodifikation durch ein Merkmal angenommen [28]. Werden nach dieser Methodik zwischen den Subgruppen signifikant

unterschiedliche Effekte ($p < 0,05$) ermittelt, werden die Ergebnisse auf Subgruppenebene aufgeführt und bewertet.

Die Interpretation von Subgruppenanalysen unterliegt allgemein anerkannten Einschränkungen. Dies gilt insbesondere für post-hoc-Analysen. Im Allgemeinen werden Subgruppenergebnisse als hypothesengenerierend betrachtet und nicht als konfirmatorische Überprüfung einer Hypothese. Daher ist es problematisch, aus Subgruppenergebnissen Aussagen über den Zusatznutzen nach der Methodik der frühen Nutzenbewertung abzuleiten. Ein weiteres Problem ist, dass bei Subgruppenanalysen prinzipiell nicht für multiples Testen korrigiert wird. Ein Signifikanzniveau von 0,05 ist mit einer Irrtumswahrscheinlichkeit von 5 % behaftet. Dies bedeutet, dass bei 5 % der pro Studie durchgeführten Interaktionstests falsch positive Ergebnisse zu erwarten sind, d. h. die Nullhypothese, dass keine Interaktion besteht, wird fälschlicherweise abgelehnt. Somit sind mit steigender Anzahl von Analysen zunehmend positive Interaktionstests zu erwarten, die vermutlich als Zufallsergebnisse zu interpretieren sind.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethode darstellen[8]. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche[9] oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)[10] und Rücker (2012)[11] vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten

[8] Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.
[9] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50(6): 683-691.
[10] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.
[11] Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods 2012; 3(4): 312-324.

Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist[12].

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades).

Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen[13, 14, 15]

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

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten 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).*

[12] Schöttker B, Lühmann D, Boulkhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.
[13] 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.
[14] Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study BMJ 2011; 343 :d4909
[15] Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Res Synth Methods 2013; 4(4): 291-323.

- *Art und Umfang von Sensitivitätsanalysen.*

Bei der Suche nach Studien zum Nachweis des Zusatznutzens von Trifaroten konnten keine geeigneten Studien identifiziert werden, die Trifaroten mit der zweckmäßigen Vergleichstherapie vergleichen. Bei den dargestellten Studien PERFECT-1 und -2 handelt es sich um die zulassungsbegründenden Phase-III-Studien, in welchen die Wirksamkeit und Sicherheit einer Behandlung mit Trifaroten im Vergleich mit Vehikel untersucht wurde. Grundsätzlich wäre somit die Durchführung eines adjustierten indirekten Vergleichs mit Vehikel als Brückenkomparator möglich.

Im Rahmen der frühen Nutzenbewertung bestehen hohe methodische Anforderungen an indirekte Vergleiche, welche in vielen Fällen nicht erfüllt werden konnten. Ein häufiger Grund für die Ablehnung von indirekten Vergleichen ist die nicht hinreichende Ähnlichkeit der Studien im Hinblick auf die Patientenpopulationen oder Aspekte des Studiendesigns.

In der Indikation Akne ist aufgrund folgender Besonderheiten von einer eingeschränkten Vergleichbarkeit der Studien auszugehen:

Ein wesentlicher Aspekt für die Validität eines adjustierten indirekten Vergleichs ist die Übereinstimmung der Studien hinsichtlich des Brückenkomparators. In der vorliegenden Indikation kommt meist das Vehikel als Brückenkomparator in Frage. Bei topischen Arzneimitteln hat jedoch die Trägersubstanz selbst aufgrund ihrer kosmetischen und pharmakologischen Eigenschaften eine mögliche intrinsische Wirkung auf die Akne und somit potentiell einen Einfluss auf den Behandlungserfolg [29, 30]. Die intrinsische Wirkung der verschiedenen Vehikel ist einer der Gründe für die große Variabilität des Placeboeffekts in klinischen Studien zur Behandlung der Akne [31]. Es ist somit unklar, inwieweit Studien mit unterschiedlichen Vehikeln vergleichbar sind bzw. in welchem Ausmaß die Vehikel-Effekte die Ergebnisse eines indirekten Vergleichs verzerren.

Des Weiteren kann die Vergleichbarkeit der Studienpopulationen bei Akne-Patienten dadurch erschwert sein, dass es keine allgemein gültigen Kriterien für die Bestimmung der Krankheitsschwere gibt. In der Literatur sind über 25 verschiedene Skalen bzw. Klassifikationssysteme für Akne beschrieben [32, 33]. Die Variabilität im Hinblick auf die Bestimmung der Krankheitsschwere spiegelt sich auch in der Operationalisierung der Endpunkte wider, da für die Erhebung der Wirksamkeit einer Therapie im Rahmen klinischer Studien in der Regel die gleiche Skala wie für den Einschluss der Patienten genutzt wird. Somit sind die auch Ergebnisse zur Verringerung der Krankheitsschwere unter Umständen nicht vergleichbar.

Um aus einem indirekten Vergleich valide Aussagen für die Nutzenbewertung ableiten zu können, müssen die zentralen Annahmen zur Ähnlichkeit, Homogenität und Konsistenz erfüllt sein. Die beschrieben Besonderheiten der Indikation Akne führen hingegen zu einer erhöhten Unsicherheit im Hinblick auf die Ergebnisse aus indirekten Vergleichen. Aus diesem Grund wurde im vorliegenden Dossier auf die Durchführung eines indirekten Vergleichs verzichtet.

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)

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

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, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-3: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungs-studie (ja/nein)**	**Sponsor (ja/nein)**	**Status (ab-geschlossen / abgebrochen / laufend)**	**Studiendauer ggf. Datenschnitt**	**Therapiearme**
PERFECT-1 (Studie 18251; NCT02566369; EudraCT 2016-002860-15)	ja	ja	abgeschlossen	12 Wochen	Trifaroten 50 µg/g (Creme) Vehikel (Creme)
PERFECT-2 (Studie 18252; NCT02556788; EudraCT 2015-002540-13)	ja	ja	abgeschlossen	12 Wochen	Trifaroten 50 µg/g (Creme) Vehikel (Creme)
Studie 18223 (NCT01616654)	nein(1)	ja	abgeschlossen	12 Wochen	Trifaroten 25 µg/g (Creme) Trifaroten 50 µg/g (Creme) Trifaroten 100 µg/g (Creme) Tazaroten 0,1 % Vehikel (Creme)
Studie 40126 (EudraCT 2010-022211-19)	nein(1)	ja	abgeschlossen	26 Tage	Trifaroten 100 µg/g (Gel)(2) Tazaroten 0,1 % Vehikel (Gel)
Studie 40129 (EudraCT 2010-024214-54)	nein(1)	ja	abgeschlossen	28 Tage	Trifaroten 25 µg/g (Creme)(2) Trifaroten 50 µg/g (Creme)(2) Trifaroten 50 µg/g (Creme) Trifaroten 50 µg/g (Gel)(2) Tazaroten 0,1 %
Studie 40076E (EudraCT 2008-007981-44)	nein(1)	ja	abgeschlossen	27 Tage	Trifaroten 100 µg/g (Gel)(2) Trifaroten 50 µg/g (Gel)(2) Adapalen/Benzoylperoxid 0,1 %/2,5 % Vehikel (Gel)
Studie 18213	nein(1)	ja	abgeschlossen	28 Tage	Trifaroten 25 µg/g (Creme)(2) Trifaroten 50 µg/g (Creme)(2) Trifaroten 50 µg/g (Gel)(2) Vehikel (Gel)
Studie 18214	nein(1)	ja	abgeschlossen	28 Tage	Trifaroten 25 µg/g (Creme) Trifaroten 50 µg/g (Creme) Trifaroten 50 µg/g (Gel)(2) Vehikel (Gel)

(1) Die Studien wurden als supportive Evidenz mit dem CTD an die Zulassungsbehörde übermittelt.
(2) Die Formulierung des Vehikels entspricht nicht der Formulierung des Vehikels der zugelassenen Trifaroten-Creme.

Geben Sie an, welchen Stand die Information in Tabelle 4-3 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.

Die Informationen in Tabelle 4-3 geben die Informationen mit Stand vom 16.06.2020 wieder.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-3 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

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

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
PERFECT-1 (Studie 18251)	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
PERFECT-2 (Studie 18252)	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
Studie 18223	A1: Weniger als 80 % der Studienteilnehmer entsprechen der Patientenpopulation gemäß zugelassenem Anwendungsgebiet. A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
Studie 40126	A2: Die Intervention entspricht nicht Trifaroten in der zugelassenen Formulierung und Dosierung. A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie. A6: Studie ist kürzer als 12 Wochen.
Studie 40129	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie. A6: Studie ist kürzer als 12 Wochen.
Studie 40076E	A2: Die Intervention entspricht nicht Trifaroten in der zugelassenen Formulierung. A6: Studie ist kürzer als 12 Wochen.
Studie 18213	A2: Die Intervention entspricht nicht Trifaroten in der zugelassenen Formulierung. A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie. A6: Studie ist kürzer als 12 Wochen.
Studie 18214	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie. A6: Studie ist kürzer als 12 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.

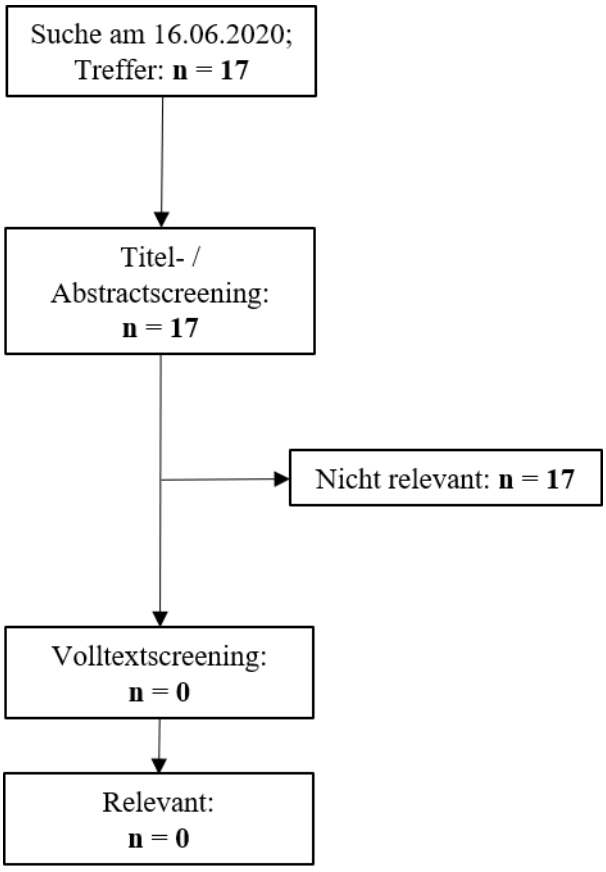


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliographische Literaturrecherche erfolgte am 16.06.2020. Die Suche in MEDLINE und EMBASE über Ovid sowie in der Cochrane Library ergab insgesamt 17 Treffer. Alle Treffer wurden ohne Ausschluss von Dubletten anhand von Titel und Abstract gescreent. Es wurde kein relevanter Treffer identifiziert.

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank 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 4-3) 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 4-5: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	**Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank und Angabe der Zitate[a])**	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Status (abgeschlossen/ abgebrochen/ laufend)**
-	-	-	-	-
a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

Geben Sie an, welchen Stand die Information in Tabelle 4-5 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.

Die Informationen in Tabelle 4-5 geben die Informationen mit Stand vom 17.06.2020 wieder. Die Suche in Studienregistern ergab insgesamt 33 Treffer. Es wurde keine relevante Studie identifiziert (Ausschlussgründe siehe Anhang 4-D).

4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-3) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

[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 4-6: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	**Relevante Quellen**[a]	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Studie durch Suche in Studienregistern / Studienergebnis datenbanken identifiziert (ja/nein)**
-	-	-	-	-
a: Quellen aus der Suche auf der Internetseite des G-BA				

Geben Sie an, welchen Stand die Information in Tabelle 4-6 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.

Die Informationen in Tabelle 4-6 geben die Informationen mit Stand vom 17.06.2020 wieder. Bei der Suche auf der Internetseite des G-BA wurde keine relevante Studie identifiziert.

4.3.1.1.5 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, 4.3.1.1.3 und 4.3.1.1.4) 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 aller verfügbaren Quellen in Abschnitt 4.3.1.4 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 Zwischenzeilenüberschriften 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.

Im Rahmen der Suche nach Studien zum Nachweis des Zusatznutzens von Trifaroten wurde keine geeignete Studie identifiziert, die Trifaroten mit der zweckmäßigen Vergleichstherapie vergleicht.

Im Folgenden werden somit die beiden zulassungsbegründenden Phase-III-Studien PERFECT-1 und -2 dargestellt, welche aufgrund des Komparators Vehikel formal nicht geeignet sind, einen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie gemäß § 35a SGB V zu belegen. Die Studien werden dennoch in den folgenden Abschnitten herangezogen, um den medizinischen Nutzen von Trifaroten nachzuweisen. Ergänzend werden zudem Ergebnisse aus der einarmigen Phase-III-Studie SATISFY dargestellt, um die Langzeitwirksamkeit und -sicherheit von Trifaroten zu zeigen.

Tabelle 4-7: 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-berichte (ja/nein [Zitat])	Register-einträge[c] (ja/nein [Zitat])	Publikation und sonstige Quellen[d] (ja/nein [Zitat])
Studien zum Nachweis des Zusatznutzens						
-						
Studien zum Nachweis des medizinischen Nutzens - placebokontrolliert						
PERFECT-1	ja	ja	nein	ja [34]	ja [35-37]	ja [38]
PERFECT-2	ja	ja	nein	ja [39]	ja [40-42]	ja [38]
Ergänzende Evidenz zur Langzeitwirksamkeit und -sicherheit						
SATISFY	ja	ja	nein	ja [43]	ja [44-46]	ja [47]

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 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (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.

d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.

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. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	**Studiendesign** <RCT, doppelblind/ einfach verblindet/ offen, parallel/ cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
PERFECT-1 (Studie 18251)	Phase-III-RCT, doppelblind, parallel, multizentrisch	Patienten ab 9 Jahren mit mittelschwerer *Acne vulgaris* im Gesicht und am Rumpf	Trifaroten: n = 612 Vehikel: n = 596	Behandlung: 12 Wochen	Deutschland, Kanada, Puerto Rico, Ungarn, USA 11/2015 – 11/2017	<u>Primäre Endpunkte</u>: Therapieerfolgsrate (IGA 0/1) Absolute Änderung der entzündlichen und nicht entzündlichen Läsionen im Gesicht <u>Sekundäre und weitere Endpunkte</u>: Therapieerfolgsrate (PGA 0/1) Absolute Änderung der entzündlichen und nicht entzündlichen Läsionen am Rumpf Prozentuale Änderung der entzündlichen und nicht entzündlichen Läsionen im Gesicht und am Rumpf Selbsteinschätzung der Patienten Lebensqualität (DLQI / CDLQI) Unerwünschte Ereignisse Lokale Verträglichkeit im Gesicht und am Rumpf
PERFECT-2 (Studie 18252)	Phase-III-RCT, doppelblind, parallel, multizentrisch	Patienten ab 9 Jahren mit mittelschwerer *Acne vulgaris* im Gesicht und am Rumpf	Trifaroten: n = 602 Vehikel: n = 610	Behandlung: 12 Wochen	Polen, Rumänien, Russland, Spanien, Tschechien, Ukraine, Ungarn, USA	<u>Primäre Endpunkte</u>: Therapieerfolgsrate (IGA 0/1) Absolute Änderung der entzündlichen und nicht entzündlichen Läsionen im

Studie	**Studiendesign** <RCT, doppelblind/ einfach verblindet/ offen, parallel/ cross-over etc.>	**Population** <relevante Charakteristika, z. B. Schweregrad>	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte** <ggf. Run-in, Behandlung, Nachbeobachtung>	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
					11/2015 – 11/2017	Gesicht <u>Sekundäre und weitere Endpunkte</u>: Therapieerfolgsrate (PGA 0/1) Absolute Änderung der entzündlichen und nicht entzündlichen Läsionen am Rumpf Prozentuale Änderung der entzündlichen und nicht entzündlichen Läsionen im Gesicht und am Rumpf Selbsteinschätzung der Patienten Lebensqualität (DLQI / CDLQI) Unerwünschte Ereignisse Lokale Verträglichkeit im Gesicht und am Rumpf
SATISFY (Studie 18250)	Phase-III-Studie, unkontrolliert, offen, multizentrisch	Patienten ab 9 Jahren mit mittelschwerer *Acne vulgaris* im Gesicht und am Rumpf	Trifaroten: n = 455	Behandlung: 52 Wochen	Deutschland, Tschechien, Ungarn, USA 02/2015 – 02/2017	Therapieerfolgsrate (IGA 0/1) Therapieerfolgsrate (PGA 0/1) Selbsteinschätzung der Patienten Lebensqualität (DLQI / CDLQI) Unerwünschte Ereignisse Lokale Verträglichkeit im Gesicht und am Rumpf

CDLQI: *Children´s Dermatology Life Quality Index*; DLQI: *Dermatology Life Quality Index*; IGA: *Investigator´s Global Assessment*; PGA: *Physician´s Global Assessment*

Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Intervention	Komparator
PERFECT-1	Creme mit Trifaroten 50 µg/g einmal täglich im Gesicht (ein Pumpenhub) und am Rumpf (zwei Pumpenhübe)	Vehikel-Creme einmal täglich im Gesicht (ein Pumpenhub) und am Rumpf (zwei Pumpenhübe)
PERFECT-2	Creme mit Trifaroten 50 µg/g einmal täglich im Gesicht (ein Pumpenhub) und am Rumpf (zwei Pumpenhübe)	Vehikel-Creme einmal täglich im Gesicht (ein Pumpenhub) und am Rumpf (zwei Pumpenhübe)
SATISFY	Creme mit Trifaroten 50 µg/g einmal täglich im Gesicht (ein Pumpenhub) und am Rumpf (zwei Pumpenhübe)	-

Tabelle 4-10: Charakterisierung der Studienpopulationen (demographische Charakteristika) – RCT mit dem zu bewertenden Arzneimittel

	PERFECT-1		**PERFECT-2**		**SATISFY**
	Trifaroten N = 612	**Vehikel N = 596**	**Trifaroten N = 602**	**Vehikel N = 610**	**Trifaroten N_{SAF} = 453**
Alter [Jahre]					
Mittelwert (SD)	19,6 (6,88)	19,3 (5,89)	19,6 (6,20)	19,9 (6,38)	18,3 (6,6)
Median	17,0	18,0	18,0	18,0	16,0
Min; Max	9; 58	10; 50	11; 49	11; 46	9,0; 54,0
Altersgruppe [n (%)]					
9 bis ≤ 11 Jahre	10 (1,6)	9 (1,5)	9 (1,5)	6 (1,0)	18 (4,0)
12 bis ≤ 17 Jahre	304 (49,7)	269 (45,1)	267 (44,4)	288 (47,2)	268 (59,2)
≥ 18 Jahre	298 (48,7)	318 (53,4)	326 (54,2)	316 (51,8)	167 (36,9)
Geschlecht [n (%)]					
Männlich	307 (50,2)	272 (45,6)	245 (40,7)	272 (44,6)	227 (50,1)
Weiblich	305 (49,8)	324 (54,4)	357 (59,3)	338 (55,4)	226 (49,9)
Abstammung [n (%)]					
Weiß	508 (83,0)	484 (81,2)	565 (93,9)	554 (90,8)	432 (95,4)
Schwarz/afroamerik.	47 (7,7)	49 (8,2)	27 (4,5)	42 (6,9)	12 (2,6)
Asiatisch	23 (3,8)	32 (5,4)	2 (0,3)	6 (1,0)	3 (0,7)
Andere	34 (5,6)	31 (5,2)	8 (1,3)	8 (1,3)	6 (1,3)
Hauttyp [n (%)]					
I	31 (5,1)	34 (5,7)	36 (6,0)	37 (6,1)	13 (2,9)
II	197 (32,2)	182 (30,5)	274 (45,5)	249 (40,8)	188 (41,5)
III	233 (38,1)	227 (38,1)	233 (38,7)	248 (40,7)	184 (40,6)

	PERFECT-1		PERFECT-2		SATISFY
	Trifaroten N = 612	**Vehikel N = 596**	**Trifaroten N = 602**	**Vehikel N = 610**	**Trifaroten N_{SAF} = 453**
IV	97 (15,8)	91 (15,3)	33 (5,5)	38 (6,2)	53 (11,7)
V	43 (7,0)	48 (8,1)	14 (2,3)	19 (3,1)	7 (1,5)
VI	11 (1,8)	14 (2,3)	12 (2,0)	19 (3,1)	2 (0,4)

N: Anzahl der randomisierten Patienten (ITT-Population); N_{SAF}: Anzahl der Patienten mit mindestens einer Anwendung der Studienmedikation (*Safety*-Population); SD: Standardabweichung (*standard deviation*)

Tabelle 4-11: Charakterisierung der Studienpopulationen (Basischarakteristika der Erkrankung) – RCT mit dem zu bewertenden Arzneimittel

	PERFECT-1		PERFECT-2		SATISFY
	Trifaroten N = 612	**Vehikel N = 596**	**Trifaroten N = 602**	**Vehikel N = 610**	**Trifaroten N_{SAF} = 453**
Krankheitsschwere Gesicht – IGA [n (%)]					
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	612 (100)	596 (100)	602 (100)	610 (100)	453 (100)
4	0	0	0	0	0
Krankheitsschwere Rumpf – PGA [n (%)]					
0	9 (1,5)	8 (1,3)	3 (0,5)	1 (0,2)	4 (0,9)
1	1 (0,2)	2 (0,3)	1 (0,2)	0	4 (0,9)
2	2 (0,3)	1 (0,2)	0	0	1 (0,2)
3	600 (98,0)	585 (98,2)	598 (99,3)	609 (99,8)	444 (98,0)
4	0	0	0	0	0
Entzündliche Läsionen im Gesicht					
Mittelwert (SD)	34,7 (13,02)	34,8 (13,61)	36,1 (12,47)	37,1 (15,06)	36,9 (15,0)
Median	31,0	31,0	33,0	34,0	32,0
Min; Max	20; 131	20; 113	10; 110	7; 200	20; 123
Nicht entzündliche Läsionen im Gesicht					
Mittelwert (SD)	54,0 (28,55)	52,8 (26,08)	50,6 (25,93)	51,2 (25,75)	58,2 (36,7)
Median	46,0	45,0	43,0	44,0	48,0
Min; Max	22; 225	21; 191	25; 232	25; 305	22; 363

	PERFECT-1		PERFECT-2		SATISFY
	Trifaroten N = 612	**Vehikel N = 596**	**Trifaroten N = 602**	**Vehikel N = 610**	**Trifaroten N_{SAF} = 453**
Entzündliche Läsionen am Rumpf					
Mittelwert (SD)	36,9 (17,89)	35,6 (16,70)	39,0 (16,16)	39,1 (17,41)	43,4 (28,6)
Median	32,0	31,0	35,0	34,0	34,0
Min; Max	0, 140	0, 115	0, 100	0, 220	0, 202
Nicht entzündliche Läsionen am Rumpf					
Mittelwert (SD)	46,4 (21,57)	47,5 (21,94)	46,1 (20,17)	45,7 (19,58)	56,1 (39,5)
Median	42,0	43,0	42,0	42,5	45,0
Min; Max	0; 125	0; 107	0; 180	0; 260	0; 350

IGA: *Investigator´s Global Assessment*; N: Anzahl der randomisierten Patienten (ITT-Population); N_{SAF}: Anzahl der Patienten mit mindestens einer Anwendung der Studienmedikation (*Safety*-Population); PGA: *Physician´s Global Assessment*; SD: Standardabweichung (*standard deviation*)

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. 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.

Studien PERFECT-1 und PERFECT-2

Bei den Studien PERFECT-1 (Studie 18251) und PERFECT-2 (Studie 18252) handelt es sich um die zulassungsbegründenden Phase-III-Studien für Trifaroten. Die beiden Studien sind hinsichtlich ihres Studiendesigns identisch und werden im Folgenden gemeinsam beschrieben. Detaillierte Angaben zum Design und der Methodik der Studien finden sich im Anhang in Tabelle 4-76 und Tabelle 4-77.

Studiendesign

Bei den Studien PERFECT-1 und -2 handelt es sich um multizentrische, randomisierte, kontrollierte, doppelblinde Studien im Parallelgruppendesign.

Beide Studien untersuchten die Wirksamkeit und Sicherheit der topischen Behandlung mit 50 µg/g Trifaroten im Vergleich zur Vehikel-Creme bei Patienten mit *Acne vulgaris*. Die Creme

wurde jeweils einmal täglich im Gesicht und am Rumpf angewendet. Die Patienten wurden im Verhältnis 1:1 auf die beiden Behandlungsarme randomisiert. Die Randomisierung erfolgte stratifiziert nach Studienzentrum. Die doppelblinde Behandlungsphase dauerte zwölf Wochen. Nach Beendigung der Behandlung endete die Studienteilnahme ohne Follow-up (siehe Abbildung 4-2).

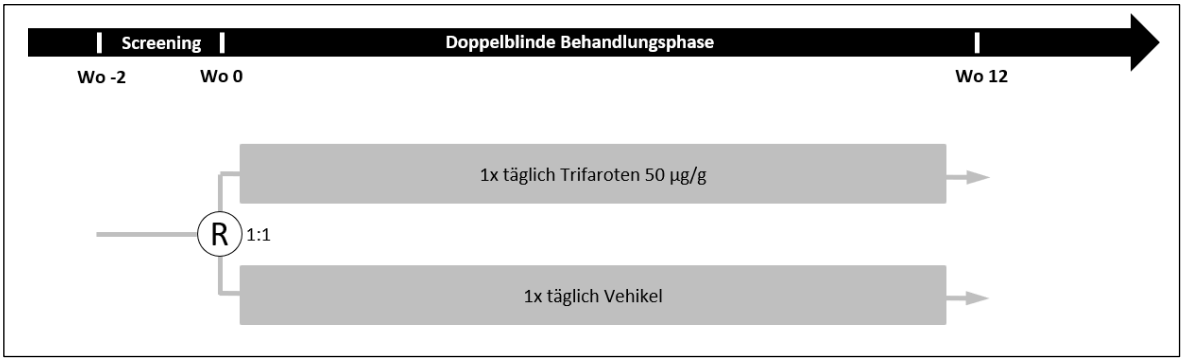


Abbildung 4-2: Design der Studien PERFECT-1 und PERFECT-2

Als primäre Endpunkte wurden die Therapieerfolgsrate im Gesicht gemäß IGA 0/1 sowie die absolute Änderung der Anzahl der entzündlichen und der nicht entzündlichen Läsionen im Gesicht nach zwölf Wochen Behandlung bestimmt. Als sekundäre Endpunkte wurden die Therapieerfolgsrate am Rumpf gemäß PGA 0/1 sowie die absolute Änderung der Anzahl der entzündlichen und der nicht entzündlichen Läsionen am Rumpf nach zwölf Wochen Behandlung bestimmt. Weitere Endpunkte waren die Selbsteinschätzung der Patienten bezüglich der Verbesserung der Akne, die Lebensqualität, erhoben mittels DLQI bzw. CDLQI, sowie die Sicherheit und lokale Verträglichkeit.

Patientenpopulation

In beide Studien wurden Patienten ab neun Jahren mit einer mittelschweren Akne im Gesicht, entsprechend einem Wert auf der IGA-Skala von 3, sowie einer mittelschweren Akne am Rumpf, entsprechend einem Wert auf der PGA-Skala von 3, eingeschlossen. Patienten im Alter von neun bis elf Jahren wurden unabhängig vom Vorliegen einer Akne am Rumpf eingeschlossen.

Die Patienten in den beiden Studien waren hinsichtlich ihrer demographischen Charakteristika sowie der Basischarakteristika der Erkrankung vergleichbar (siehe Tabelle 4-10 und Tabelle 4-11). Auch innerhalb der Studien waren die Patientencharakteristika zwischen den Behandlungsarmen ausgewogen.

Die Patienten waren im Mittel ca. 19 Jahre alt und der Anteil der Jugendlichen und Erwachsenen betrug jeweils etwa 50 %. Der Anteil der 9- bis 11-Jährigen war mit weniger als 2 % sehr gering. Etwa die Hälfte der Studienteilnehmer war männlich. Über 80 % der Teilnehmer war weißer Abstammung. Die Hauttypen II und III waren mit jeweils ca. 30 % bis 45 % die vorherrschenden Hauttypen.

Zu Studienbeginn wiesen alle Patienten eine mittelschwere Akne im Gesicht (IGA 3) und mindestens 98 % der Patienten eine mittelschwere Akne am Rumpf (PGA 3) auf. Die Anzahl der entzündlichen bzw. nicht entzündlichen Läsionen im Gesicht betrug im Mittel 35 bis 37 bzw. 51 bis 54. Am Rumpf wiesen die Patienten im Mittel 36 bis 39 entzündliche bzw. 46 bis 48 nicht entzündliche Läsionen auf.

Studie SATISFY

Die Studie SATISFY (Studie 18250) ist eine Phase-III-Studie zur Untersuchung der Langzeitsicherheit und -wirksamkeit von Trifaroten.

Studiendesign

Bei der Studie SATISFY handelt es sich um eine unkontrollierte, offene, multizentrische Studie zur topischen Langzeitanwendung von 50 µg/g Trifaroten bei Patienten mit *Acne vulgaris*. Die Creme wurde über 52 Wochen jeweils einmal täglich im Gesicht und am Rumpf angewendet.

Patienten, die in Woche 12 bereits eine vollständige Abheilung der Haut im Gesicht und am Rumpf erreicht hatten, sollten die Behandlung aussetzen und diese bei einem erneuten Auftreten der Akne, d. h. bei Feststellung eines Wertes von ≥ 1 auf der IGA- oder PGA-Skala bei einer der nachfolgenden Visiten, wieder aufnehmen.

Primäre Endpunkte der Studie waren die Sicherheit und lokale Verträglichkeit der Behandlung. Als weitere Endpunkte wurden die Krankheitsschwere im Gesicht und am Rumpf mittels IGA- und PGA-Skala, die Selbsteinschätzung der Patienten bezüglich der Verbesserung der Akne, die Lebensqualität mittels DLQI bzw. CDLQI sowie Parameter der Pharmakokinetik erhoben.

Patientenpopulation

In die Studie wurden Patienten ab neun Jahren mit einer mittelschweren Akne im Gesicht, entsprechend einem Wert auf der IGA-Skala von 3, sowie einer mittelschweren Akne am Rumpf, entsprechend einem Wert auf der PGA-Skala von 3, eingeschlossen. Patienten im Alter von neun bis elf Jahren wurden unabhängig vom Vorliegen einer Akne am Rumpf eingeschlossen. Die Ein- und Ausschlusskriterien entsprechen somit denen der Studien PERFECT-1 und -2.

Die Patienten der Studie SATISFY waren hinsichtlich ihrer demographischen Charakteristika sowie der Basischarakteristika der Erkrankung mit den Patienten der Studien PERFECT-1 und -2 vergleichbar (siehe Tabelle 4-10 und Tabelle 4-11).

Das mittlere Alter der Patienten betrug 18 Jahre. Bei etwa 37 % der Patienten handelte es sich um Erwachsene. Der Anteil der unter Zwölfjährigen war mit 4 % gering. Etwa die Hälfte der Studienteilnehmer war männlich. Über 95 % der Teilnehmer waren weißer Abstammung. Die Hauttypen II und III waren mit jeweils ca. 40 % die vorherrschenden Hauttypen.

Alle Patienten hatten zu Studienbeginn eine mittelschwere Akne im Gesicht (IGA 3) und 98 % der Patienten eine mittelschwere Akne am Rumpf (PGA 3). Im Gesicht wiesen die Patienten

im Mittel 37 entzündliche und 58 nicht entzündliche Läsionen auf. Am Rumpf wiesen die Patienten im Mittel 43 entzündliche bzw. 56 nicht entzündliche Läsionen auf.

Übertragbarkeit auf den deutschen Versorgungskontext

Die Studien PERFECT-1 und -2 sowie SATISFY wurden weltweit, unter anderem auch in Deutschland, durchgeführt. Viele Studienzentren rekrutierten Patienten in westlichen Ländern mit einem ähnlichen Standard der Gesundheitsversorgung wie in Deutschland. Über 80 % der Studienteilnehmer waren weißer Abstammung und wiesen einen Hauttyp I bis III auf.

Die in die Studien PERFECT-1 und -2 eingeschlossenen Patienten wurden im Hinblick auf demografische und klinische Faktoren von den am dezentralen Zulassungsverfahren beteiligten europäischen Behörden als repräsentative Population für Patienten mit *Acne vulgaris* in Europa angesehen, da keine relevanten regionalen Unterschiede im Hinblick auf die Akne-Läsionen zu erwarten seien [17].

Es gibt somit keine Hinweise auf eine eingeschränkte Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

4.3.1.2.2 Verzerrungspotenzial 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 4-12: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
PERFECT-1	ja	ja	ja	ja	ja	ja	niedrig
PERFECT-2	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei den Studien PERFECT-1 und -2 handelt es sich um randomisierte, kontrollierte, doppelblinde Phase-III-Studien. Die Randomisierung auf die Behandlungsgruppen wurde

mittels eines *Interactive-Response-Technology*-Systems realisiert. Patienten und behandelnde Personen waren verblindet. Die Erhebung und Auswertung der Endpunkte erfolgte verblindet. Eine ergebnisunabhängige Berichterstattung liegt vor. Abschließend wird das Verzerrungspotential für beide Studien als niedrig eingestuft.

Die ergänzend dargestellte Studie SATISFY ist aufgrund ihres unkontrollierten und offenen Studiendesigns grundsätzlich mit einem hohen Verzerrungspotential behaftet.

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.

Tabelle 4-13: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Morbidität – Therapieerfolgsrate im Gesicht gemäß IGA	Morbidität – Therapieerfolgsrate am Rumpf gemäß PGA	Morbidität – Anzahl der Läsionen im Gesicht	Morbidität – Anzahl der Läsionen am Rumpf	Morbidität – Selbsteinschätzung der Verbesserung der Akne	Gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI	Sicherheit[1]	Lokale Verträglichkeit
PERFECT-1	ja	ja	ja	ja	ja	ja	ja	ja
PERFECT-2	ja	ja	ja	ja	ja	ja	ja	ja
SATISFY	ja	ja	nein	nein	ja	ja	ja	ja

(1) Dargestellt als jegliche unerwünschte Ereignisse (UE), schwerwiegende UE (SUE) und Therapieabbrüche wegen UE
CDLQI: *Children´s Dermatology Life Quality Index*; DLQI: *Dermatology Life Quality Index*; IGA: *Investigator´s Global Assessment*; PGA: *Physician´s Global Assessment*;

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 inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
– 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.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]).

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,

2. Gesamtrate schwerwiegender UE (SUE),

3. Gesamtrate der Abbrüche wegen UE,

4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,

5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind

- Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind

- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

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 4.2.5.3). 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 Therapieerfolgsrate im Gesicht gemäß IGA – RCT

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

Tabelle 4-14: Operationalisierung des Endpunkts Therapieerfolgsrate im Gesicht gemäß IGA

<table>
<tr><th>Studie</th><th>Operationalisierung</th></tr>
<tr><td>PERFECT-1</td><td>Die globale Krankheitsschwere der Akne im Gesicht wurde von geschulten Prüfern über die IGA-Skala (Investigator's Global Assessment) erhoben.
Die IGA-Skala umfasst fünf Stufen, die in der Studie wie folgt definiert waren:
<table>
<tr><td>0</td><td>Abgeheilt</td><td>Abgeheilte Haut ohne entzündliche oder nicht entzündliche Läsionen.</td></tr>
<tr><td>1</td><td>Fast abgeheilt</td><td>Haut mit einigen vereinzelten Komedonen und einigen kleinen Papeln.</td></tr>
<tr><td>2</td><td>Leicht</td><td>Leicht erkennbar; weniger als die Hälfte der Oberfläche ist betroffen. Einige Komedonen sowie einige Papeln und Pusteln.</td></tr>
<tr><td>3</td><td>Mittelschwer</td><td>Mehr als die Hälfte der Oberfläche ist betroffen. Viele Komedonen, Papeln und Pusteln. Ein Knoten kann vorhanden sein.</td></tr>
<tr><td>4</td><td>Schwer</td><td>Die gesamte Oberfläche ist betroffen. Bedeckt mit Komedonen, zahlreichen Papeln und Pusteln. Einige Knoten können vorhanden sein.</td></tr>
</table>
Die Einstufung erfolgte anhand definierter Bereiche von Stirn, Wangen und Kinnbereich, jeweils vor der Bestimmung der genauen Anzahl der Läsionen im Gesicht.
Der IGA-Wert wurde zur Screening- und Baseline-Visite sowie in Woche 1, 2, 4, 8 und 12 erhoben.
Der Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA ist operationalisiert als Anteil der Patienten, die in Woche 12 einen IGA-Wert von 0 oder 1 und eine Verbesserung um mindestens zwei Stufen erreicht haben.</td></tr>
<tr><td>PERFECT-2</td><td>Siehe PERFECT-1</td></tr>
<tr><td>SATISFY</td><td>Siehe PERFECT-1 mit folgenden Abweichungen:
Der IGA-Wert wurde zur Screening- und Baseline-Visite sowie in Woche 12, 20, 26, 38 und 52 bzw. EOT (end of treatment)-Visite erhoben.</td></tr>
<tr><td colspan="2">EOT: End of Treatment; IGA: Investigator's Global Assessment</td></tr>
</table>

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 4-15: Bewertung des Verzerrungspotenzials für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA ist somit als niedrig einzustufen.

Das Verzerrungspotential der ergänzend dargestellten Studie SATISFY ist bereits auf Studienebene als hoch einzustufen, somit erfolgt keine Bewertung des Verzerrungspotentials auf Endpunktebene.

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.

Die Ergebnisse zum Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA aus den Studien PERFECT-1 und PERFECT-2 sind in der nachfolgenden Tabelle 4-16 dargestellt.

Nach zwölfwöchiger Behandlung mit Trifaroten erreichen in den Studien PERFECT-1 und -2 29,4 % bzw. 42,3 % der Patienten das Therapieziel IGA 0/1 und somit eine vollständig abgeheilte oder nahezu abgeheilte Haut im Gesicht. Im Vehikel-Arm trifft dies lediglich auf 19,5 % bzw. 25,7 % der Patienten zu. In beiden Studien ist der Behandlungsunterschied im Vergleich zu Vehikel statistisch signifikant zugunsten von Trifaroten (RR [95 %-KI]: 1,53 [1,24; 1,87], $p < 0,001$ bzw. 1,71 [1,45; 2,02], $p < 0,001$).

Tabelle 4-16: Ergebnisse zum Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA aus RCT (Woche 12, PERFECT-1 und -2)

Studie	**Trifaroten**			**Vehikel**			**Trifaroten vs. Vehikel**		
	N	**N′**	**Patienten mit IGA 0/1 (%)**[1]	**N**	**N′**	**Patienten mit IGA 0/1 (%)**[1]	**OR**[2] **[95 %-KI]**	**RR**[2] **[95 %-KI]**	**ARR**[3] **[95 %-KI]**
PERFECT-1	612	543	29,4	596	539	19,5	1,91 [1,40; 2,60]	1,53 [1,24; 1,87]	9,82 [4,82; 14,81]
							$p < 0,001$[4]		
PERFECT-2	602	563	42,3	610	573	25,7	2,43 [1,85; 3,19]	1,71 [1,45; 2,02]	16,63 [11,29; 21,98]
							$p < 0,001$[4]		

(1) Fehlende Daten wurden mittels multipler Imputation ersetzt.

(2) Odds Ratios, Relative Risiken und die zugehörigen Konfidenzintervalle wurden basierend auf einem CMH-Test stratifiziert nach Zentrum berechnet. Die Effektschätzer wurden für die multipel imputierten Datensätze berechnet und mittels Rubins Methode zusammengefasst.

(3) Risikodifferenzen und die zugehörigen Konfidenzintervalle basieren auf einem Approximationsverfahren für binäre Daten innerhalb einer großen Stichprobe und ohne Anwendung einer Nullzellenkorrektur. Die Effektschätzer wurden für die multipel imputierten Datensätze berechnet und mittels Rubins Methode zusammengefasst.

(4) Die p-Werte basieren auf einem CMH-Test stratifiziert nach Zentrum. Der CMH-Test wurde für jeden multipel imputierten Datensatz berechnet und mittels Rubins Methode unter Anwendung der Wilson-Hilferty-Transformation zusammengefasst.

ARR: Absolute Risikoreduktion in %; CMH: Cochran-Mantel-Haenszel; IGA: *Investigator´s Global Assessment*; N: Patienten in der ITT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; OR: Odds Ratio; RR: Relatives Risiko

In der Langzeitstudie SATISFY zeigt sich zudem, dass der Anteil der Patienten mit IGA 0/1 bei einer Behandlung über Woche 12 hinaus weiter zunimmt (siehe Tabelle 4-17).

Unter Berücksichtigung der Patienten, für die zum Auswertungszeitpunkt Daten vorlagen, erreichen nach zwölf Wochen 26,6 % der Patienten das Therapieziel IGA 0/1. Diese Rate steigt bis Woche 52 auf 65,1 % an. Somit zeigt sich, dass etwa zwei Drittel der Patienten, die eine Behandlung mit Trifaroten beginnen, nach 52-wöchiger Behandlung eine vollständig oder nahezu vollständig abgeheilte Haut im Gesicht aufweisen.

Tabelle 4-17: Ergänzende Ergebnisse zum Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA (Woche 12 und 52, Studie SATISFY)

Studie	**Zeitpunkt**	**N**	**N′**	**Anteil Patienten mit IGA 0/1 n (%)**
SATISFY	Woche 12	453	402	107 (26,6)
	Woche 52	453	350	228 (65,1)

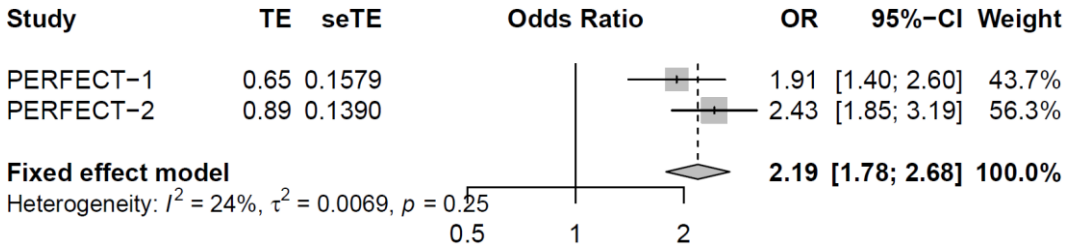
IGA: *Investigator´s Global Assessment*; N: Patienten in der SAF-Population; N′: Patienten mit Wert in Woche 12 bzw. 52

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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-3 sind die Meta-Analysen zum Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel hinsichtlich der Therapieerfolgsrate im Gesicht nach zwölf Wochen Behandlung (RR [95 %-KI]: 1,64 [1,44; 1,86]). Das Ergebnis der Meta-Analyse spiegelt somit die Einzelergebnisse der Studien wider.

a)



b)

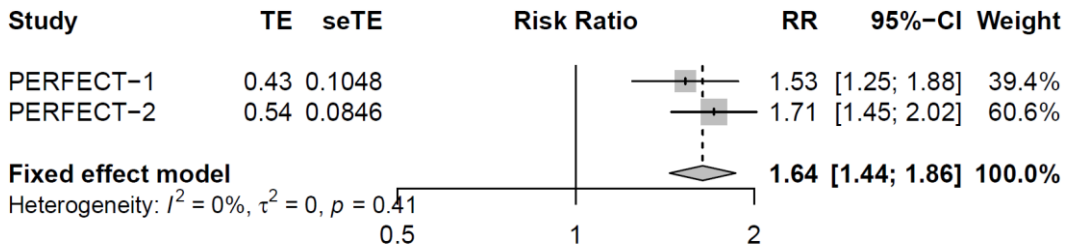


Abbildung 4-3: Meta-Analyse für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

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 Therapieerfolgsrate am Rumpf gemäß PGA – RCT

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

Tabelle 4-18: Operationalisierung des Endpunkts Therapieerfolgs am Rumpf gemäß PGA

<table>
<tr><th>Studie</th><th>Operationalisierung</th></tr>
<tr><td>PERFECT-1</td><td>Die globale Krankheitsschwere der Akne am Rumpf wurde von geschulten Prüfern über die PGA-Skala (Physician’s Global Assessment) erhoben.
Die PGA-Skala umfasst fünf Stufen, die in der Studie wie folgt definiert waren:
<table>
<tr><td>0</td><td>Abgeheilt</td><td>Abgeheilte Haut ohne entzündliche oder nicht entzündliche Läsionen.</td></tr>
<tr><td>1</td><td>Fast abgeheilt</td><td>Haut mit einigen vereinzelten Komedonen und einigen kleinen Papeln.</td></tr>
<tr><td>2</td><td>Leicht</td><td>Leicht erkennbar; weniger als die Hälfte der Oberfläche ist betroffen. Einige Komedonen sowie einige Papeln und Pusteln.</td></tr>
<tr><td>3</td><td>Mittelschwer</td><td>Mehr als die Hälfte der Oberfläche ist betroffen. Viele Komedonen, Papeln und Pusteln. Ein Knoten kann vorhanden sein.</td></tr>
<tr><td>4</td><td>Schwer</td><td>Die gesamte Oberfläche ist betroffen. Bedeckt mit Komedonen, zahlreichen Papeln und Pusteln. Einige Knoten können vorhanden sein.</td></tr>
</table>
Die Einstufung erfolgte anhand definierter Bereiche an den Schultern, dem oberen Rücken und dem oberen Brustkorb (d. h. anhand der Bereiche am oberen Rumpf, die für eine Selbstanwendung der Creme leicht zugänglich sind), jeweils vor der Bestimmung der genauen Anzahl der Läsionen am Rumpf.
Der PGA-Wert wurde zur Screening- und Baseline-Visite sowie in Woche 1, 2, 4, 8 und 12 erhoben.
Der Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA ist operationalisiert als Anteil der Patienten, die in Woche 12 einen PGA-Wert von 0 oder 1 und eine Verbesserung um mindestens zwei Stufen erreicht haben.</td></tr>
<tr><td>PERFECT-2</td><td>Siehe PERFECT-1</td></tr>
<tr><td>SATISFY</td><td>Siehe PERFECT-1 mit folgenden Abweichungen:
Der PGA-Wert wurde zur Screening- und Baseline-Visite sowie in Woche 12, 20, 26, 38 und 52 bzw. EOT-Visite erhoben.</td></tr>
<tr><td colspan="2">EOT: End of Treatment; PGA: Physician’s Global Assessment</td></tr>
</table>

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 4-19: Bewertung des Verzerrungspotenzials für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Die ITTT-Population umfasst mindestens 98 % der Patienten der ITT-Population der jeweiligen Behandlungsarme. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA ist somit als niedrig einzustufen.

Das Verzerrungspotential der ergänzend dargestellten Studie SATISFY ist bereits auf Studienebene als hoch einzustufen, somit erfolgt keine Bewertung des Verzerrungspotentials auf Endpunktebene.

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.

Die Ergebnisse zum Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA aus den Studien PERFECT-1 und PERFECT-2 sind in der nachfolgenden Tabelle 4-20 dargestellt.

Nach zwölfwöchiger Behandlung mit Trifaroten erreichen in den Studien PERFECT-1 und -2 35,7 % bzw. 42,6 % der Patienten das Therapieziel PGA 0/1, d. h. eine vollständig abgeheilte oder nahezu abgeheilte Haut am Rumpf. Im Vehikel-Arm trifft dies lediglich auf 25,0 % bzw. 29,9 % der Patienten zu. Der Behandlungsunterschied im Vergleich zu Vehikel ist in beiden Studien statistisch signifikant zugunsten von Trifaroten (RR [95 %-KI]: 1,41 [1,18; 1,69], $p < 0,001$ bzw. 1,50 [1,29; 1,75], $p < 0,001$).

Tabelle 4-20: Ergebnisse zum Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA aus RCT (Woche 12, PERFECT-1 und -2)

Studie	**Trifaroten**			**Vehikel**			**Trifaroten vs. Vehikel**		
	N	**N′**	**Patienten mit PGA 0/1 (%)**[1]	**N**	**N′**	**Patienten mit PGA 0/1 (%)**[1]	**OR**[2] **[95 %-KI]**	**RR**[2] **[95 %-KI]**	**ARR**[3] **[95 %-KI]**
PERFECT-1	600	533	35,7	585	530	25,0	1,73 [1,31; 2,29]	1,41 [1,18; 1,69]	10,74 [5,37; 16,1]
								$p < 0,001$[4]	
PERFECT-2	598	559	42,6	609	572	29,9	2,06 [1,58; 2,69]	1,50 [1,29; 1,75]	12,70 [7,21; 18,2]
								$p < 0,001$[4]	

(1) Fehlende Daten wurden mittels multipler Imputation ersetzt.

(2) Odds Ratios, Relative Risiken und die zugehörigen Konfidenzintervalle wurden basierend auf einem CMH-Test stratifiziert nach Zentrum berechnet. Die Effektschätzer wurden für die multipel imputierten Datensätze berechnet und mittels Rubins Methode zusammengefasst.

(3) Risikodifferenzen und die zugehörigen Konfidenzintervalle basieren auf einem Approximationsverfahren für binäre Daten innerhalb einer großen Stichprobe und ohne Anwendung einer Nullzellenkorrektur. Die Effektschätzer wurden für die multipel imputierten Datensätze berechnet und mittels Rubins Methode zusammengefasst.

(4) Die p-Werte basieren auf einem CMH-Test stratifiziert nach Zentrum. Der CMH-Test wurde für jeden multipel imputierten Datensatz berechnet und mittels Rubins Methode unter Anwendung der Wilson-Hilferty-Transformation zusammengefasst.

ARR: Absolute Risikoreduktion in %; CMH: Cochran-Mantel-Haenszel; PGA: *Physician´s Global Assessment*; N: Patienten in der ITTT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; OR: Odds Ratio; RR: Relatives Risiko

In der Langzeitstudie SATISFY zeigt sich zudem, dass der Anteil der Patienten mit PGA 0/1 bei einer Behandlung über Woche 12 hinaus weiter zunimmt (siehe Tabelle 4-21).

Unter Berücksichtigung der Patienten, für die zum jeweiligen Auswertungszeitpunkt Daten vorlagen, erreichen nach zwölf Wochen 38,6 % der Patienten das Therapieziel PGA 0/1. Diese Rate steigt bis Woche 52 auf 66,9 % an. Somit zeigt sich bei etwa zwei Drittel der Patienten, die eine Behandlung mit Trifaroten beginnen, nach 52-wöchiger Behandlung eine vollständig oder nahezu vollständig abgeheilte Haut am Rumpf.

Tabelle 4-21: Ergänzende Ergebnisse zum Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA (Woche 12 und 52, Studie SATISFY)

Studie	**Zeitpunkt**	**N**	**N′**	**Anteil Patienten mit PGA 0/1 n (%)**
SATISFY	Woche 12	444	396	153 (38,6)
	Woche 52	444	344	230 (66,9)

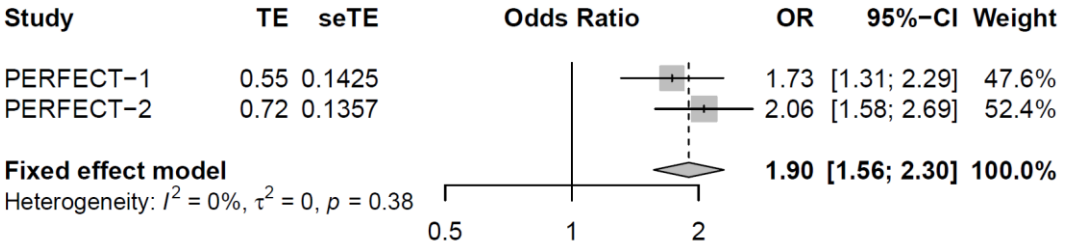
PGA: *Physician´s Global Assessment*; N: Patienten in der SAFT-Population; N′: Patienten mit Wert in Woche 12 bzw. 52

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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-4 sind die Meta-Analysen zum Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel hinsichtlich der Therapieerfolgsrate am Rumpf nach zwölf Wochen Behandlung (RR [95 %-KI]: 1,46 [1,30; 1,64]). Das Ergebnis der Meta-Analyse spiegelt somit die Einzelergebnisse der Studien wider.

a)



b)

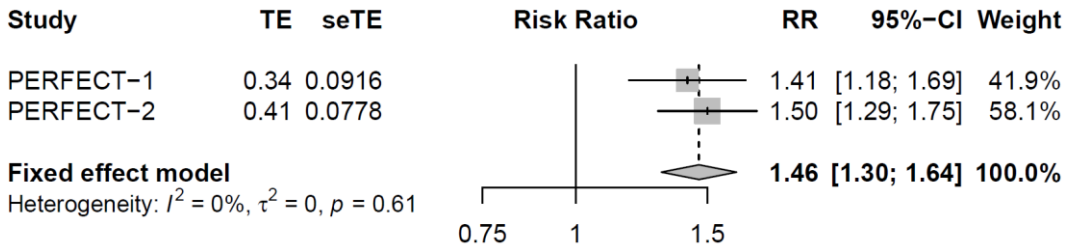


Abbildung 4-4: Meta-Analyse für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

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.

4.3.1.3.3 Änderung der Anzahl der Läsionen im Gesicht – RCT

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

Tabelle 4-22: Operationalisierung der Änderung der Anzahl der Läsionen im Gesicht

Studie	Operationalisierung
PERFECT-1	Die Anzahl der Läsionen im Gesicht wurde von geschulten Prüfern durch visuelle Begutachtung ggf. mit Ertasten der Läsionen bestimmt. Dabei wurde zwischen entzündlichen und nicht entzündlichen Läsionen unterschieden. Diese waren wie folgt definiert: • Entzündliche Läsionen o Papeln: kleine, rote, feste Erhebung mit einem Durchmesser < 0,5 cm o Pusteln: kleine, begrenzte Erhebung, die gelblich-weiße Flüssigkeit enthält • Nicht entzündliche Läsionen o Offene Komedonen: pigmentierte, erweiterte Talgdrüsenöffnung („*blackhead*“) o Geschlossene Komedonen: kleines, weißes Knötchen („*whitehead*“) Die Anzahl der Läsionen wurde zur Screening- und Baseline-Visite sowie in Woche 1, 2, 4, 8 und 12 bestimmt. Folgende Auswertungen wurden durchgeführt: • Absolute Änderung der Anzahl entzündlicher Läsionen im Gesicht in Woche 12 im Vergleich zum Ausgangswert • Absolute Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht in Woche 12 im Vergleich zum Ausgangswert • Prozentuale Änderung der Anzahl entzündlicher Läsionen im Gesicht in Woche 12 im Vergleich zum Ausgangswert • Prozentuale Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht in Woche 12 im Vergleich zum Ausgangswert
PERFECT-2	Siehe PERFECT-1
SATISFY	Der Endpunkt wurde nicht erhoben.

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 4-23: Bewertung des Verzerrungspotenzials für den Endpunkt Änderung der Anzahl der Läsionen im Gesicht in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Die Endpunkte wurden verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Änderung der Anzahl der Läsionen im Gesicht ist somit als niedrig einzustufen.

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.

Die Ergebnisse aus den Studien PERFECT-1 und PERFECT-2 zur Änderung der Anzahl der entzündlichen und nicht entzündlichen Läsionen im Gesicht sind in Tabelle 4-24 als absolute Änderungen und in Tabelle 4-25 als prozentuale Änderungen dargestellt.

Tabelle 4-24: Ergebnisse zur absoluten Änderung der Anzahl der Läsionen im Gesicht aus RCT (Woche 12, PERFECT-1 und -2)

Studie	Trifaroten					Vehikel					Trifaroten vs. Vehikel
	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Absolute Änderung LSM (SE)	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Absolute Änderung LSM (SE)	LSMD(1) [95 %-KI] p-Wert
Absolute Änderung entzündlicher Läsionen											
PERFECT-1	612	543	34,7 (13,02)	15,1 (12,34)	-19,0 (0,50)	596	539	34,8 (13,61)	18,9 (15,34)	-15,4 (0,51)	-3,6 [-4,9; -2,2] < 0,001
PERFECT-2	602	563	36,1 (12,47)	11,7 (12,57)	-24,2 (0,51)	610	574	37,1 (15,06)	17,3 (13,85)	-18,7 (0,51)	-5,6 [-6,9; -4,3] < 0,001
Absolute Änderung nicht entzündlicher Läsionen											
PERFECT-1	612	543	54,0 (28,55)	27,1 (25,41)	-25,0 (0,87)	596	539	52,8 (26,08)	33,7 (28,27)	-17,9 (0,87)	-7,1 [-9,4; -4,8] < 0,001
PERFECT-2	602	563	50,6 (25,93)	20,1 (17,22)	-30,1 (0,71)	610	574	51,2 (25,75)	28,2 (23,06)	-21,6 (0,71)	-8,5 [-10,3; -6,6] < 0,001

(1) Fehlende Daten wurden mittels multipler Imputation ersetzt. P-Werte und Konfidenzintervalle basieren auf einem ANCOVA-Modell mit Anzahl der Läsionen zu Studienbeginn, Zentrum und Behandlung als Effekte.

ANCOVA: *analysis of covariance*; LSM: *least squares mean*; LSMD: *least squares mean difference*; MW: Mittelwert; N: Patienten in der ITT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; SD: Standardabweichung (*standard deviation*); SE: Standardfehler (*standard error*)

Tabelle 4-25: Ergebnisse zur prozentualen Änderung der Anzahl der Läsionen im Gesicht aus RCT (Woche 12, PERFECT-1 und -2)

Studie	Trifaroten					Vehikel					Trifaroten vs. Vehikel
	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Prozentuale Änderung LSM (SE)	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Prozentuale Änderung LSM (SE)	LSMD[(1)] [95 %-KI] p-Wert
Prozentuale Änderung entzündlicher Läsionen											
PERFECT-1	612	543	34,7 (13,02)	15,1 (12,34)	-54,0 (1,47)	596	539	34,8 (13,61)	18,9 (15,34)	-44,4 (1,49)	-9,6 [-13,5; -5,7] < 0,001
PERFECT-2	602	563	36,1 (12,47)	11,7 (12,57)	-65,5 (1,41)	610	574	37,1 (15,06)	17,3 (13,85)	-49,5 (1,41)	-16,0 [-19,7; -12,4] < 0,001
Prozentuale Änderung nicht entzündlicher Läsionen											
PERFECT-1	612	543	54,0 (28,55)	27,1 (25,41)	-48,1 (1,56)	596	539	52,8 (26,08)	33,7 (28,27)	-34,1 (1,57)	-14,0 [-18,1; -9,8] < 0,001
PERFECT-2	602	563	50,6 (25,93)	20,1 (17,22)	-58,1 (1,36)	610	574	51,2 (25,75)	28,2 (23,06)	-42,6 (1,36)	-15,5 [-19,0; -12,0] < 0,001

(1) Fehlende Daten wurden mittels multipler Imputation ersetzt. P-Werte und Konfidenzintervalle basieren auf einem ANCOVA-Modell mit Anzahl der Läsionen zu Studienbeginn, Zentrum und Behandlung als Effekte.

ANCOVA: *analysis of covariance*; LSM: *least squares mean*; LSMD: *least squares mean difference*; MW: Mittelwert; N: Patienten in der ITT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; SD: Standardabweichung (*standard deviation*); SE: Standardfehler (*standard error*)

Zu Studienbeginn weisen die Patienten im Mittel 35 bis 37 entzündliche und 51 bis 54 nicht entzündliche Läsionen im Gesicht auf. Die Auswertungen zeigen in allen Studienarmen eine Reduktion der Anzahl der Läsionen im Gesicht im Verlauf der zwölfwöchigen Behandlung.

Unter Trifaroten ist eine statistisch signifikant stärkere Verringerung der Zahl der entzündlichen Läsionen im Vergleich zu Vehikel in beiden Studien zu beobachten (LSMD [95 %-KI]: -3,6 [-4,9; -2,2], $p < 0,001$ bzw. -5,6 [-6,9; -4,3], $p < 0,001$). Auch die Zahl der nicht entzündlichen Läsionen verringert sich unter Trifaroten statistisch signifikant stärker als unter Vehikel (LSMD [95 %-KI]: -7,1 [-9,4; -4,8], $p < 0,001$ bzw. -8,5 [-10,3; -6,6], $p < 0,001$).

Die prozentuale Änderung der Anzahl der entzündlichen Läsionen beträgt 54,0 % bzw. 65,5 % unter Trifaroten gegenüber 44,4 % bzw. 49,5 % unter Vehikel. Der Behandlungsunterschied ist in beiden Studien statistisch signifikant (LSMD [95 %-KI]: -9,6 [-13,5; -5,7], $p < 0,001$ bzw. -16,0 [-19,7; -12,4], $p < 0,001$). Auch die prozentuale Änderung der Anzahl der nicht entzündlichen Läsionen ist in beiden Studien unter Trifaroten statistisch signifikant größer (LSMD [95 %-KI]: -14,0 [-18,1; -9,8], $p < 0,001$ bzw. -15,5 [-19,0; -12,0], $p < 0,001$); unter Trifaroten verringert sich die Zahl der nicht entzündlichen Läsionen um 48,1 % bzw. 58,1 % gegenüber 34,1 % bzw. 42,6 % unter Vehikel.

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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-5 und Abbildung 4-6 sind die Meta-Analysen zur absoluten und prozentualen Änderung der Anzahl der Läsionen im Gesicht dargestellt.

Für die beiden Auswertungen zur Änderung der Anzahl der entzündlichen Läsionen zeigt sich anhand der Heterogenitätstests jeweils eine bedeutsame Heterogenität der Effekte ($p = 0,04$ bzw. $p = 0,02$). Die Ursache für eine solche Heterogenität kann gegebenenfalls auf methodische oder klinische Unterschiede zwischen den Studien zurückzuführen sein, andernfalls liegt eine zufallsbedingte Variabilität vor. Da die Studien ein identisches Studiendesign aufweisen, ist eine methodische Heterogenität eher unwahrscheinlich. Die bekannten Patientencharakteristika sind hinreichend ähnlich, allerdings gibt es regionale Unterschiede zwischen den Studien. Ob diese die Ursache für die festgestellte Heterogenität darstellen oder ob es sich um zufallsbedingte Unterschiede handelt, ist unklar.

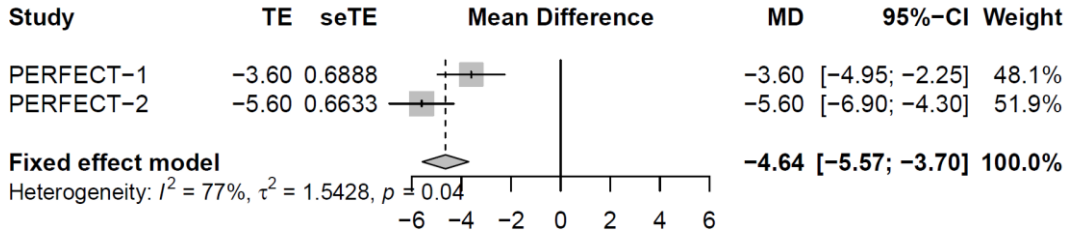
Da die Einzelergebnisse der Studien jedoch gleichgerichtet und deutlich statistisch signifikant sind, ist die Bildung eines gemeinsamen Effektschätzers dennoch plausibel. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich sowohl hinsichtlich der absoluten als auch der prozentualen Änderung der Anzahl der entzündlichen Läsionen jeweils ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel (LSMD [95 %-KI]: -4,64 [-5,57; -3,70] bzw. -13,01 [-15,68; -10,35]).

Für die beiden Auswertungen zur Änderung der Anzahl der nicht entzündlichen Läsionen zeigt sich anhand der Heterogenitätstests keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich sowohl hinsichtlich der absoluten als auch der prozentualen Änderung der Anzahl der nicht entzündlichen Läsionen jeweils ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel (LSMD [95 %-KI]: -7,95 [-9,39; -6,51] bzw. -14,88 [-17,55; -12,20]).

Zusammenfassend bestätigen die Ergebnisse der Meta-Analyse die Einzelergebnisse der beiden Studien.

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

a)



b)

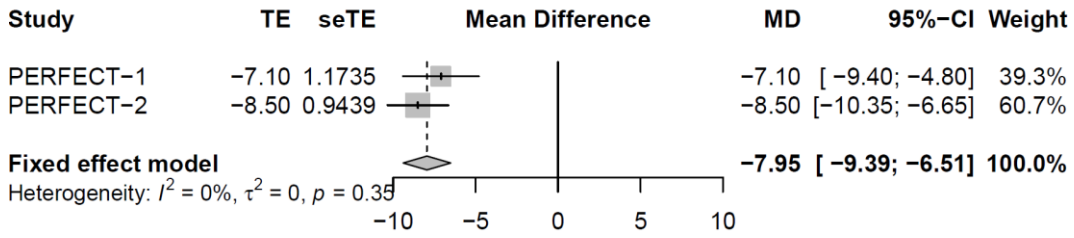
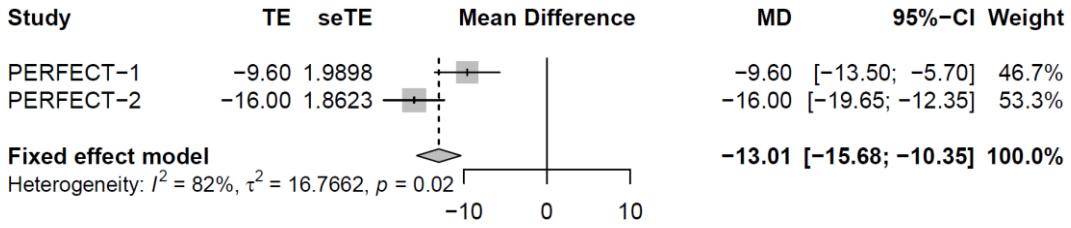


Abbildung 4-5: Meta-Analyse für den Endpunkt absolute Änderung der Anzahl der Läsionen im Gesicht aus RCT; Trifaroten vs. Vehikel
a) entzündliche Läsionen, b) nicht entzündliche Läsionen

a)



b)

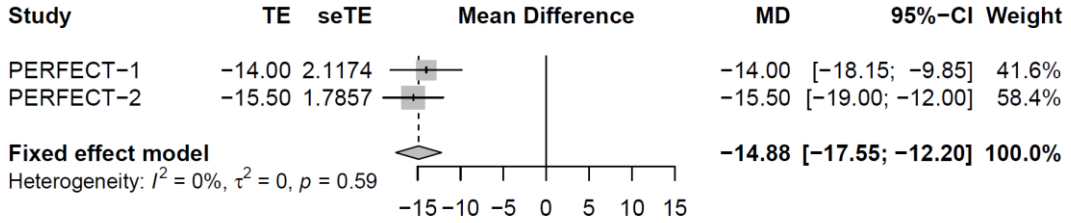


Abbildung 4-6: Meta-Analyse für den Endpunkt prozentuale Änderung der Anzahl der Läsionen im Gesicht aus RCT; Trifaroten vs. Vehikel
a) entzündliche Läsionen, b) nicht entzündliche Läsionen

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.

4.3.1.3.4 Änderung der Anzahl der Läsionen am Rumpf – RCT

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

Tabelle 4-26: Operationalisierung der Änderung der Anzahl der Läsionen am Rumpf

Studie	Operationalisierung
PERFECT-1	Die Anzahl der Läsionen am Rumpf wurde von geschulten Prüfern durch visuelle Begutachtung ggf. mit Ertasten der Läsionen bestimmt. Dabei wurde zwischen entzündlichen und nicht entzündlichen Läsionen unterschieden. Diese waren wie folgt definiert: • Entzündliche Läsionen o Papeln: kleine, rote, feste Erhebung mit einem Durchmesser < 0,5 cm o Pusteln: kleine, begrenzte Erhebung, die gelblich-weiße Flüssigkeit enthält • Nicht entzündliche Läsionen o Offene Komedonen: pigmentierte, erweiterte Talgdrüsenöffnung („*blackhead*“) o Geschlossene Komedonen: kleines, weißes Knötchen („*whitehead*“) Die Anzahl der Läsionen wurde zur Screening- und Baseline-Visite sowie in Woche 1, 2, 4, 8 und 12 bestimmt. Folgende Auswertungen wurden durchgeführt: • Absolute Änderung der Anzahl entzündlicher Läsionen am Rumpf in Woche 12 im Vergleich zum Ausgangswert • Absolute Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf in Woche 12 im Vergleich zum Ausgangswert • Prozentuale Änderung der Anzahl entzündlicher Läsionen am Rumpf in Woche 12 im Vergleich zum Ausgangswert • Prozentuale Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf in Woche 12 im Vergleich zum Ausgangswert
PERFECT-2	Siehe PERFECT-1
SATISFY	Der Endpunkt wurde nicht erhoben.

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 4-27: Bewertung des Verzerrungspotenzials für den Endpunkt Änderung der Anzahl der Läsionen am Rumpf in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Die Endpunkte wurden verblindet erhoben und auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Die ITTT-Population umfasst mindestens 98 % der Patienten der ITT-Population der jeweiligen Behandlungsarme. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Änderung der Anzahl der Läsionen am Rumpf ist somit als niedrig einzustufen.

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.

Die Ergebnisse aus den Studien PERFECT-1 und PERFECT-2 zur Änderung der Anzahl der entzündlichen und nicht entzündlichen Läsionen am Rumpf sind in Tabelle 4-28 als absolute Änderung und in Tabelle 4-29 als prozentuale Änderung dargestellt.

Tabelle 4-28: Ergebnisse zur absoluten Änderung der Anzahl der Läsionen am Rumpf aus RCT (Woche 12, PERFECT-1 und -2)

Studie	Trifaroten					Vehikel					Trifaroten vs. Vehikel
	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Absolute Änderung LSM (SE)	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Absolute Änderung LSM (SE)	LSMD(1) [95 %-KI] p-Wert
Absolute Änderung entzündlicher Läsionen											
PERFECT-1	600	533	37,5 (17,32)	15,7 (14,47)	-21,4 (0,54)	585	530	36,2 (16,19)	17,5 (14,98)	-18,8 (0,55)	-2,5 [-4,0; -1,1] < 0,001
PERFECT-2	598	557	39,3 (15,93)	13,3 (13,93)	-25,5 (0,59)	609	573	39,1 (17,36)	18,5 (17,38)	-19,8 (0,58)	-5,7 [-7,2; -4,2] < 0,001
Absolute Änderung nicht entzündlicher Läsionen											
PERFECT-1	600	533	47,0 (20,82)	23,9 (24,28)	-21,9 (0,93)	585	530	48,3 (21,29)	28,5 (26,65)	-17,8 (0,94)	-4,1 [-6,6; -1,7] 0,001
PERFECT-2	598	558	46,4 (19,93)	20,0 (18,55)	-25,9 (0,67)	609	573	45,8 (19,51)	24,1 (18,14)	-20,8 (0,66)	-5,0 [-6,8; -3,3] < 0,001

(1) Fehlende Daten wurden mittels multipler Imputation ersetzt. P-Werte und Konfidenzintervalle basieren auf einem ANCOVA-Modell mit Anzahl der Läsionen zu Studienbeginn, Zentrum und Behandlung als Effekte.

ANCOVA: *analysis of covariance*; LSM: *least squares mean*; LSMD: *least squares mean difference*; MW: Mittelwert; N: Patienten in der ITTT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; SD: Standardabweichung (*standard deviation*); SE: Standardfehler (*standard error*)

Tabelle 4-29: Ergebnisse zur prozentualen Änderung der Anzahl der Läsionen am Rumpf aus RCT (Woche 12, PERFECT-1 und -2)

Studie	Trifaroten					Vehikel					Trifaroten vs. Vehikel
	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Prozentuale Änderung LSM (SE)	N	N′	Anzahl zu Studienbeginn MW (SD)	Anzahl in Woche 12 MW (SD)	Prozentuale Änderung LSM (SE)	LSMD[(1)] [95 %-KI] p-Wert
Prozentuale Änderung entzündlicher Läsionen											
PERFECT-1	600	533	37,5 (17,32)	15,7 (14,47)	-57,8 (1,51)	585	530	36,2 (16,19)	17,5 (14,98)	-50,3 (1,54)	-7,5 [-11,5; -3,5] < 0,001
PERFECT-2	598	557	39,3 (15,93)	13,3 (13,93)	-64,1 (1,82)	609	573	39,1 (17,36)	18,5 (17,38)	-48,9 (2,00)	-15,1 [-20,1; -10,2] < 0,001
Prozentuale Änderung nicht entzündlicher Läsionen											
PERFECT-1	600	533	47,0 (20,82)	23,9 (24,28)	-48,0 (1,76)	585	530	48,3 (21,29)	28,5 (26,65)	-39,3 (1,75)	-8,7 [-13,3; -4,1] < 0,001
PERFECT-2	598	558	46,4 (19,93)	20,0 (18,55)	-55,6 (1,50)	609	573	45,8 (19,51)	24,1 (18,14)	-43,2 (1,49)	-12,4 [-16,3; -8,5] < 0,001

(1) Fehlende Daten wurden mittels multipler Imputation ersetzt. P-Werte und Konfidenzintervalle basieren auf einem ANCOVA-Modell mit Anzahl der Läsionen zu Studienbeginn, Zentrum und Behandlung als Effekte.

ANCOVA: *analysis of covariance*; LSM: *least squares mean*; LSMD: *least squares mean difference*; MW: Mittelwert; N: Patienten in der ITTT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; SD: Standardabweichung (*standard deviation*); SE: Standardfehler (*standard error*)

Zu Studienbeginn weisen die Patienten im Mittel 36 bis 40 entzündliche und 46 bis 48 nicht entzündliche Läsionen am Rumpf auf. Die Auswertungen zeigen in allen Studienarmen eine Reduktion der Anzahl der Läsionen am Rumpf im Verlauf der zwölfwöchigen Behandlung.

Unter Trifaroten ist eine statistisch signifikant stärkere Verringerung der Zahl der entzündlichen Läsionen im Vergleich zu Vehikel in beiden Studien zu beobachten (LSMD [95 %-KI]: -2,5 [-4,0; -1,1], $p < 0,001$ bzw. -5,7 [-7,2; -4,2], $p < 0,001$). Auch die Zahl der nicht entzündlichen Läsionen verringert sich unter Trifaroten statistisch signifikant stärker als unter Vehikel (LSMD [95 %-KI]: -4,1 [-6,6; -1,7], $p < 0,001$ bzw. -5,0 [-6,8; -3,3], $p < 0,001$).

Die prozentuale Änderung der Anzahl der entzündlichen Läsionen beträgt 57,8 % bzw. 64,1 % unter Trifaroten gegenüber 50,3 % bzw. 48,9 % unter Vehikel. Der Behandlungsunterschied ist in beiden Studien statistisch signifikant (LSMD [95 %-KI]: -7,5 [-11,5; -3,5], $p < 0,001$ bzw. -15,1 [-20,1; -10,2], $p < 0,001$). Auch die prozentuale Änderung der Anzahl der nicht entzündlichen Läsionen ist unter Trifaroten in beiden Studien statistisch signifikant größer (LSMD [95 %-KI]: -8,7 [-13,3; -4,1], $p < 0,001$ bzw. -12,4 [-16,3; -8,5], $p < 0,001$); unter Trifaroten verringert sich die Zahl der nicht entzündlichen Läsionen um 48,0 % bzw. 55,6 % gegenüber 39,3 % bzw. 43,2 % unter Vehikel.

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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-7 und Abbildung 4-8 sind die Meta-Analysen zur absoluten und prozentualen Änderung der Anzahl der Läsionen am Rumpf dargestellt.

Für die beiden Auswertungen zur Änderung der Anzahl der entzündlichen Läsionen zeigt sich anhand der Heterogenitätstests jeweils eine bedeutsame Heterogenität der Effekte ($p < 0,01$ bzw. $p = 0,02$). Die Ursache für eine solche Heterogenität kann gegebenenfalls auf methodische oder klinische Unterschiede zwischen den Studien zurückzuführen sein, andernfalls liegt eine zufallsbedingte Variabilität vor. Da die Studien ein identisches Studiendesign aufweisen, ist eine methodische Heterogenität eher unwahrscheinlich. Die bekannten Patientencharakteristika sind hinreichend ähnlich, allerdings gibt es regionale Unterschiede zwischen den Studien. Ob diese die Ursache für die festgestellte Heterogenität darstellen oder ob es sich um zufallsbedingte Unterschiede handelt, ist unklar.

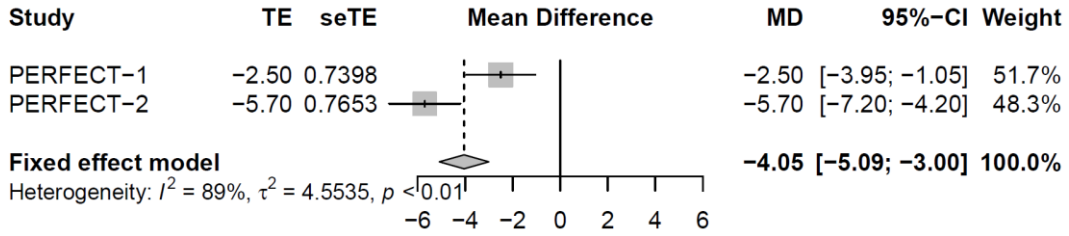
Da die Einzelergebnisse der Studien jedoch gleichgerichtet und deutlich statistisch signifikant sind, ist die Bildung eines gemeinsamen Effektschätzers dennoch plausibel. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich sowohl hinsichtlich der absoluten als auch der prozentualen Änderung der Anzahl der entzündlichen Läsionen jeweils ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel (LSMD [95 %-KI]: -4,05 [-5,09; -3,00] bzw. -10,50 [-13,61; -7,39]).

Für die beiden Auswertungen zur Änderung der Anzahl der nicht entzündlichen Läsionen zeigt sich anhand der Heterogenitätstests keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich sowohl hinsichtlich der absoluten als auch der prozentualen Änderung der Anzahl der nicht entzündlichen Läsionen jeweils ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel (LSMD [95 %-KI]: -4,70 [-6,12; -3,27] bzw. -10,85 [-13,83; -7,88]).

Zusammenfassend bestätigen die Ergebnisse der Meta-Analyse die Einzelergebnisse der beiden Studien.

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

a)



b)

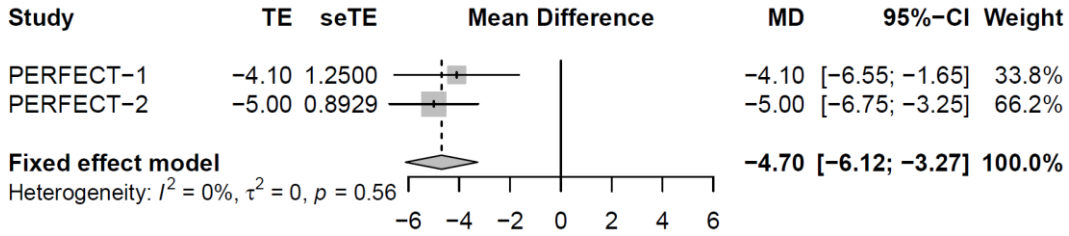
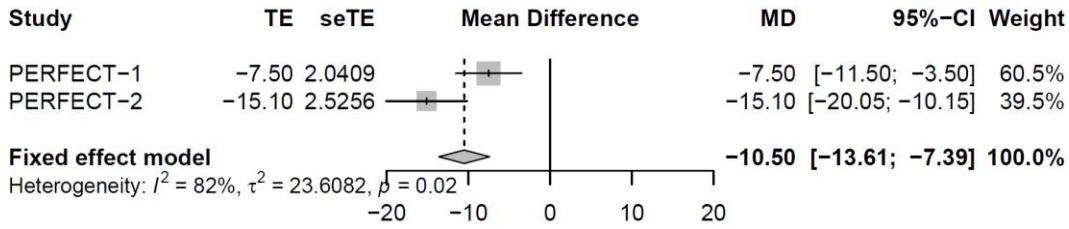


Abbildung 4-7: Meta-Analyse für den Endpunkt absolute Änderung der Anzahl der Läsionen am Rumpf aus RCT; Trifaroten vs. Vehikel
a) entzündliche Läsionen, b) nicht entzündliche Läsionen

a)



b)

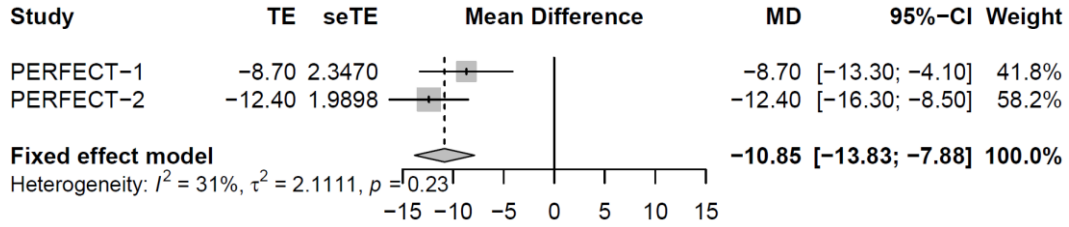


Abbildung 4-8: Meta-Analyse für den Endpunkt prozentuale Änderung der Anzahl der Läsionen am Rumpf aus RCT; Trifaroten vs. Vehikel
a) entzündliche Läsionen, b) nicht entzündliche Läsionen

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.

4.3.1.3.5 Selbsteinschätzung der Verbesserung der Akne – RCT

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

Tabelle 4-30: Operationalisierung der Selbsteinschätzung der Verbesserung der Akne

Studie	Operationalisierung
PERFECT-1	Die Patienten bewerteten die Verbesserung ihrer Akne im Gesicht in Woche 12 bzw. nach Ende der Behandlung zur EOT-Visite im Vergleich zum Studienbeginn anhand einer sechsstufigen Skala: 0: vollständige Verbesserung (*complete improvement*) 1: deutliche Verbesserung (*marked improvement*) 2: moderate Verbesserung (*moderate improvement*) 3: leichte Verbesserung (*minimal improvement)* 4: keine Änderung (*no change*) 5: Verschlechterung (*worse*) Die Selbsteinschätzung der Patienten wurde vor der Einschätzung durch den Prüfarzt erhoben. Die Selbsteinschätzung der Verbesserung wurde in der Studie deskriptiv ausgewertet. Post hoc wurde folgende Auswertung vorgenommen: • Anteil der Patienten mit einer vollständigen oder deutlichen Verbesserung der Akne in Woche 12
PERFECT-2	Siehe PERFECT-1
SATISFY	Siehe PERFECT-1 mit folgenden Abweichungen: Die Selbsteinschätzung der Patienten wurde in Woche 12, 26 und 52 bzw. zur EOT-Visite erhoben.
EOT: *End of Treatment*	

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 4-31: Bewertung des Verzerrungspotenzials für den Endpunkt Selbsteinschätzung der Verbesserung der Akne in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Selbsteinschätzung der Verbesserung der Akne ist somit als niedrig einzustufen.

Das Verzerrungspotential der ergänzend dargestellten Studie SATISFY ist bereits auf Studienebene als hoch einzustufen, somit erfolgt keine Bewertung des Verzerrungspotentials auf Endpunktebene.

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.

Die Ergebnisse aus den Studien PERFECT-1 und PERFECT-2 zur Selbsteinschätzung der Verbesserung der Akne sind in der nachfolgenden Tabelle 4-32 dargestellt.

Nach zwölfwöchiger Behandlung mit Trifaroten geben 36,1 % bzw. 42,0 % der Patienten eine vollständige oder deutliche Besserung der Akne an. Unter Vehikel trifft dies lediglich auf 22,8 % bzw. 27,4 % der Patienten zu. Der Behandlungsunterschied ist in beiden Studien zugunsten von Trifaroten statistisch signifikant (RR [95 %-KI]: 1,57 [1,30; 1,88], $p < 0,001$ bzw. 1,56 [1,32; 1,83], $p < 0,001$).

Tabelle 4-32: Ergebnisse zum Endpunkt Selbsteinschätzung der Verbesserung der Akne aus RCT (Woche 12, PERFECT-1 und -2)

Studie	Trifaroten			Vehikel			Trifaroten vs. Vehikel		
	N	N′	Patienten mit Verbes-serung (%)[1,2]	N	N′	Patienten mit Verbes-serung (%)[1,2]	OR[3] [95 %-KI]	RR[3] [95 %-KI]	ARR[4] [95 %-KI]
PERFECT-1	612	542	36,1	596	537	22,8	1,94 [1,49; 2,53]	1,57 [1,30; 1,88]	13,29 [8,21; 18,4]
								p < 0,001[5]	
PERFECT-2	602	562	42,0	610	574	27,4	2,01 [1,56; 2,59]	1,56 [1,32; 1,83]	14,65 [9,35; 20,0]
								p < 0,001[5]	

(1) Patienten mit einer vollständigen oder deutlichen Verbesserung der Akne

(2) Fehlende Daten wurden mittels Non-Responder-Imputation ersetzt.

(3) Odds Ratios, Relative Risiken und die zugehörigen Konfidenzintervalle wurden basierend auf einem CMH-Test stratifiziert nach Zentrum berechnet.

(4) Risikodifferenzen und die zugehörigen Konfidenzintervalle basieren auf einem Approximationsverfahren für binäre Daten innerhalb einer großen Stichprobe und ohne Anwendung einer Nullzellenkorrektur.

(5) Die p-Werte basieren auf einem CMH-Test stratifiziert nach Zentrum.

ARR: Absolute Risikoreduktion in %; CMH: Cochran-Mantel-Haenszel; N: Patienten in der ITT-Population; N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; OR: Odds Ratio; RR: Relatives Risiko

In der Langzeitstudie SATISFY zeigt sich zudem, dass der Anteil der Patienten mit einer vollständigen oder deutlichen Verbesserung der Akne bei einer Behandlung über Woche 12 hinaus weiter zunimmt (siehe Tabelle 4-33).

Unter Berücksichtigung der Patienten, für die zum jeweiligen Auswertungszeitpunkt Daten vorlagen, geben nach zwölf Wochen 41,1 % der Patienten eine mindestens deutliche Verbesserung ihrer Akne an. Diese Rate steigt bis Woche 52 auf 66,6 % an. Somit zeigt sich, dass etwa zwei Drittel der Patienten, die eine Behandlung mit Trifaroten beginnen, nach 52-wöchiger Behandlung eine vollständige oder deutliche Besserung der Akne erreichen. Diese Zahlen entsprechen in etwa den Therapieerfolgsraten gemäß IGA und PGA (siehe Abschnitte 4.3.1.3.1 und 4.3.1.3.2), wodurch die Übereinstimmung der Einschätzung der Prüfer mit der Selbsteinschätzung der Patienten bestätigt wird.

Tabelle 4-33: Ergänzende Ergebnisse zum Endpunkt Selbsteinschätzung der Verbesserung der Akne (Woche 12 und 52, Studie SATISFY)

Studie	**Zeitpunkt**	**N**	**N′**	**Anteil Patienten mit vollständiger/deutlicher Verbesserung n (%)**
SATISFY	Woche 12	453	401	166 (41,4)
	Woche 52	453	350	233 (66,6)
N: Patienten in der SAF-Population; N′: Patienten mit Wert in Woche 12 bzw. 52				

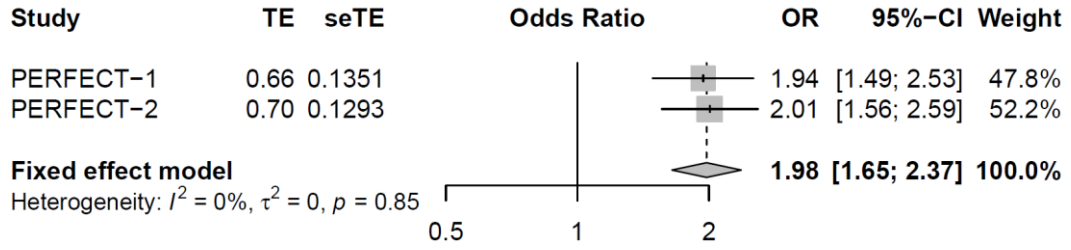
Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse 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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-9 sind die Meta-Analysen zum Endpunkt Selbsteinschätzung der Verbesserung der Akne dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität ($p > 0,05$), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich ein statistisch signifikanter Vorteil für Trifaroten gegenüber Vehikel hinsichtlich des Anteils der Patienten, die nach zwölf Wochen Behandlung eine mindestens deutliche Verbesserung ihrer Akne erreichen (RR [95 %-KI]: 1,56 [1,38; 1,77]). Das Ergebnis der Meta-Analyse spiegelt somit die Einzelergebnisse der Studien wider.

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

a)



b)

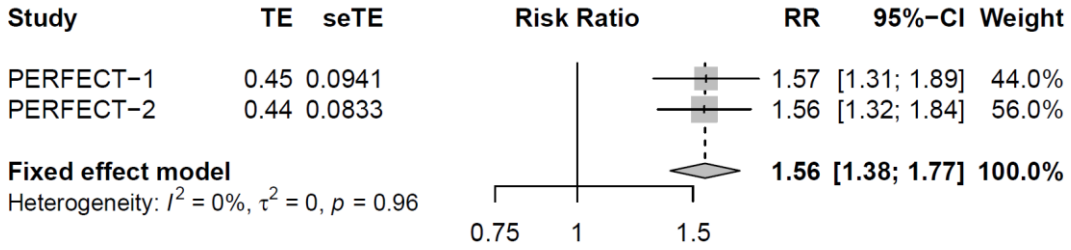


Abbildung 4-9: Meta-Analyse für den Endpunkt Selbsteinschätzung der Verbesserung der Akne aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

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.

4.3.1.3.6 Gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI – RCT

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

Tabelle 4-34: Operationalisierung von gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI

Studie	Operationalisierung
PERFECT-1	Die Patienten bewerteten ihre gesundheitsbezogene Lebensqualität zur Baseline-Visite sowie in Woche 12 bzw. nach Ende der Behandlung zur EOT-Visite anhand des Fragebogens *Dermatology Life Quality Index* (DLQI). Patienten bis einschließlich 16 Jahre nutzten die Version CDLQI (*Children´s Dermatology Life Quality Index*). Der DLQI / CDLQI ist ein Fragebogen mit 10 Fragen zur Ermittlung der krankheitsspezifischen Lebensqualität bei Patienten mit Hauterkrankungen. Die 10 Fragen beziehen sich auf folgende sechs Domänen: Symptome und Befinden, tägliche Aktivitäten, Freizeit, Arbeit/Schule, persönliche Beziehungen und Auswirkungen der Behandlung auf den Alltag. Jede Frage hat vier Antwortkategorien von 0 (gar nicht) bis 3 (sehr stark). Der DLQI / CDLQI kann Werte von 0 bis 30 annehmen, der Wert ergibt sich aus der Summe der Antwortkategorien der zehn Fragen. Je höher der Score ist, desto größer ist die Beeinträchtigung der gesundheitsbezogenen Lebensqualität. Sowohl bei Anwendung des DLQI als auch des CDLQI kann ein Score von 0 oder 1 als „nicht eingeschränkte Lebensqualität“ interpretiert werden. Die gesundheitsbezogene Lebensqualität gemäß DLQI oder CDLQI wurde in der Studie deskriptiv ausgewertet. Post hoc wurde folgende Auswertung vorgenommen: • Anteil der Patienten mit DLQI 0/1 in Woche 12 • Anteil der Patienten mit CDLQI 0/1 in Woche 12
PERFECT-2	Siehe PERFECT-1
SATISFY	Siehe PERFECT-1 mit folgenden Abweichungen: Die gesundheitsbezogene Lebensqualität gemäß DLQI bzw. CDLQI wurde zur Baseline-Visite sowie in Woche 12, 26 und 52 bzw. zur EOT-Visite erhoben.
CDLQI *Children´s Dermatology Life Quality Index;* DLQI: *Dermatology Life Quality Index;* EOT: *End of Treatment*	

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 4-35: Bewertung des Verzerrungspotenzials für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. In allen Behandlungsgruppen liegt die Rücklaufquote für den DLQI bzw. den CDLQI in Woche 12 bei über 80 %. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI ist somit als niedrig einzustufen.

Das Verzerrungspotential der ergänzend dargestellten Studie SATISFY ist bereits auf Studienebene als hoch einzustufen, somit erfolgt keine Bewertung des Verzerrungspotentials auf Endpunktebene.

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.

In den Studien PERFECT-1 und -2 wurde die gesundheitsbezogene Lebensqualität mittels der Fragebögen DLQI bzw. CDLQI erhoben.

Zu Studienbeginn weisen die Patienten, die 17 Jahre oder älter waren, im Mittel einen DLQI von 6 bis 7 und die unter 17-Jährigen einen CDLQI von etwa 4 auf (siehe Tabelle 4-36). Diese Werte zeigen eine moderate Einschränkung der Lebensqualität infolge der Akne bei den älteren bzw. eine geringe Beeinträchtigung der Lebensqualität bei den jüngeren Studienteilnehmern. Nach zwölfwöchiger Behandlung sowohl mit Trifaroten als auch mit Vehikel verringern sich die Werte in allen Behandlungsgruppen, was eine Verbesserung der Lebensqualität im Laufe der Behandlung anzeigt.

Tabelle 4-36: Ergebnisse zur mittleren Änderung des DLQI und CDLQI aus RCT (PERFECT-1 und -2)

Studie	Trifaroten				Vehikel			
	N	N′	Baseline-Wert MW (SD)	Wert in Woche 12 MW (SD)	N	N′	Baseline-Wert MW (SD)	Wert in Woche 12 MW (SD)
DLQI								
PERFECT-1	358	306	5,9 (5,31)	3,0 (3,58)	366	321	6,0 (5,27)	2,7 (3,28)
PERFECT-2	384	360	6,9 (5,22)	3,7 (4,34)	374	347	7,0 (5,45)	3,7 (4,38)
CDLQI								
PERFECT-1	254	234	3,9 (3,93)	2,9 (3,08)	230	215	3,9 (3,77)	2,2 (2,65)
PERFECT-2	218	203	4,5 (3,94)	3,1 (3,29)	236	225	4,6 (4,41)	2,7 (3,39)

CDLQI: *Children´s Dermatology Life Quality Index*; DLQI: *Dermatology Life Quality Index*; MW: Mittelwert; N: Patienten in der ITT-Population ≤ 16 Jahre (CDLQI) bzw. > 16 Jahre (DLQI); N′: Patienten mit Wert in Woche 12; SD: Standardabweichung (*standard deviation*)

Für die statistische Auswertung der Verbesserung der gesundheitsbezogenen Lebensqualität wurde die Anzahl der Patienten, die einen DLQI bzw. CDLQI von 0 oder 1 – und somit eine nicht beeinträchtigte Lebensqualität – in Woche 12 erreichen, ermittelt (siehe Tabelle 4-37).

Dabei zeigt sich bei den Patienten ab 17 Jahren, dass 37,2 % bzw. 33,3 % der Patienten unter Trifaroten und 39,3 % bzw. 37,2 % der Patienten unter Vehikel einen DLQI 0/1 erreichen. Der Behandlungsunterschied ist in beiden Studien nicht statistisch signifikant (RR [95 %-KI]: 0,95 [0,79; 1,13], p = 0,549 bzw. 0,90 [0,74; 1,09], p = 0,274).

Bei den unter 17-Jährigen erreichen 37,8 % bzw. 35,3 % der Patienten unter Trifaroten und 47,4 % bzw. 44,1 % der Patienten unter Vehikel einen CDLQI 0/1. In beiden Studien ist der Behandlungsunterschied nicht statistisch signifikant (RR [95 %-KI]: 0,84 [0,67; 1,04], p = 0,110 bzw. 0,82 [0,66; 1,02], p = 0,072).

Tabelle 4-37: Ergebnisse zum Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI aus RCT (Woche 12, PERFECT-1 und -2)

Studie	**Trifaroten**			**Vehikel**			**Trifaroten vs. Vehikel**		
	N	**N′**	**Patienten mit Wert 0/1 n (%)**[1]	**N**	**N′**	**Patienten mit Wert 0/1 n (%)**[1]	**OR**[2] **[95 %-KI]**	**RR**[2] **[95 %-KI]**	**ARR**[3] **[95 %-KI]**
DLQI 0/1									
PERFECT-1	358	306	133 (37,2)	366	321	144 (39,3)	0,90 [0,64; 1,26]	0,95 [0,79; 1,13]	-2,19 [-9,27; 4,88]
							p = 0,549[4]		
PERFECT-2	384	360	128 (33,3)	374	347	139 (37,2)	0,83 [0,60; 1,16]	0,90 [0,74; 1,09]	-3,83 [-10,63; 2,97]
							p = 0,274[4]		
CDLQI 0/1									
PERFECT-1	254	234	96 (37,8)	230	215	109 (47,4)	0,72 [0,49; 1,07]	0,84 [0,67; 1,04]	-9,60 [-18,4; -0,81]
							p = 0,110[4]		
PERFECT-2	218	203	77 (35,3)	236	225	104 (44,1)	0,68 [0,45; 1,04]	0,82 [0,66; 1,02]	-8,75 [-17,71; 0,22]
							p = 0,072[4]		

(1) Fehlende Daten wurden mittels Non-Responder-Imputation ersetzt.

(2) Odds Ratios, Relative Risiken und die zugehörigen Konfidenzintervalle wurden basierend auf einem CMH-Test stratifiziert nach Zentrum berechnet.

(3) Risikodifferenzen und die zugehörigen Konfidenzintervalle basieren auf einem Approximationsverfahren für binäre Daten innerhalb einer großen Stichprobe und ohne Anwendung einer Nullzellenkorrektur.

(4) Die p-Werte basieren auf einem CMH-Test stratifiziert nach Zentrum.

ARR: Absolute Risikoreduktion in %; CDLQI: *Children´s Dermatology Life Quality Index*; CMH: Cochran-Mantel-Haenszel; DLQI: *Dermatology Life Quality Index*; N: Patienten in der ITT-Population ≤ 16 Jahre (CDLQI) bzw. > 16 Jahre (DLQI); N′: Patienten mit Wert in Woche 12; KI: Konfidenzintervall; OR: Odds Ratio; RR: Relatives Risiko

In der Langzeitstudie SATISFY zeigt sich, dass der Anteil der Patienten mit DLQI 0/1 bzw. CDLQI 0/1 bei einer Behandlung über Woche 12 hinaus weiter zunimmt (siehe Tabelle 4-38).

Unter Berücksichtigung der Patienten, für die zum jeweiligen Auswertungszeitpunkt Daten vorlagen, beträgt in Woche 12 der Anteil der Patienten mit DLQI 0/1 37,6 %. Diese Rate steigt bis Woche 52 auf 53,8 % an.

Bei den unter 17-Jährigen beträgt der Anteil der Patienten mit CDLQI 0/1 42,5 % in Woche 12 und 54,2 % in Woche 52.

Tabelle 4-38: Ergänzende Ergebnisse zum Endpunkt Lebensqualität gemäß DLQI / CDLQI (Woche 12 und 52, Studie SATISFY)

Studie	Zeitpunkt	N	N′	Anteil Patienten mit DLQI bzw. CDLQI 0/1 n (%)
DLQI 0/1				
SATISFY	Woche 12	208	189	71 (37,6)
	Woche 52	208	171	92 (53,8)
CDLQI 0/1				
SATISFY	Woche 12	246	212	90 (42,5)
	Woche 52	246	177	96 (54,2)

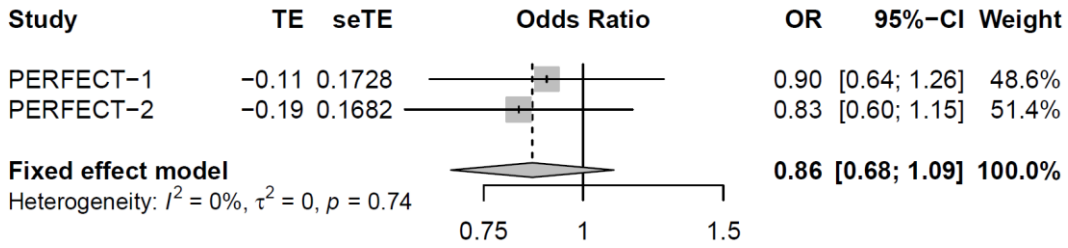
CDLQI: *Children´s Dermatology Life Quality Index*; DLQI: *Dermatology Life Quality Index*; N: Patienten in der SAF-Population mit Baseline-Wert (≤ 16 / > 16 Jahre); N′: Patienten mit Wert in Woche 12 bzw. 52

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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-10 sind die Meta-Analysen zum Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität ($p > 0,05$), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen hinsichtlich des Anteils der Patienten, die nach zwölf Wochen Behandlung einen DLQI 0/1 erreichen (RR [95 %-KI]: 0,93 [0,81; 1,06]). Das Ergebnis der Meta-Analyse bestätigt somit die Einzelergebnisse der Studien.

a)



b)

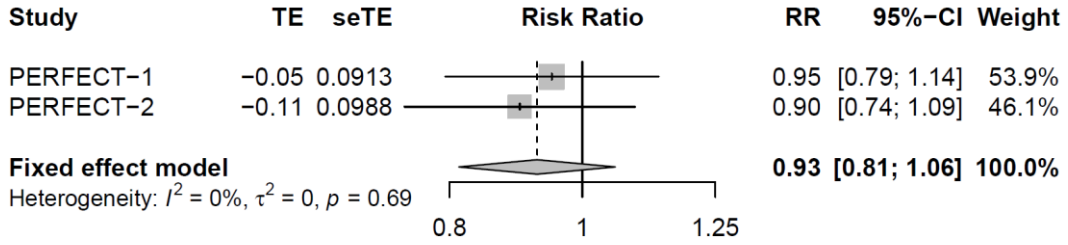


Abbildung 4-10: Meta-Analyse für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

In Abbildung 4-11 sind die Meta-Analysen zum Endpunkt gesundheitsbezogene Lebensqualität gemäß CDLQI dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich ein statistisch signifikanter Nachteil für Trifaroten gegenüber Vehikel hinsichtlich des Anteils der Patienten, die nach zwölf Wochen Behandlung einen CDLQI 0/1 aufweisen (RR [95 %-KI]: 0,83 [0,71; 0,97]).

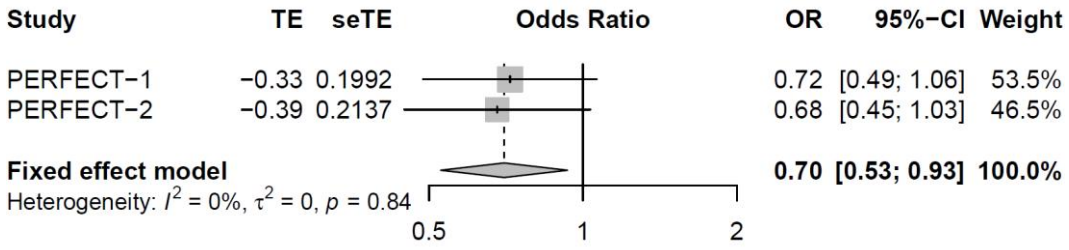
Bei der Interpretation dieses Ergebnisses ist zu beachten, dass die mittlere Beeinträchtigung der Lebensqualität bei den jüngeren Studienteilnehmern zu Studienbeginn gering ist. Somit ist keine starke Verbesserung infolge der Behandlung zu erwarten. Bereits zu Studienbeginn schätzen etwa drei Viertel der unter 17-Jährigen ihre Lebensqualität als gar nicht oder nur gering beeinträchtigt ein (siehe ergänzende Daten in Anhang 4-G). Dies kann zum einen daran liegen, dass Kinder und Jugendliche die Belastung durch Akne anders wahrnehmen als Erwachsene. Sie sind sich nicht bewusst, dass es eine chronische Erkrankung der Haut mit dem Risiko der Entstehung bleibender Narben ist, die man medikamentös behandeln kann und sollte. Die Akne wird in diesem Lebensalter sowohl von betroffenen jungen Patienten als auch von Personen im sozialen Umfeld weitestgehend als assoziierte, temporäre Veränderung oder Unreinheit der Haut wahrgenommen und akzeptiert. Zum anderen ist die Akne am Rumpf bei Kindern und Jugendlichen im frühen Alter wenig ausgeprägt und entwickelt sich mit dem Älterwerden. Patienten, bei denen die Akne auf das Gesicht beschränkt ist, zeigen eine

geringere Beeinträchtigung der gesundheitsbezogenen Lebensqualität als Patienten mit großflächiger Akne im Gesicht und am Rumpf [48-50].

Die Behandlung mit einer aktiven Substanz wie Trifaroten ist im Gegensatz zur Behandlung mit Vehikel mit transienten und überwiegend milden oder moderaten Nebenwirkungen der Haut verbunden (siehe Abschnitte 4.3.1.3.7 und 4.3.1.3.8). Es ist anzunehmen, dass bei Patienten, die ihre Akne nicht als behandlungswürdige Hautveränderung wahrnehmen, das temporäre Auftreten von Nebenwirkungen die Bewertung der gesundheitsbezogenen Lebensqualität stärker beeinflusst als das Empfinden der Verbesserung einer chronischen Erkrankung. Daraus könnte der beobachtete Nachteil für Trifaroten in der Subpopulation der unter 17-Jährigen resultieren.

Zusammenfassend zeigt sich in allen Behandlungsarmen unabhängig von der verwendeten DLQI-Version eine Verbesserung der gesundheitsbezogenen Lebensqualität im Vergleich zum Studienbeginn. Nach Ende der zwölfwöchigen Behandlungsphase schätzen etwa 80 % der Patienten ab 17 Jahren und etwa 90 % der unter 17-Jährigen ihre Lebensqualität als gar nicht oder nur gering beeinträchtigt ein (siehe ergänzende Daten in Anhang 4-G).

a)



b)

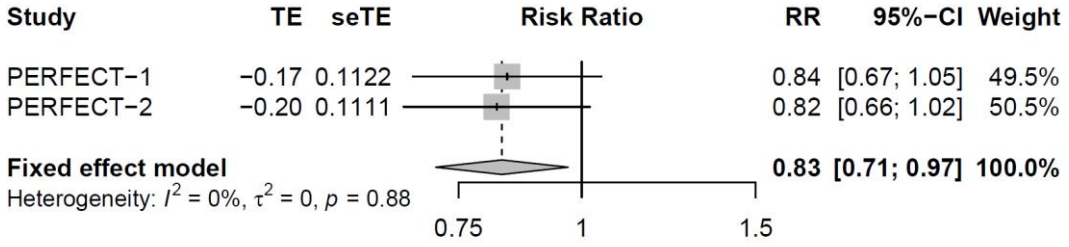


Abbildung 4-11: Meta-Analyse für den Endpunkt gesundheitsbezogene Lebensqualität gemäß CDLQI aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

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.

4.3.1.3.7 Sicherheit – RCT

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

Tabelle 4-39: Operationalisierung von Sicherheit

Studie	Operationalisierung
PERFECT-1	Unerwünschte Ereignisse (UE) wurden während des gesamten Studienzeitraums dokumentiert. Die Kodierung der UE erfolgte nach MedDRA-Version 18.0. • Ein UE war definiert als jegliches neu auftretendes unerwünschtes medizinisches Ereignis oder als Verschlechterung einer bestehenden Vorerkrankung unter Studienmedikation, unabhängig von einem möglichen Kausalzusammenhang mit der Studienmedikation. • Ein schwerwiegendes UE (SUE) war definiert als UE, das - zum Tod führte - lebensbedrohlich war (definiert durch das Risiko, dass das Ereignis zum Zeitpunkt seines Eintretens zum Tod führen konnte; die Definition schließt nicht das hypothetische Risiko ein, dass ein schwerwiegenderer Verlauf des Ereignisses zum Tode hätte führen können) - eine Krankenhausaufnahme erforderlich machte oder eine Verlängerung eines Krankenhausaufenthalts verursachte (bestimmte Ausnahmen waren im Protokoll definiert) - in anhaltender oder signifikanter Behinderung oder Invalidität resultierte - eine kongenitale Anomalie/Geburtsfehler war - ein wesentliches unerwünschtes medizinisches Ereignis war (definiert als medizinisches Ereignis, das nicht unmittelbar lebensbedrohlich war oder zum Tod oder einer Krankenhauseinweisung geführt hat und das, gestützt auf medizinische und wissenschaftliche Beurteilung, den Studienteilnehmer gefährdet hat oder eine medikamentöse oder chirurgische Intervention erforderte, um eine der oben aufgeführten schwerwiegenden Ereignisse zu vermeiden) • Ein UE von besonderem Interesse war definiert als - klinisch relevanter Laborwert außerhalb des Normbereichs, der als mit der Studienmedikation in Zusammenhang stehend erachtet wurde - kutanes UE, das als mit der Studienmedikation in Zusammenhang stehend erachtet wurde und zum Therapieabbruch führte - allergische Reaktion, die als potentiell mit der Studienmedikation in Zusammenhang stehend erachtet wurde Der Schweregrad der UE wurde basierend auf der klinischen Beurteilung durch den Prüfarzt in mild, moderat und schwer unterteilt. Die Endpunkte zur Sicherheit wurden in der Studie deskriptiv unter Einbeziehung aller Patienten, die mindestens eine Dosis entsprechend der randomisiert zugeordneten Behandlung erhielten (SAF-Population), ausgewertet. Post hoc wurden folgende Auswertungen vorgenommen: • Anteil der Patienten mit jeglichen UE • Anteil der Patienten mit SUE • Anteil der Patienten mit Therapieabbrüchen wegen UE Zusätzlich werden die Inzidenzen der UE von besonderem Interesse gemäß der Definition im Studienprotokoll sowie die Inzidenzen für häufige UE und SUE (mindestens 10 Ereignisse in einem Behandlungsarm) und Therapieabbrüche wegen UE auf Ebene der SOC und PT dargestellt.

PERFECT-2	Siehe PERFECT-1
SATISFY	Siehe PERFECT-1
PT: *Preferred Term*; SAF: *Safety*; SOC: Systemorganklasse (*system organ class*); SUE: Schwerwiegendes Unerwünschtes Ereignis; UE: Unerwünschtes Ereignis	

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 4-40: Bewertung des Verzerrungspotenzials für Endpunkte der Sicherheit in RCT mit dem zu bewertenden Arzneimittel

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial Endpunkt**
PERFECT-1						
Jegliche UE	niedrig	ja	ja	ja	ja	niedrig
SUE	niedrig	ja	ja	ja	ja	niedrig
Therapieabbrüche wegen UE	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2						
Jegliche UE	niedrig	ja	ja	ja	ja	niedrig
SUE	niedrig	ja	ja	ja	ja	niedrig
Therapieabbrüche wegen UE	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Die Sicherheitsendpunkte wurden verblindet erhoben und auf Basis der SAF-Population adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für die Endpunkte zur Sicherheit ist somit als niedrig einzustufen.

Das Verzerrungspotential der ergänzend dargestellten Studie SATISFY ist bereits auf Studienebene als hoch einzustufen, somit erfolgt keine Bewertung des Verzerrungspotentials auf Endpunktebene.

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.

In der nachfolgenden Tabelle 4-41 sind die Ergebnisse bezüglich der Gesamtraten jeglicher UE, SUE und Therapieabbrüchen wegen UE aus den Studien PERFECT-1 und -2 zusammengefasst.

In der Studie PERFECT-1 wurde bei 20,8 % der Patienten mit Vehikel und 33,9 % der Patienten mit Trifaroten ein unerwünschtes Ereignis innerhalb der zwölfwöchigen Behandlungsphase dokumentiert. Der Unterschied zwischen den Behandlungsgruppen ist statistisch signifikant (RR [95 %-KI]: 1,63 [1,34; 1,97] p < 0,001). In der Studie PERFECT-2 hingegen traten UE in beiden Behandlungsarmen mit 19,2 % unter Vehikel und 20,2 % unter Trifaroten etwa gleich häufig auf. Der Unterschied zwischen den Behandlungsgruppen ist statistisch nicht signifikant (RR [95 %-KI]: 1,05 [0,84; 1,32], p = 0,665).

SUE waren sowohl unter Vehikel als auch unter Trifaroten sehr selten, und keines der schwerwiegenden Ereignisse wurde als mit der Studienmedikation in Zusammenhang stehend eingestuft. In beiden Studien zeigt sich kein statistisch signifikanter Unterschied zwischen der Behandlung mit Trifaroten und Vehikel (RR [95 %-KI]: 1,92 [0,35; 10,42], p = 0,687 bzw. 0,50 [0,09; 2,75]; p = 0,687).

Todesfälle traten in beiden Studien nicht auf.

Therapieabbrüche aufgrund unerwünschter Ereignisse waren ebenfalls selten und traten unter Trifaroten bei 1,7 % bzw. 2,3 % der Patienten auf, wohingegen im Vehikel-Arm in beiden Studien 0,2 % der Patienten die Behandlung aufgrund von UE abbrachen. Der Unterschied zwischen den Behandlungsarmen ist in beiden Studien statistisch signifikant (RR [95 %-KI]: 13,41 [1,77; 101,66], p < 0,001 bzw. 10,10 [1,30; 78,65], p = 0,006).

Tabelle 4-41: Ergebnisse zu unerwünschten Ereignissen aus RCT (12 Wochen, PERFECT-1 und -2)

Studie	Trifaroten		Vehikel		Trifaroten vs. Vehikel		
	N	Patienten mit Ereignis n (%)	N	Patienten mit Ereignis n (%)	OR[1] [95 %-KI]	RR[1] [95 %-KI]	ARR[2] [95 %-KI]
Jegliche UE							
PERFECT-1	617	209 (33,9)	591	123 (20,8)	1,95 [1,50; 2,53]	1,63 [1,34; 1,97]	13,06 [8,10; 18,03]
						p < 0,001[3]	
PERFECT-2	603	122 (20,2)	609	117 (19,2)	1,07 [0,80; 1,42]	1,05 [0,84; 1,32]	1,02 [-3,46; 5,50]
						p = 0,665[3]	
SUE							
PERFECT-1	617	4 (0,6)	591	2 (0,3)	1,92 [0,35; 10,53]	1,92 [0,35; 10,42]	0,31 [-0,48; 1,10]
						p = 0,687[3]	
PERFECT-2	603	2 (0,3)	609	4 (0,7)	0,50 [0,09; 2,76]	0,50 [0,09; 2,75]	-0,33 [-1,11; 0,46]
						p = 0,687[3]	
Therapieabbrüche wegen UE							
PERFECT-1	617	14 (2,3)	591	1 (0,2)	13,70 [1,80; 104,50]	13,41 [1,77; 101,66]	2,10 [0,88; 3,32]
						p < 0,001[3]	
PERFECT-2	603	10 (1,7)	609	1 (0,2)	10,25 [1,31; 80,34]	10,10 [1,30; 78,65]	1,49 [0,43; 2,56]
						p = 0,006[3]	

(1) Odds Ratios, Relative Risiken und die zugehörigen Konfidenzintervalle basieren auf einer Normalverteilungsapproximation für log-transformierte Daten ohne Anwendung einer Nullzellenkorrektur.

(2) Risikodifferenzen und die zugehörigen Konfidenzintervalle basieren auf einer Normalverteilungsapproximation ohne Anwendung einer Nullzellenkorrektur.

(3) P-Werte basieren auf einem exakten Test nach Fisher.

ARR: Absolute Risikoreduktion; N: Patienten in der SAF-Population; n: Patienten mit Ereignis; KI: Konfidenzintervall; OR: Odds Ratio; RR: Relatives Risiko; SAF: *Safety*; SUE: schwerwiegende unerwünschte Ereignisse; UE: unerwünschte Ereignisse

In Tabelle 4-42 sind die häufigsten UE für die Studien PERFECT-1 und -2 mit den jeweiligen Systemorganklassen (SOC, *system organ class*) und bevorzugten Bezeichnungen (PT, *preferred term*) aufgelistet.

In der Studie PERFECT-1 wurden am häufigsten unerwünschte Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“, dabei insbesondere die PT „Reizung an der Applikationsstelle“ und „Pruritus an der Applikationsstelle“, dokumentiert. Diese Ereignisse traten unter Trifaroten deutlich häufiger auf als unter Vehikel. In der Studie PERFECT-2 waren Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ bei insgesamt niedrigerer Inzidenz als in der Studie PERFECT-1 unter Trifaroten ebenfalls häufiger als unter Vehikel.

Weitere häufige UE waren in beiden Studien Ereignisse der SOC „Infektionen und parasitäre Erkrankungen“ mit etwa gleicher Inzidenz in den Behandlungsarmen sowie Ereignisse der SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“, insbesondere „Sonnenbrand“, welche unter Trifaroten häufiger als unter Vehikel dokumentiert wurden.

Tabelle 4-42: Ergebnisse zu häufigen unerwünschten Ereignissen aus RCT (≥ 10 Ereignisse in einem Behandlungsarm, PERFECT-1 und -2)

Studie	**PERFECT-1**		**PERFECT-2**	
SOC / PT	**Trifaroten (N = 617) n (%)**	**Vehikel (N = 591) n (%)**	**Trifaroten (N = 603) n (%)**	**Vehikel (N = 609) n (%)**
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	92 (14,9)	15 (2,5)	29 (4,8)	4 (0,7)
PT: Reizung an der Applikationsstelle	66 (10,7)	4 (0,7)	18 (3,0)	0
PT: Pruritus an der Applikationsstelle	24 (3,9)	8 (1,4)	5 (0,8)	2 (0,3)
SOC: Infektionen und parasitäre Erkrankungen	70 (11,3)	63 (10,7)	59 (9,3)	73 (12,0)
PT: Nasopharyngitis	24 (3,9)	27 (4,6)	26 (4,3)	29 (4,8)
PT: Infektion der oberen Atemwege	10 (1,6)	8 (1,4)	9 (1,5)	8 (1,3)
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	50 (8,1)	15 (2,5)	13 (2,2)	4 (0,7)
PT: Sonnenbrand	27 (4,4)	5 (0,8)	6 (1,0)	1 (0,2)
SOC: Erkrankungen der Haut und des Unterhautgewebes	24 (3,9)	6 (1,0)	6 (1,0)	5 (0,8)
SOC: Erkrankungen des Nervensystems	12 (1,9)	12 (2,0)	12 (2,0)	4 (0,7)
PT: Kopfschmerzen	6 (1,0)	12 (2,0)	10 (1,7)	4 (0,7)
SOC: Erkrankungen des Gastrointestinaltrakts	12 (1,9)	8 (1,4)	12 (2,0)	14 (2,3)
SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums	7 (1,1)	14 (2,4)	2 (0,3)	6 (1,0)

N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SOC: *System Organ Class*; PT: *Preferred Term*; UE: unerwünschte Ereignisse

Bei den in den Studien PERFECT-1 und -2 aufgetretenen SUE handelt es sich um sehr wenige Einzelereignisse, die sämtlich als nicht mit der Studienmedikation in Zusammenhang stehend eingestuft wurden, so dass auf eine Auflistung der Ereignisse auf Ebene der SOC und PT verzichtet wird.

Die Gründe für Therapieabbrüche aufgrund von unerwünschten Ereignissen sind in Tabelle 4-43 aufgelistet. In beiden Studien führten am häufigsten Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ und insbesondere das Ereignis „Reizung an der Applikationsstelle“ zu einem Abbruch der Behandlung mit Trifaroten.

Tabelle 4-43: Ergebnisse zu Therapieabbrüchen wegen unerwünschter Ereignisse aus RCT (PERFECT-1 und -2)

Studie	**PERFECT-1**		**PERFECT-2**	
SOC / PT	**Trifaroten (N = 617) n (%)**	**Vehikel (N = 591) n (%)**	**Trifaroten (N = 603) n (%)**	**Vehikel (N = 609) n (%)**
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	8 (1,3)	0	7 (1,2)	0
PT: Reizung an der Applikationsstelle	8 (1,3)	0	5 (0,8)	0
PT: Erosion an der Verabreichungsstelle	0	0	1 (0,2)	0
PT: Erythem am Verabreichungsort	0	0	1 (0,2)	0
PT: Schmerzen an der Verabreichungsstelle	0	0	1 (0,2)	0
SOC: Infektionen und parasitäre Erkrankungen	0	1 (0,2)	1 (0,2)	1 (0,2)
PT: Bronchitis	0	1 (0,2)	1 (0,2)	0
PT: Sinusitis	0	0	0	1 (0,2)
SOC: Erkrankungen des Immunsystems	0	0	1 (0,2)	0
PT: Arzneimittelüberempfindlichkeit	0	0	1 (0,2)	0
SOC: Psychiatrische Erkrankungen	0	0	1 (0,2)	0
PT: Endogene Depression	0	0	1 (0,2)	0
PT: Suizidversuch	0	0	1 (0,2)	0
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	1 (0,2)	0	0	0
PT: Muskelriss	1 (0,2)	0	0	0
PT: Tibiafraktur	1 (0,2)	0	0	0

Studie	**PERFECT-1**		**PERFECT-2**	
SOC / PT	**Trifaroten (N = 617) n (%)**	**Vehikel (N = 591) n (%)**	**Trifaroten (N = 603) n (%)**	**Vehikel (N = 609) n (%)**
SOC: Erkrankungen der Haut und des Unterhautgewebes	5 (0,8)	0	0	0
PT: Akne	2 (0,3)	0	0	0
PT: Hautreizung	2 (0,3)	0	0	0
PT: Dermatitis allergisch	1 (0,2)	0	0	0

N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SOC: *System Organ Class*; PT: *Preferred Term*; UE: unerwünschte Ereignisse

In Tabelle 4-44 sind die Ergebnisse zu den UE von besonderem Interesse, die in den Studien PERFECT-1 und -2 prädefiniert waren, zusammengefasst. In beiden Studien sind die Mehrheit der UE von besonderem Interesse kutane Ereignisse, die als mit der Studienmedikation in Zusammenhang stehend erachtet wurden und zum Therapieabbruch führten; diese Ereignisse traten ausschließlich unter Trifaroten auf.

Tabelle 4-44: Ergebnisse zu UE von besonderem Interesse aus RCT (PERFECT-1 und -2)

Studie	**PERFECT-1**		**PERFECT-2**	
SOC / PT	**Trifaroten (N = 617) n (%)**	**Vehikel (N = 591) n (%)**	**Trifaroten (N = 603) n (%)**	**Vehikel (N = 609) n (%)**
Patienten mit UE von besonderem Interesse (Gesamtrate)	16 (2,6)	2 (0,3)	9 (1,5)	2 (0,3)
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	9 (1,5)	0	8 (1,3)	0
PT: Reizung an der Applikationsstelle	9 (1,5)	0	5 (0,8)	0
PT: Trockenheit an der Verabreichungsstelle	0	0	1 (0,2)	0
PT: Erythem am Verabreichungsort	0	0	1 (0,2)	0
PT: Erosion an der Verabreichungsstelle	0	0	1 (0,2)	0
PT: Schmerzen an der Verabreichungsstelle	0	0	1 (0,2)	0
SOC: Erkrankungen der Haut und des Unterhautgewebes	7 (1.1)	0	1 (0,2)	0
PT: Dermatitis allergisch	3 (0,5)	0	0	0
PT: Akne	2 (0,3)	0	0	0

Studie	PERFECT-1		PERFECT-2	
SOC / PT	Trifaroten (N = 617) n (%)	Vehikel (N = 591) n (%)	Trifaroten (N = 603) n (%)	Vehikel (N = 609) n (%)
PT: Hautreizung	2 (0,3)	0	1 (0,2)	0
SOC: Untersuchungen	0	2 (0,3)	0	2 (0,3)
PT: Kreatinin im Blut erhöht	0	2 (0,3)	0	0
PT: Leberfunktionstest anormal	0	2 (0,3)	0	0
PT: Bilirubin im Blut erhöht	0	0	0	1 (0,2)
PT: Kreatinin im Blut erhöht	0	0	0	1 (0,2)
SOC: Stoffwechsel- und Ernährungsstörungen	0	1 (0,2)	0	0
PT: Hyperurikämie	0	1 (0,2)	0	0

N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SOC: *System Organ Class*; PT: *Preferred Term*; UE: unerwünschte Ereignisse

In Tabelle 4-45 sind die Gesamtraten für jegliche UE, SUE und Therapieabbrüche wegen UE aus der Studie SATISFY für das erste Quartal der Studie und den gesamten Studienzeitraum dargestellt. Insgesamt trat bei knapp der Hälfte der Patienten im Studienverlauf ein UE auf, wobei die Mehrheit dieser Ereignisse im 1. Studienquartal dokumentiert wurde. Die Gesamtraten der SUE und Therapieabbrüche wegen UE waren in den ersten zwölf Wochen vergleichbar niedrig wie in den Studien PERFECT-1 und -2. Die Gesamtraten steigen bis Woche 52 etwas an. Insbesondere bei den Therapieabbrüchen wegen UE ist zu beobachten, dass fast alle Ereignisse, die zu einem Therapieabbruch führen, in den ersten drei Monaten der Behandlung dokumentiert wurden.

Tabelle 4-45: Ergänzende Ergebnisse zu unerwünschten Ereignissen (12 und 52 Wochen, SATISFY)

Studie	Endpunkt	Anteil Patienten mit Ereignis im 1. Studienquartal[(1)] (N = 453) n (%)	Anteil Patienten mit Ereignis im gesamten Studienzeitraum bis Woche 52 (N = 453) n (%)
SATISFY	Jegliche UE	154 (34,0)	218 (48,1)
	SUE	4 (0,9)	10 (2,2)
	Therapieabbrüche wegen UE	13 (2,9)	16 (3,5)

(1) Ereignisse mit Beginn von Tag 1 bis Tag 89

Studie	Endpunkt	Anteil Patienten mit Ereignis im 1. Studienquartal[1] (N = 453) n (%)	Anteil Patienten mit Ereignis im gesamten Studienzeitraum bis Woche 52 (N = 453) n (%)
N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SUE: schwerwiegende unerwünschte Ereignisse; UE: unerwünschte Ereignisse			

Die häufigsten UE in der Studie SATISFY stellen wie auch in den Studien PERFECT-1 und -2 Ereignisse der SOC „Infektionen und parasitäre Erkrankungen“, „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“, dabei insbesondere die PT „Reizung an der Applikationsstelle“ und „Pruritus an der Applikationsstelle“, sowie Ereignisse der SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ dar (siehe Tabelle 4-46).

Tabelle 4-46: Ergänzende Ergebnisse zu häufigen unerwünschten Ereignissen (≥ 10 Ereignisse, 52 Wochen, SATISFY)

SOC / PT	Trifaroten (N = 453) n (%)
SOC: Infektionen und parasitäre Erkrankungen	116 (25,6)
PT: Nasopharyngitis	48 (10,6)
PT: Infektion der oberen Atemwege	13 (2,9)
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	57 (12,6)
PT: Pruritus an der Applikationsstelle	23 (5,1)
PT: Reizung an der Applikationsstelle	22 (4,9)
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	50 (11,0)
PT: Sonnenbrand	27 (6,0)
SOC: Erkrankungen der Haut und des Unterhautgewebes	25 (5,5)
SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums	23 (5,1)
SOC: Erkrankungen des Gastrointestinaltrakts	19 (4,2)
SOC: Erkrankungen des Nervensystems	10 (2,2)
N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SOC: *System Organ Class*; PT: *Preferred Term*	

Der häufigste Grund für den Abbruch der Therapie waren Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ und insbesondere das Ereignis „Reizung an der Applikationsstelle“ (siehe Tabelle 4-47).

Tabelle 4-47: Ergänzende Ergebnisse zu Therapieabbrüchen wegen unerwünschter Ereignisse (52 Wochen, SATISFY)

SOC / PT	**Trifaroten (N = 453) n (%)**
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	10 (2,2)
PT: Reizung an der Applikationsstelle	10 (2,2)
SOC: Erkrankungen der Haut und des Unterhautgewebes	5 (1,1)
PT: Akne	5 (1,1)
SOC: Erkrankungen der Geschlechtsorgane und der Brustdrüse	1 (0,2)
PT: Polyzystische Eierstöcke	1 (0,2)
N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SOC: *System Organ Class*; PT: *Preferred Term*	

Als UE von besonderem Interesse wurden ausschließlich kutane Ereignisse, die als mit der Studienmedikation in Zusammenhang stehend erachtet wurden und zum Therapieabbruch führten, dokumentiert (siehe Tabelle 4-48).

Tabelle 4-48: Ergänzende Ergebnisse zu UE von besonderem Interesse (52 Wochen, SATISFY)

SOC / PT	**Trifaroten (N = 453) n (%)**
Patienten mit UE von besonderem Interesse (Gesamtrate)	13 (2,9)
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	10 (2,2)
PT: Reizung an der Applikationsstelle	10 (2,2)
SOC: Erkrankungen der Haut und des Unterhautgewebes	3 (0,7)
PT: Akne	3 (0,7)
N: Patienten in der SAF-Population; n: Patienten mit Ereignis; SAF: *Safety*; SOC: *System Organ Class*; PT: *Preferred Term*	

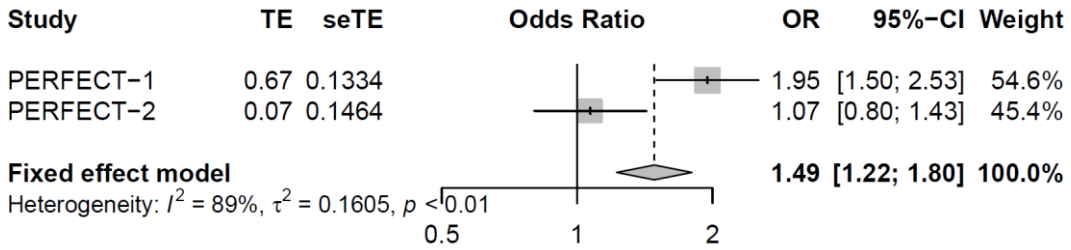
Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ

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.

Die Ergebnisse aus den Studien PERFECT-1 und -2 sind aus medizinischer und methodischer Sicht grundsätzlich für eine Meta-Analyse geeignet (siehe Abschnitt 4.2.5.3).

In Abbildung 4-12 sind die Meta-Analysen zum Endpunkt jegliche UE dargestellt. Anhand der Heterogenitätstests zeigt sich für beide Effektschätzer eine bedeutsame Heterogenität ($p < 0,01$) und es liegen keine deutlich gleichgerichteten Effekte in den beiden Studien vor. Somit ist eine meta-analytische Zusammenfassung nicht sinnvoll und zur Auswertung des Endpunkts jegliche UE sollten die Einzelergebnisse der Studien betrachtet werden (siehe Tabelle 4-41). Die Ursache für die Heterogenität kann gegebenenfalls auf methodische oder klinische Unterschiede zwischen den Studien zurückzuführen sein, oder es liegt eine zufallsbedingte Variabilität vor. Da die Studien ein identisches Studiendesign aufweisen, ist eine methodische Heterogenität eher unwahrscheinlich. Die bekannten Patientencharakteristika sind hinreichend ähnlich, allerdings gibt es regionale Unterschiede zwischen den Studien. Ob diese die Ursache für die festgestellte Heterogenität darstellen oder ob es sich um zufallsbedingte Unterschiede handelt, ist unklar.

a)



b)

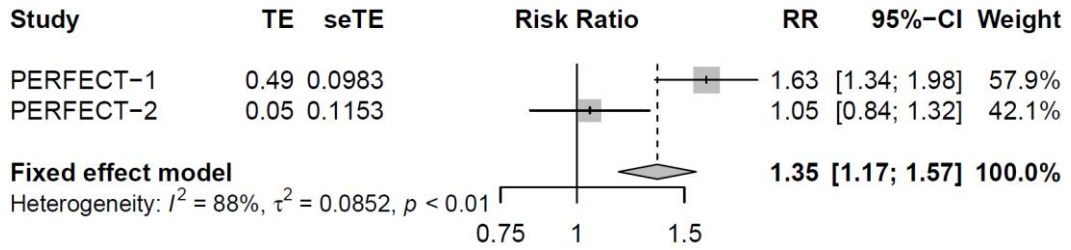
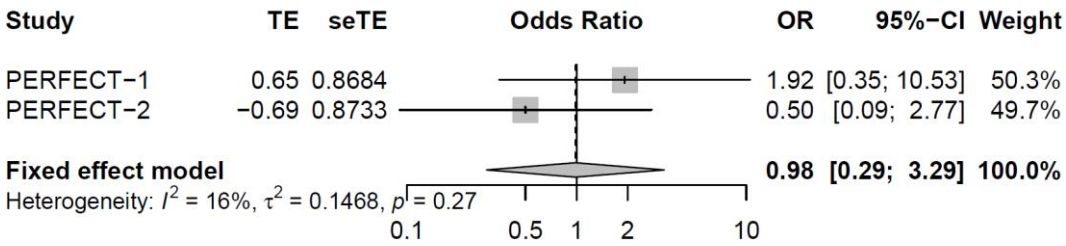


Abbildung 4-12: Meta-Analyse für den Endpunkt jegliche UE aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

In Abbildung 4-13 sind die Meta-Analysen zum Endpunkt SUE dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen (RR [95 %-KI]: 0,98 [0,29; 3,28]). Die Meta-Analyse bestätigt somit die Einzelergebnisse der beiden Studien hinsichtlich der Häufigkeit von SUE.

a)



b)

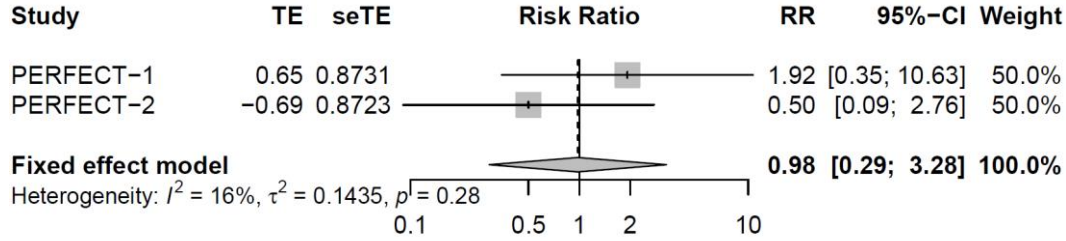
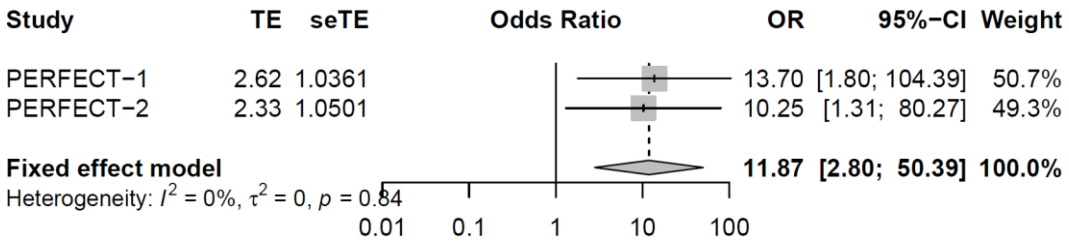


Abbildung 4-13: Meta-Analyse für den Endpunkt SUE aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

In Abbildung 4-14 sind die Meta-Analysen zum Endpunkt Therapieabbrüche wegen UE dargestellt. Anhand der Heterogenitätstests zeigt sich für die untersuchten Effektschätzer keine bedeutsame Heterogenität (p > 0,05), sodass eine meta-analytische Zusammenfassung sinnvoll ist. In der Meta-Analyse auf Basis eines Modells mit festen Effekten ergibt sich ein statistisch signifikanter Nachteil für Trifaroten gegenüber Vehikel hinsichtlich des Anteils der Patienten, die aufgrund von UE die Behandlung abbrechen (RR [95 %-KI]: 11,66 [2,76; 49,27]). Die Meta-Analyse bestätigt somit die Einzelergebnisse der beiden Studien hinsichtlich der Häufigkeit von Therapieabbrüchen wegen UE.

Die Ergebnisse der Studien PERFECT-1 und -2 sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1), dies gilt folglich auch für die Ergebnisse der Meta-Analyse.

a)



b)

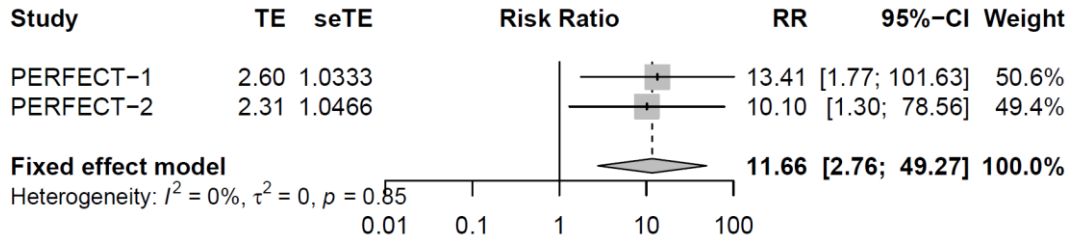


Abbildung 4-14: Meta-Analyse für den Endpunkt Therapieabbrüche wegen UE aus RCT; Trifaroten vs. Vehikel
a) Odds Ratio, b) Relatives Risiko

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.

4.3.1.3.8 Lokale Verträglichkeit – RCT

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 lokaler Verträglichkeit

<table>
<tr><th>Studie</th><th>Operationalisierung</th></tr>
<tr><td>PERFECT-1</td><td>Die lokale Verträglichkeit der Behandlung wurde separat für das Gesicht und den Rumpf ausgewertet. Für beide Hautareale wurden die Symptome Rötung (erythema), Schuppung (scaling), Trockenheit (dryness) und Stechen/Brennen (stinging/burning) jeweils auf einer vierstufigen Skala nach folgenden Kriterien erhoben:
<table>
<tr><td colspan="3">Rötung: abnormale Rötung der Haut</td></tr>
<tr><td>Keine</td><td>0</td><td>Keine Rötung</td></tr>
<tr><td>Mild</td><td>1</td><td>Leichte Rosafärbung vorhanden</td></tr>
<tr><td>Moderat</td><td>2</td><td>Deutliche Rötung, leicht zu erkennen</td></tr>
<tr><td>Schwer</td><td>3</td><td>Intensive Rötung</td></tr>
<tr><td colspan="3">Schuppung: abnormale Schuppung des Stratum corneum</td></tr>
<tr><td>Keine</td><td>0</td><td>Keine Schuppung</td></tr>
<tr><td>Mild</td><td>1</td><td>Kaum wahrnehmbare Schuppung, nur bei leichtem Kratzen oder Reiben spürbar</td></tr>
<tr><td>Moderat</td><td>2</td><td>Deutliche aber nicht übermäßige Schuppung</td></tr>
<tr><td>Schwer</td><td>3</td><td>Starke Schuppenbildung</td></tr>
<tr><td colspan="3">Trockenheit: Sprödigkeit und/oder Spannungsgefühl</td></tr>
<tr><td>Keine</td><td>0</td><td>Keine Trockenheit</td></tr>
<tr><td>Mild</td><td>1</td><td>Leichte aber deutliche Rauheit</td></tr>
<tr><td>Moderat</td><td>2</td><td>Moderate Rauheit</td></tr>
<tr><td>Schwer</td><td>3</td><td>Ausgeprägte Rauheit</td></tr>
<tr><td colspan="3">Stechen/Brennen: kribbelndes Schmerzempfinden nach der Anwendung</td></tr>
<tr><td>Keine</td><td>0</td><td>Kein Stechen/Brennen</td></tr>
<tr><td>Mild</td><td>1</td><td>Leicht warmes, kribbelndes/stechendes Gefühl, nicht wirklich störend</td></tr>
<tr><td>Moderat</td><td>2</td><td>Deutliches warmes, kribbelndes/stechendes Gefühl, etwas störend</td></tr>
<tr><td>Schwer</td><td>3</td><td>Heißes, kribbelndes/stechendes Gefühl, deutlich unangenehm</td></tr>
</table>
Die lokale Verträglichkeit im Gesicht und am Rumpf wurde zur Baseline-Visite sowie bei jeder Post-Baseline-Visite erhoben. Die Symptome Rötung, Schuppung und Trockenheit wurden dabei durch den Prüfarzt bewertet. Das Symptom Brennen wurde vom Prüfarzt nach Angaben des Patienten bewertet.
Die Endpunkte zur lokalen Verträglichkeit wurden in der Studie unter Einbeziehung aller Patienten, die mindestens eine Dosis entsprechend der randomisiert zugeordneten Behandlung erhielten (SAF- bzw. SAFT-Population), ausgewertet. Dabei wurde jeweils der schlechteste Post-Baseline-Wert der Patienten dargestellt, wobei nur die Ereignisse berücksichtigt wurden, die eine Verschlechterung gegenüber dem Studienbeginn darstellten.</td></tr>
<tr><td>PERFECT-2</td><td>Siehe PERFECT-1</td></tr>
<tr><td>SATISFY</td><td>Siehe PERFECT-1</td></tr>
<tr><td colspan="2">SAF: Safety; SAFT: Safety on the trunk</td></tr>
</table>

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 Verzerrungspotenzials für den Endpunkt Lokale Verträglichkeit in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PERFECT-1	niedrig	ja	ja	ja	ja	niedrig
PERFECT-2	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential auf Studienebene ist für die Studien PERFECT-1 und -2 als niedrig einzuschätzen. Die Endpunkte zur lokalen Verträglichkeit wurden verblindet erhoben und auf Basis der SAF- bzw. SAFT-Population adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für die Endpunkte zur lokalen Verträglichkeit ist somit als niedrig einzustufen.

Das Verzerrungspotential der ergänzend dargestellten Studie SATISFY ist bereits auf Studienebene als hoch einzustufen, somit erfolgt keine Bewertung des Verzerrungspotentials auf Endpunktebene.

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.

Die Ergebnisse der Auswertung zur lokalen Verträglichkeit im Gesicht bzw. am Rumpf aus den Studien PERFECT-1 und -2 sind in Tabelle 4-51 bzw. Tabelle 4-52 dargestellt.

Bei allen dokumentierten Symptomen zeigt sich erwartungsgemäß häufiger eine Verschlechterung bei Anwendung von Trifaroten im Vergleich zum Vehikel. Die Ereignisse sind jedoch in den meisten Fällen von milder oder moderater Schwere.

Schwere Unverträglichkeiten im Gesicht traten bei 2,5 % bis maximal 10,0 % der Patienten unter Trifaroten auf, im Vehikel-Arm waren 0 % bis maximal 1,6 % von schweren Unverträglichkeiten betroffen.

Am Rumpf sind Unverträglichkeiten insgesamt seltener als im Gesicht. Bei 0,3 % bis maximal 7,2 % der Patienten unter Trifaroten traten schwere Unverträglichkeiten am Rumpf auf, im Vehikel-Arm waren 0 % bis maximal 0,8 % von schweren Unverträglichkeiten betroffen.

Tabelle 4-51: Ergebnisse zur lokalen Verträglichkeit im Gesicht aus RCT (schlechtester Post-Baseline-Wert über 12 Wochen; PERFECT-1 und -2)

Studie	**PERFECT-1**		**PERFECT-2**	
Symptom / Schweregrad	**Trifaroten (N = 617 / 612) n (%)**	**Vehikel (N = 591 / 587) n (%)**	**Trifaroten (N = 603 / 602) n (%)**	**Vehikel (N = 609 / 607) n (%)**
Rötung				
Mild (1)	212 (34,6)	137 (23,3)	159 (26,4)	114 (18,8)
Moderat (2)	145 (23,7)	38 (6,5)	200 (33,2)	43 (7,1)
Schwer (3)	15 (2,5)	0	60 (10,0)	10 (1,6)
Schuppung				
Mild (1)	235 (38,4)	136 (23,2)	220 (36,5)	147 (24,2)
Moderat (2)	131 (21,4)	16 (2,7)	198 (32,9)	55 (9,1)
Schwer (3)	18 (2,9)	0	41 (6,8)	4 (0,7)
Trockenheit				
Mild (1)	269 (44,0)	159 (27,1)	204 (33,9)	198 (32,6)
Moderat (2)	141 (23,0)	23 (3,9)	219 (36,4)	58 (9,6)
Schwer (3)	15 (2,5)	0	43 (7,1)	9 (1,5)
Stechen/ Brennen				
Mild (1)	227 (37,1)	85 (14,5)	205 (34,1)	105 (17,3)
Moderat (2)	100 (16,3)	10 (1,7)	150 (24,9)	35 (5,8)
Schwer (3)	26 (4,2)	1 (0,2)	46 (7,6)	5 (0,8)

N: Patienten in der SAF-Population / Anzahl Patienten in der SAF-Population mit Post-Baseline-Wert; n: Patienten mit Ereignis; SAF: *Safety*

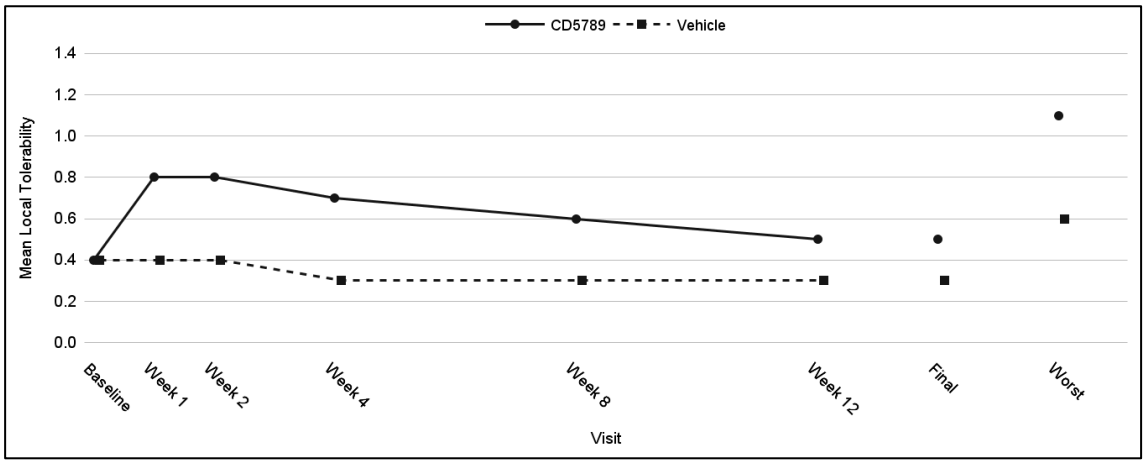
Tabelle 4-52: Ergebnisse zur lokalen Verträglichkeit am Rumpf aus RCT (schlechtester Post-Baseline-Wert über 12 Wochen; PERFECT-1 und -2)

Studie	**PERFECT-1**		**PERFECT-2**	
Symptom / Schweregrad	**Trifaroten (N = 608 / 603) n (%)**	**Vehikel (N = 583 / 579) n (%)**	**Trifaroten (N = 600 / 599) n (%)**	**Vehikel (N = 608 / 606) n (%)**
Rötung				
Mild (1)	156 (25,9)	84 (14,5)	162 (27,0)	66 (10,9)
Moderat (2)	88 (14,6)	24 (4,1)	139 (23,2)	28 (4,6)
Schwer (3)	20 (3,3)	0	43 (7,2)	5 (0,8)
Schuppung				
Mild (1)	143 (23,7)	70 (12,1)	214 (35,7)	87 (14,4)
Moderat (2)	65 (10,8)	6 (1,0)	100 (16,7)	25 (4,1)
Schwer (3)	2 (0,3)	0	18 (3,0)	1 (0,2)
Trockenheit				
Mild (1)	183 (30,3)	91 (15,7)	213 (35,6)	120 (19,8)
Moderat (2)	68 (11,3)	10 (1,7)	125 (20,9)	36 (5,9)
Schwer (3)	7 (1,2)	0	15 (2,5)	1 (0,2)
Stechen/Brennen				
Mild (1)	143 (23,7)	44 (7,6)	171 (28,5)	65 (10,7)
Moderat (2)	54 (9,0)	7 (1,2)	77 (12,9)	19 (3,1)
Schwer (3)	18 (3,0)	2 (0,3)	34 (5,7)	4 (0,7)

N: Patienten in der SAFT-Population / Anzahl Patienten in der SAFT-Population mit Post-Baseline-Wert; n: Patienten mit Ereignis; SAFT: *Safety on the trunk*

In den folgenden Abbildungen (Abbildung 4-15 bis Abbildung 4-22) ist die mittlere Änderung der lokalen Verträglichkeitsparameter im Gesicht und am Rumpf im Zeitverlauf für beide Studien dargestellt. Dabei zeigen die Kurven für Trifaroten jeweils ein Maximum der Symptome im Gesicht in den ersten beiden Wochen und eine Abnahme im weiteren Verlauf bis Woche 12. Für die Symptome am Rumpf zeigt sich hingegen das Maximum in Woche 2 bis 4. Für alle Verträglichkeitsparameter gilt jedoch unabhängig von der Lokalisation, dass die auftretenden Hautreaktionen in der überwiegenden Mehrheit der Fälle transient sind und im Laufe der Behandlung schnell abklingen.

a)



b)

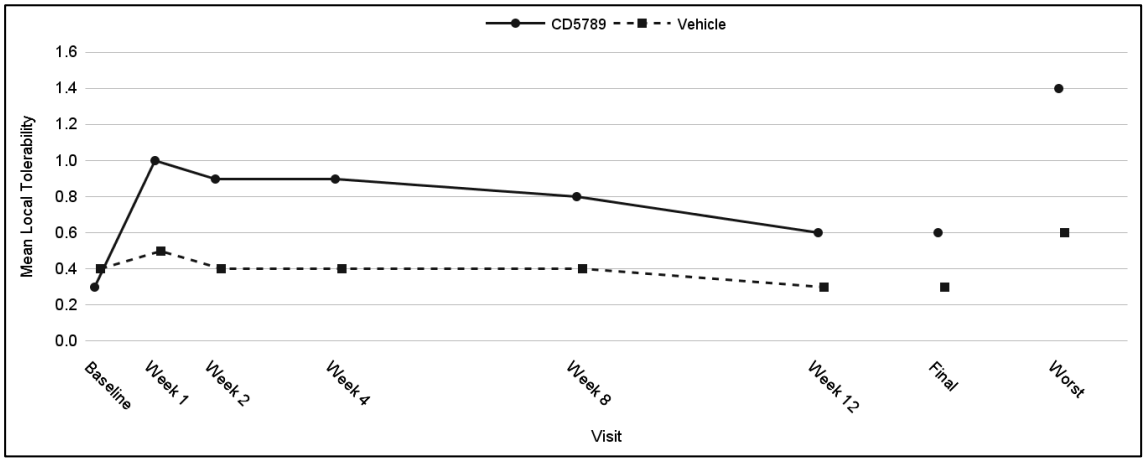
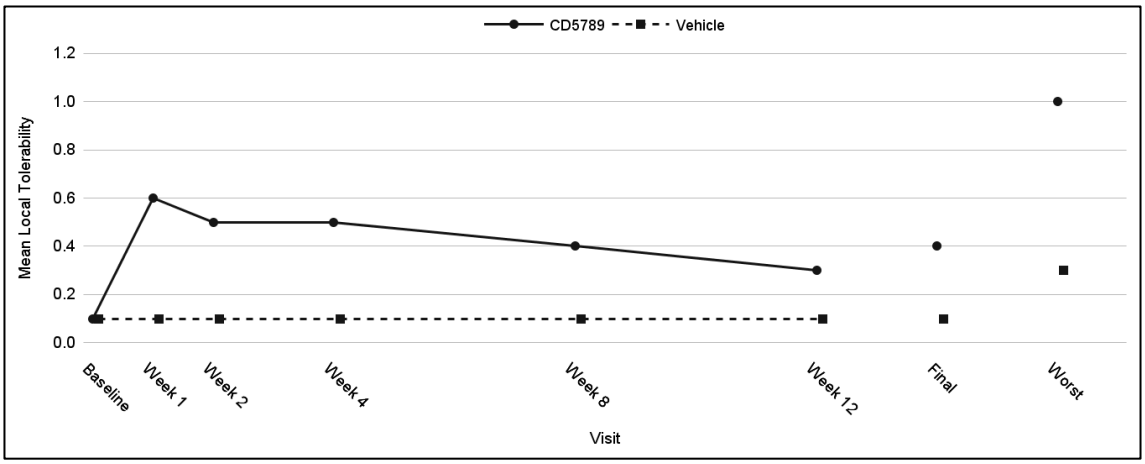


Abbildung 4-15: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Rötung
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

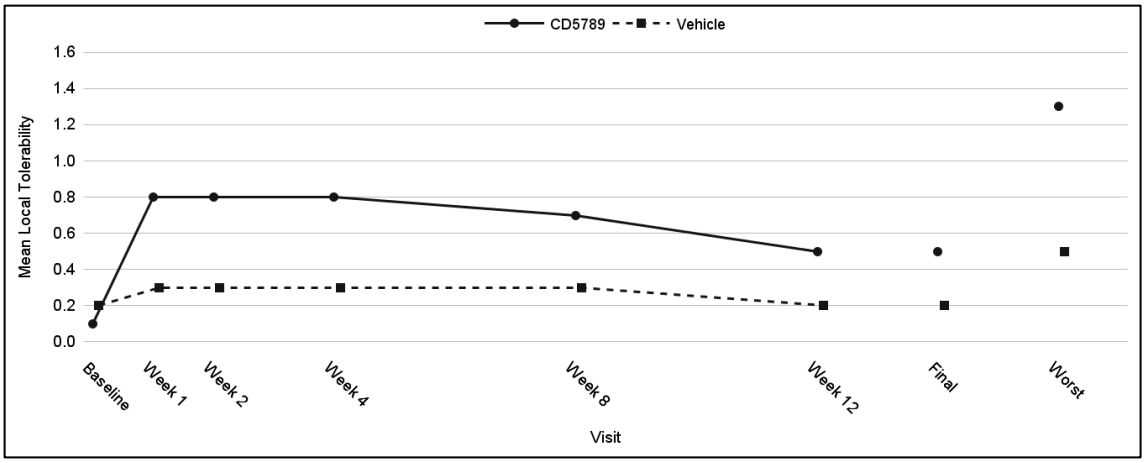
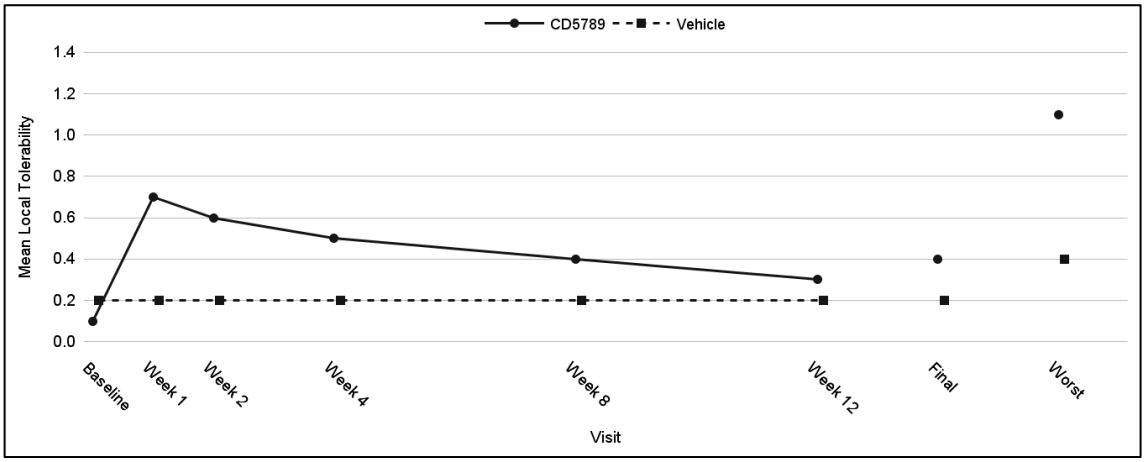


Abbildung 4-16: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Schuppung
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

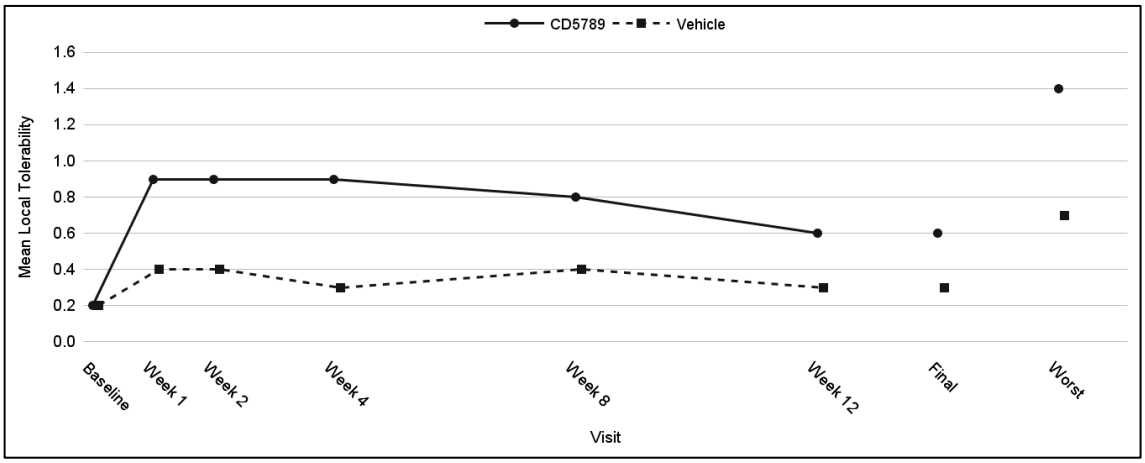
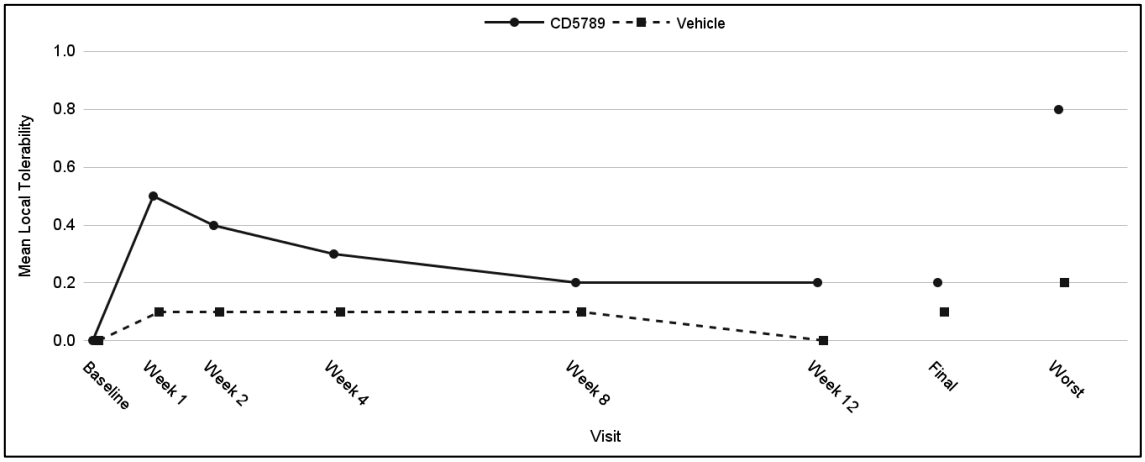


Abbildung 4-17: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Trockenheit
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

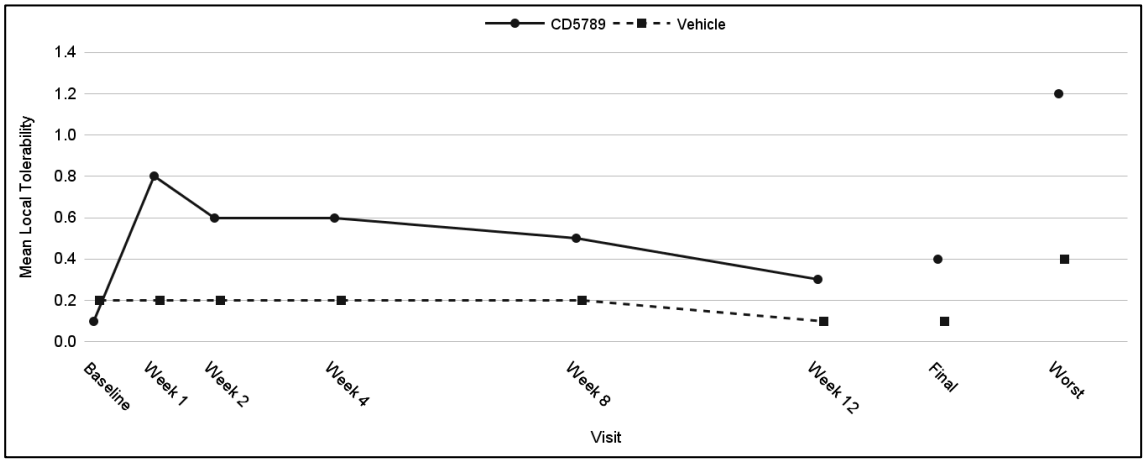
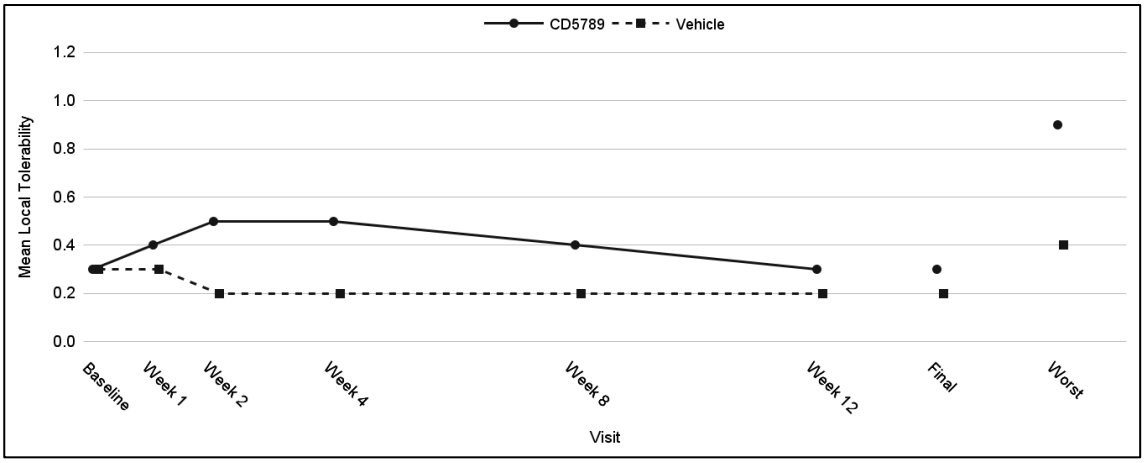


Abbildung 4-18: Lokale Verträglichkeit im Gesicht im Zeitverlauf – Stechen/Brennen
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

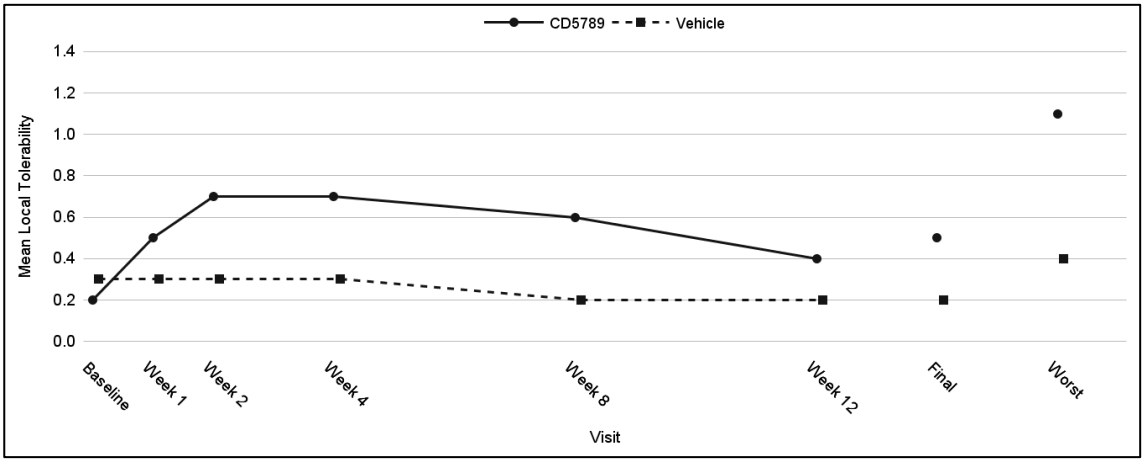
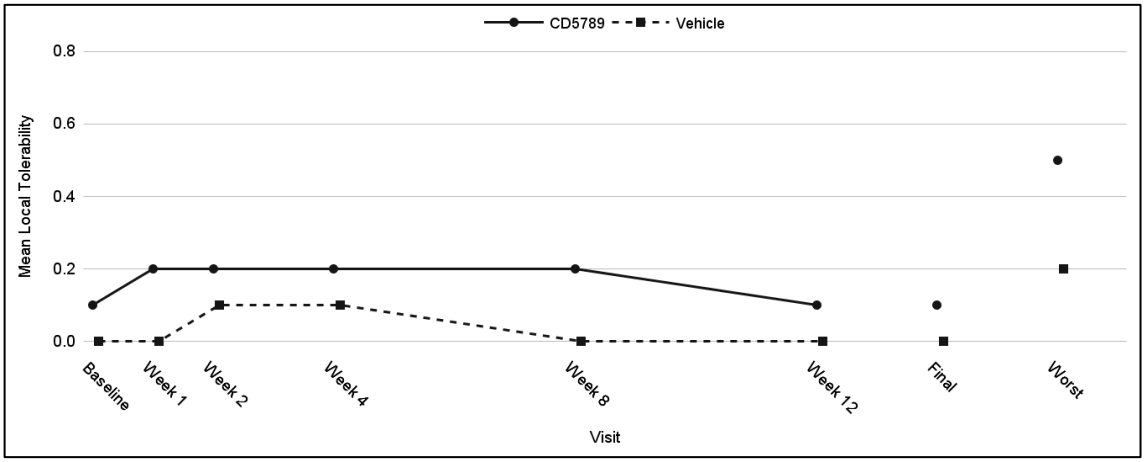


Abbildung 4-19: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Rötung
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

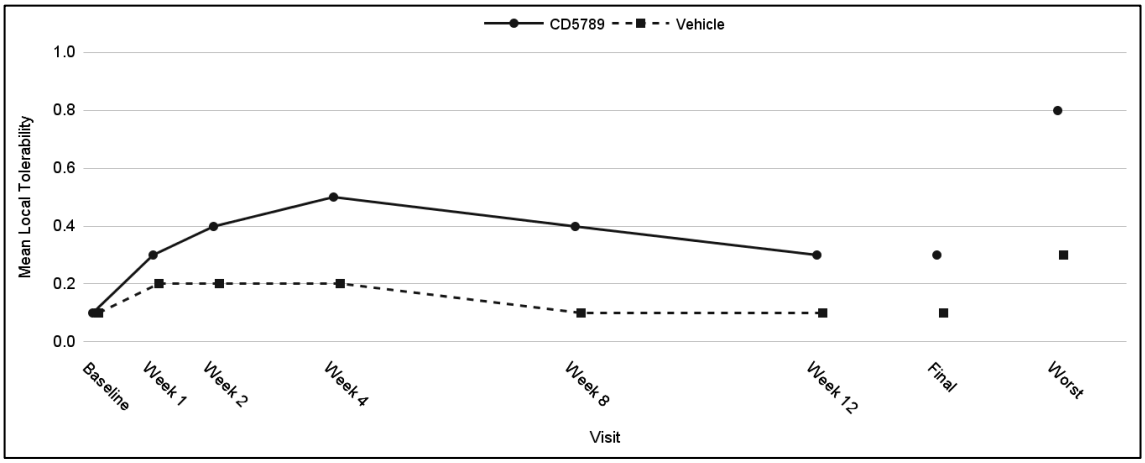
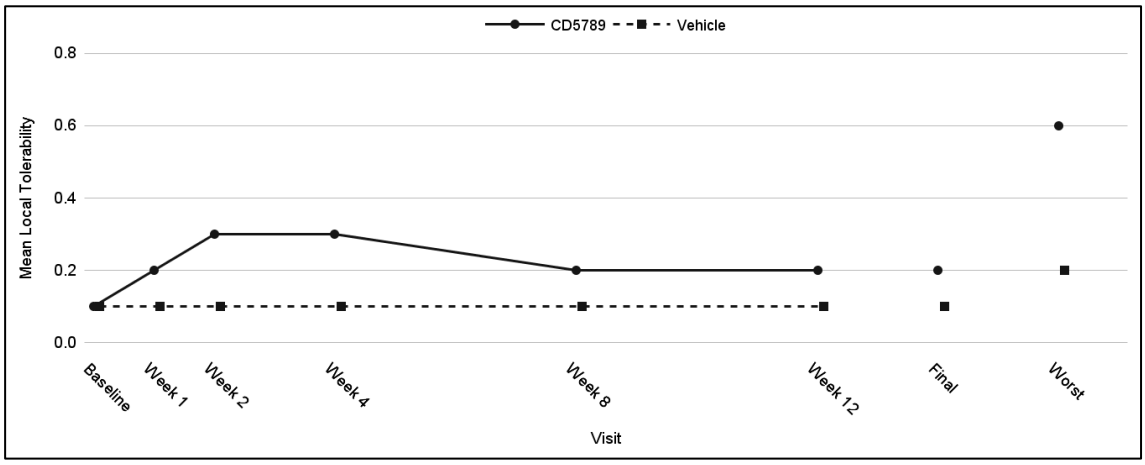


Abbildung 4-20: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Schuppung
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

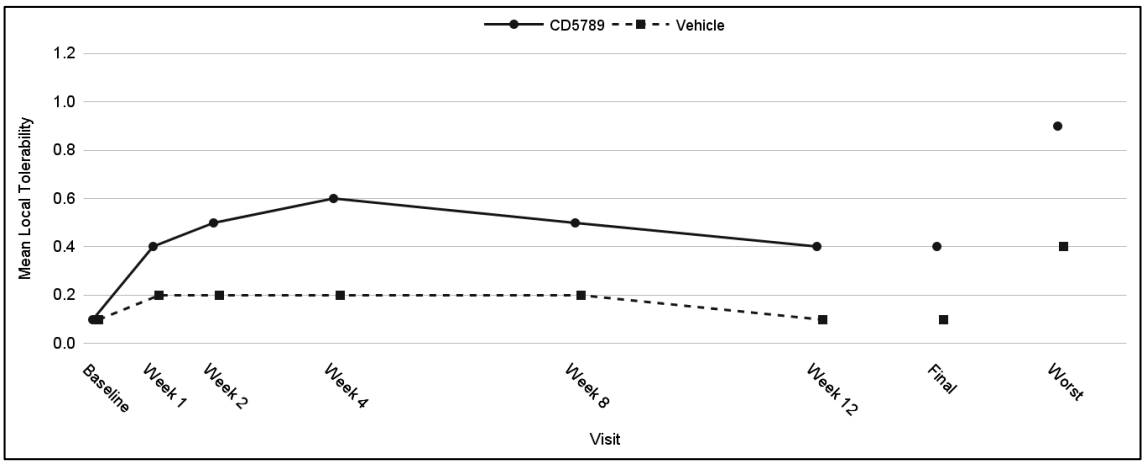
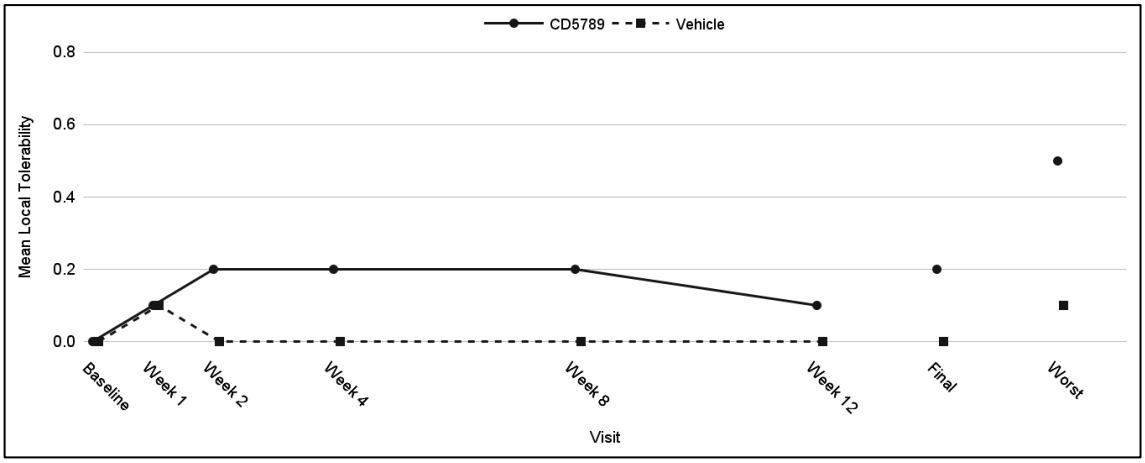


Abbildung 4-21: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Trockenheit
a) Studie PERFECT-1, b) Studie PERFECT-2

a)



b)

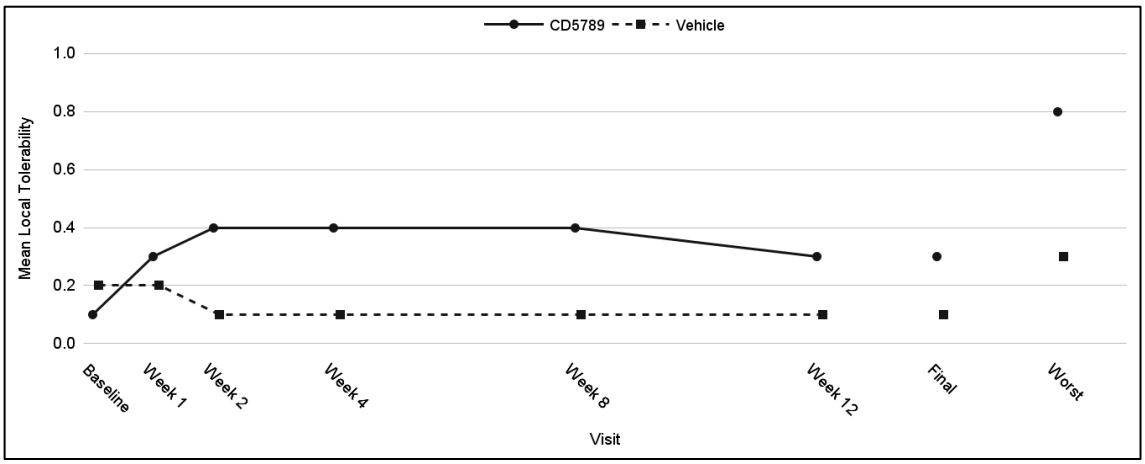


Abbildung 4-22: Lokale Verträglichkeit am Rumpf im Zeitverlauf – Stechen/Brennen
a) Studie PERFECT-1, b) Studie PERFECT-2

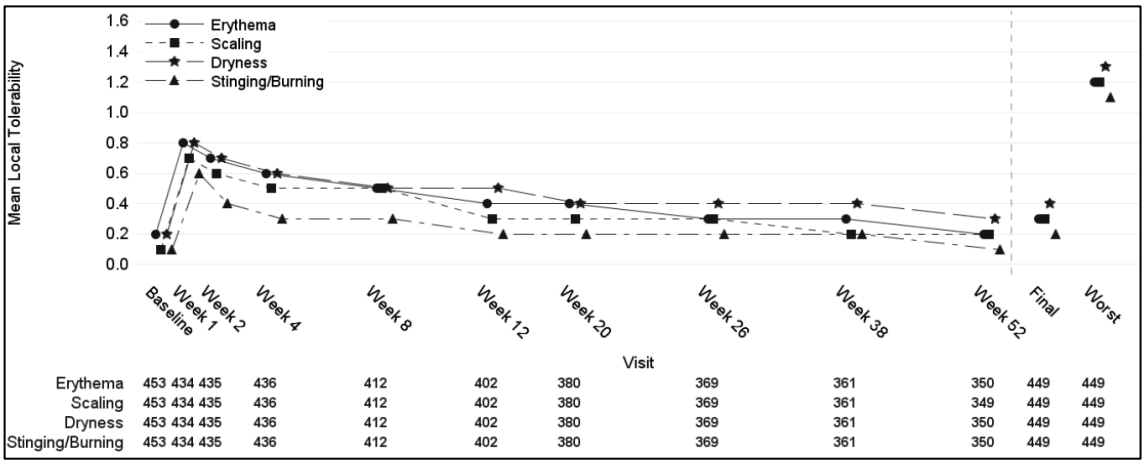
In Tabelle 4-53 und Abbildung 4-23 sind ergänzend die Ergebnisse zur lokalen Verträglichkeit aus der Studie SATISFY dargestellt. Die Ergebnisse bestätigen die Ergebnisse aus den Studien PERFECT-1 und -2: Hautreaktionen treten im Gesicht häufiger auf als am Rumpf, die meisten Ereignisse sind von milder oder moderater Schwere und die Unverträglichkeiten sind transient und treten mit einem Maximum in Woche 1 im Gesicht bzw. Woche 4 am Rumpf und somit frühzeitig im Laufe der Behandlung auf. Bei langfristiger Anwendung von Trifaroten ist keine Zunahme der Hautreaktionen zu beobachten.

Tabelle 4-53: Ergänzende Ergebnisse zur lokalen Verträglichkeit von Trifaroten im Gesicht und am Rumpf (schlechtester Post-Baseline-Wert über 52 Wochen; Studie SATISFY)

Symptom / Schweregrad	**Gesicht (N = 453 / 449) n (%)**	**Rumpf (N = 446 / 442) n (%)**
Rötung		
Mild (1)	210 (46,8)	131 (29,6)
Moderat (2)	111 (24,7)	76 (17,2)
Schwer (3)	10 (2,2)	24 (5,4)
Schuppung		
Mild (1)	210 (46,8)	137 (31,0)
Moderat (2)	131 (29,2)	48 (10,9)
Schwer (3)	10 (2,2)	11 (2,5)
Trockenheit		
Mild (1)	195 (43,4)	152 (34,4)
Moderat (2)	140 (31,2)	63 (14,3)
Schwer (3)	26 (5,8)	12 (2,7)
Stechen/Brennen		
Mild (1)	169 (37,6)	111 (25,1)
Moderat (2)	95 (21,2)	42 (9,5)
Schwer (3)	32 (7,1)	21 (4,8)

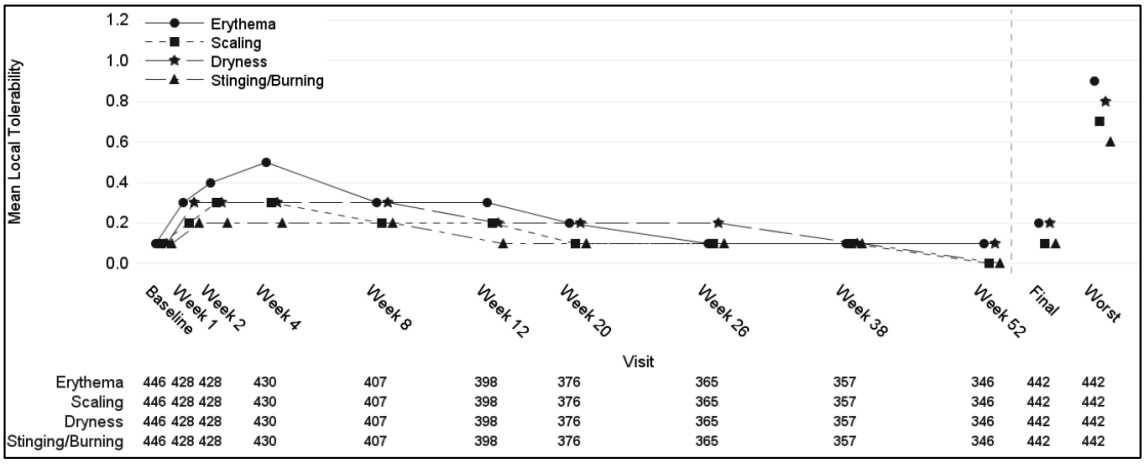
N: Patienten in der SAF- bzw. SAFT-Population / Anzahl Patienten in der SAF- bzw. SAFT-Population mit Post-Baseline-Wert; n: Patienten mit Ereignis; SAF: *Safety*; SAFT: *Safety on the trunk*

a)



Erythema	453	434	435	436	412	402	380	369	361	350	449	449
Scaling	453	434	435	436	412	402	380	369	361	349	449	449
Dryness	453	434	435	436	412	402	380	369	361	350	449	449
Stinging/Burning	453	434	435	436	412	402	380	369	361	350	449	449

b)



Erythema	446	428	428	430	407	398	376	365	357	346	442	442
Scaling	446	428	428	430	407	398	376	365	357	346	442	442
Dryness	446	428	428	430	407	398	376	365	357	346	442	442
Stinging/Burning	446	428	428	430	407	398	376	365	357	346	442	442

Abbildung 4-23: Ergänzende Ergebnisse zur lokalen Verträglichkeit aus SATISFY im Zeitverlauf
Rötung, Schuppung, Trockenheit und Stechen/Brennen im Gesicht (a) und am Rumpf (b)

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.

Für die Endpunkte der lokalen Verträglichkeit im Gesicht und am Rumpf wurde keine Meta-Analyse durchgeführt, da die Ergebnisse der Studien PERFECT-1 und PERFECT-2 rein deskriptiv dargestellt wurden.

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.

4.3.1.3.9 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1. [16]

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.
- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.
- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.
- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

[16] unbesetzt

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.

Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.

Für die Darstellung des medizinischen Nutzens im vorliegenden Dossier werden von den a priori geplanten Subgruppenanalysen die Faktoren Alter und Geschlecht als relevant erachtet. Aus den weiteren a priori geplanten Subgruppenanalysen ist aufgrund der demographischen Merkmale der deutschen Bevölkerung kein relevanter Erkenntnisgewinn für Deutschland zu erwarten (Details siehe Abschnitt 4.2.5.5). Subgruppenanalysen nach Krankheitsschwere sind nicht erforderlich, da ausschließlich Patienten mit mittelschwerer Akne eingeschlossen wurden.

Im vorliegenden Dossier werden somit lediglich die Ergebnisse für die Subgruppenanalysen nach Alter und Geschlecht für die Endpunkte Therapieerfolgsrate im Gesicht (IGA 0/1), Therapieerfolgsrate am Rumpf (PGA 0/1) sowie die Änderung der Anzahl der entzündlichen und nicht entzündlichen Läsionen im Gesicht bzw. am Rumpf dargestellt.

Tabelle 4-54: Matrix der durchgeführten Subgruppenanalysen

Endpunkt	**Studie**	**Alter [9 – 11 / 12 – 17 / ≥ 18 J.]**	**Geschlecht [Männer / Frauen]**
Therapieerfolgsrate im Gesicht (IGA 0/1)	PERFECT-1	●	●
	PERFECT-2	●	●
Therapieerfolgsrate am Rumpf (PGA 0/1)	PERFECT-1	○	○
	PERFECT-2	○	○
Anzahl der entzündlichen Läsionen im Gesicht	PERFECT-1	●	●
	PERFECT-2	●	●
Anzahl der nicht entzündlichen Läsionen im Gesicht	PERFECT-1	●	●
	PERFECT-2	●	●
Anzahl der entzündlichen Läsionen am Rumpf	PERFECT-1	○	○
	PERFECT-2	○	○
Anzahl der nicht entzündlichen Läsionen am Rumpf	PERFECT-1	○	○
	PERFECT-2	○	○

●: A priori im SAP geplante Subgruppenanalyse ○: Posthoc durchgeführte Subgruppenanalyse (im Studienbericht enthalten)

IGA: *Investigator´s Global Assessment*; PGA: *Physician´s Global Assessment*

Stellen Sie anschließend in Tabelle 4-55 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante (p < 0,05) Interaktionsterme.

In Tabelle 4-55 sind die Ergebnisse der Interaktionstests für die durchgeführten Subgruppenanalysen aufgelistet. Keiner der p-Werte unterschreitet das Signifikanzniveau von 0,05, somit liegen keine Belege für eine Effektmodifikation durch die Faktoren Alter und Geschlecht vor.

Tabelle 4-55: Ergebnisse des Interaktionsterms der Subgruppenanalysen je Endpunkt (Studien PERFECT-1 und -2)

Endpunkt	**Studie**	**p-Werte**	
		Alter [9 – 11 / 12 – 17 / ≥ 18 J.]	**Geschlecht [Männer / Frauen]**
Therapieerfolgsrate im Gesicht (IGA 0/1)	PERFECT-1	0,785	0,407
	PERFECT-2	0,775	0,629
Therapieerfolgsrate am Rumpf (PGA 0/1)	PERFECT-1	0,938	0,176
	PERFECT-2	0,755	0,696
Anzahl der entzündlichen Läsionen im Gesicht	PERFECT-1	0,144	0,454
	PERFECT-2	0,709	0,377
Anzahl der nicht entzündlichen Läsionen im Gesicht	PERFECT-1	0,231	0,740
	PERFECT-2	0,737	0,548
Anzahl der entzündlichen Läsionen am Rumpf	PERFECT-1	0,488	0,941
	PERFECT-2	0,173	0,634
Anzahl der nicht entzündlichen Läsionen am Rumpf	PERFECT-1	0,413	0,816
	PERFECT-2	0,691	0,598
IGA: *Investigator´s Global Assessment*; PGA: *Physician´s Global Assessment*			

Stellen Sie schließlich alle Subgruppenergebnisse dar.

Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich“ und „männlich“) als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig“ und „hoch“) eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt

dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.

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.

Auf eine detaillierte Darstellung der Ergebnisse der einzelnen Subgruppenanalysen wird an dieser Stelle verzichtet (tabellarische Darstellung siehe Anhang 4-G). Zum einen hat keiner der durchgeführten Interaktionstests einen Beleg für eine Effektmodifikation ergeben. Die Ergebnisse der einzelnen Subgruppenanalysen sind somit nicht relevant für die Bewertung der Studienergebnisse.

Zum anderen können die Studien PERFECT-1 und -2 nicht für die Identifikation von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen herangezogen werden, da in den Studien jeweils Trifaroten mit Vehikel und nicht mit der zweckmäßigen Vergleichstherapie verglichen wurde. Die Studien PERFECT-1 und -2 werden lediglich zur Darstellung des medizinischen Nutzens berücksichtigt. Aus diesem Grund erfolgt auch keine weitere Zusammenfassung der Subgruppenergebnisse in Meta-Analysen.

4.3.1.4 Liste der eingeschlossenen Studien - RCT

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).

Studie PERFECT-1 (18251)

- Studienbericht [34]
- Registereinträge [35-37]
- Publikation [38]
- Post-hoc-Auswertungen in Abschnitt 4.3.1.3

Studie PERFECT-2 (18252)

- Studienbericht [39]
- Registereinträge [40-42]

- Publikation [38]
- Post-hoc-Auswertungen in Abschnitt 4.3.1.3

Studie SATISFY (18250)

- Studienbericht [43]
- Registereinträge [44-46]
- Publikation [47]

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 sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der Internetseite des G-BA*
- *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 identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. ***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 4-56: 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 drei 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.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-57: 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 4-58: 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 4-59: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	**Verzerrungspotenzial auf Studienebene**	**Verblindung Endpunkterheber**	**Adäquate Umsetzung des ITT-Prinzips**	**Ergebnisunabhängige Berichterstattung**	**Keine sonstigen Aspekte**	**Verzerrungspotenzial 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 4-60: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	**Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3)**
<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: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

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 4.3.1.3.9.***

Nicht zutreffend.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

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).

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/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *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 4-61: 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

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-62: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	**<Mortalität>**	**<Gesundheits-bezogene Lebensqualität>**	**<Endpunkt>**	**<Endpunkt>**	**<Endpunkt>**
<Studie 1>	nein	ja	ja	ja	nein

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 4-63: 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 4-64: 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 4.3.1.3.9.***

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende 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).

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 sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *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

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-65: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	**<Mortalität>**	**<Gesundheits-bezogene Lebensqualität>**	**<Endpunkt>**	**<Endpunkt>**	**<Endpunkt>**
<Studie 1>	nein	ja	ja	ja	nein

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 4-66: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	**Operationalisierung**
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. 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.

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 4.3.1.3.9.***

Nicht zutreffend.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

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).

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.

Bei den Studien PERFECT-1 und -2 handelt es sich um randomisierte, kontrollierte, doppelblinde, multizentrische und internationale Phase-III-Zulassungsstudien. Beide Studien entsprechen der Evidenzstufe Ib und damit der höchsten Evidenzstufe für klinische Studien gemäß Verfahrensordnung des G-BA. Die Studie SATISFY ist eine unkontrollierte, multizentrische und internationale Phase-III-Studie. Die Studie entspricht der Evidenzstufe IV gemäß Verfahrensordnung des G-BA.

Die Ergebnisse der Studien sind auf den deutschen Versorgungskontext übertragbar (siehe Abschnitt 4.3.1.2.1).

In den Studien PERFECT-1 und -2 erfolgte die Zuteilung der Patienten zu den Behandlungsgruppen randomisiert und stratifiziert nach Zentrum, um eine Gleichheit der Behandlungsarme innerhalb der Studien zu gewährleisten. Die Vergleichbarkeit der Behandlungsarme zeigt sich in der Gegenüberstellung der Patientencharakteristika (siehe Abschnitt 4.3.1.2.1). Die interne Validität der Studien war somit gewährleistet. Das Verzerrungspotential auf Studienebene ist für beide Studien als niedrig einzustufen.

Die dargestellten Endpunkte sind valide und patientenrelevant (siehe Abschnitt 4.2.5.2). Für alle Endpunkte kann von einem niedrigen Verzerrungspotential ausgegangen werden.

Zusammenfassend kann für die vorliegenden Nachweise zum medizinischen Nutzen von Trifaroten aus den Studien PERFECT-1 und -2 von einer hohen qualitativen Ergebnissicherheit und somit einer hohen Aussagekraft ausgegangen werden. Für die ergänzend dargestellten Ergebnisse der Studie SATISFY ist aufgrund der niedrigen Evidenzstufe von einer geringeren qualitativen Ergebnissicherheit und somit einer geringeren Aussagekraft auszugehen.

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.

Die Studien PERFECT-1 und -2 erfüllen aufgrund des fehlenden Vergleichs mit der zweckmäßigen Vergleichstherapie nicht die formalen Anforderungen des G-BA an die Evidenz zum Nachweis eines Zusatznutzens. Somit stehen keine Daten zur Wirksamkeit und Sicherheit von Trifaroten zur Verfügung, die geeignet wären, einen Zusatznutzen gemäß § 35a SGB V zu belegen und dessen Ausmaß zu quantifizieren.

Die Studien PERFECT-1 und -2 werden zur Darstellung des medizinischen Nutzens von Trifaroten herangezogen. Ergänzend wird die unkontrollierte Studie SATISFY dargestellt, um Ergebnisse zur Langzeitwirksamkeit und -sicherheit von Trifaroten zu zeigen. Die Ergebnisse zum Nutzen von Trifaroten sind im ersten Teil dieses Abschnitts zusammengefasst.

Auch wenn auf Basis der bisher verfügbaren Evidenz der Nachweis bzw. die Quantifizierung eines Zusatznutzens formal nicht möglich ist, zeigen die bisher verfügbaren Daten zur Wirksamkeit und Sicherheit von Trifaroten Anhaltspunkte für einen Zusatznutzen. Diese werden im zweiten Teil dieses Abschnitts erläutert.

Schlussfolgerungen zum medizinischen Nutzen von Trifaroten

In der folgenden Tabelle 4-67 sind die Ergebnisse aus den Studien PERFECT-1 und -2 zur Wirksamkeit und Sicherheit von Trifaroten im Vergleich zur Behandlung mit Vehikel sowie die Ergebnisse der Meta-Analysen zusammengefasst.

Tabelle 4-67: Zusammenfassung der Ergebnisse aus PERFECT-1 und -2

Endpunktkategorie **Endpunkt**	**Trifaroten vs. Vehikel**	
	PERFECT-1 **PERFECT-2** **Effektschätzer [95 %-KI] p-Wert**	**Meta-Analyse** **Effektschätzer [95 %-KI]**
Morbidität		
Therapieerfolgsrate im Gesicht (IGA 0/1)	RR: 1,53 [1,24; 1,87] p < 0,001 RR: 1,71 [1,45; 2,02] p < 0,001	RR: 1,64 [1,44; 1,86]
Therapieerfolgsrate am Rumpf (PGA 0/1)	RR: 1,41 [1,18; 1,69] p < 0,001 RR: 1,50 [1,29; 1,75] p < 0,001	RR: 1,46 [1,30; 1,64]
Anzahl der Läsionen im Gesicht:		
Absolute Änderung entzündlicher Läsionen	LSMD: -3,6 [-4,9; -2,2] p < 0,001 LSMD: -5,6 [-6,9; -4,3] p < 0,001	LSMD: -4,64 [-5,57; -3,70]
Prozentuale Änderung entzündlicher Läsionen	LSMD: -9,6 [-13,5; -5,7] p < 0,001 LSMD: -16,0 [-19,7; -12,4] p < 0,001	LSMD: -13,01 [-15,68; -10,35]
Absolute Änderung nicht entzündlicher Läsionen	LSMD: -7,1 [-9,4; -4,8] p < 0,001 LSMD: -8,5 [-10,3; -6,6] p < 0,001	LSMD: -7,95 [-9,39; -6,51]
Prozentuale Änderung nicht entzündlicher Läsionen	LSMD: -14,0 [-18,1; -9,8] p < 0,001 LSMD: -15,5 [-19,0; -12,0] p < 0,001	LSMD: -14,88 [-17,55; -12,20]
Anzahl der Läsionen am Rumpf:		
Absolute Änderung entzündlicher Läsionen	LSMD: -2,5 [-4,0; -1,1] p < 0,001 LSMD: -5,7 [-7,2; -4,2] p < 0,001	LSMD: -4,05 [-5,09; -3,00]
Prozentuale Änderung entzündlicher Läsionen	LSMD: -7,5 [-11,5; -3,5] p < 0,001 LSMD: -15,1 [-20,1; -10,2] p < 0,001	LSMD: -10,50 [-13,61; -7,39]
Absolute Änderung nicht entzündlicher Läsionen	LSMD: -4,1 [-6,6; -1,7] p < 0,001 LSMD: -5,0 [-6,8; -3,3] p < 0,001	LSMD: -4,70 [-6,12; -3,27]
Prozentuale Änderung nicht entzündlicher Läsionen	LSMD: -8,7 [-13,3; -4,1] p < 0,001 LSMD: -12,4 [-16,3; -8,5] p < 0,001	LSMD: -10,85 [-13,83; -7,88]
Selbsteinschätzung der Verbesserung der Akne	RR: 1,57 [1,30; 1,88] p < 0,001 RR: 1,56 [1,32; 1,83] p < 0,001	RR: 1,56 [1,38; 1,77]

Endpunktkategorie Endpunkt	Trifaroten vs. Vehikel	
	PERFECT-1 PERFECT-2 Effektschätzer [95 %-KI] p-Wert	Meta-Analyse Effektschätzer [95 %-KI]
(vollständige oder deutliche Verbesserung)		
Gesundheitsbezogene Lebensqualität		
DLQI 0/1	RR: 0,95 [0,79; 1,13] p = 0,549 RR: 0,90 [0,74; 1,09] p = 0,274	RR: 0,93 [0,81; 1,06]
CDLQI 0/1	RR: 0,84 [0,67; 1,04] p = 0,110 RR: 0,82 [0,66; 1,02] p = 0,072	RR: 0,83 [0,71; 0,97]
Sicherheit/Verträglichkeit		
Jegliche UE	RR: 1,63 [1,34; 1,97] p < 0,001 RR: 1,05 [0,84; 1,32] p = 0,665	- (1)
SUE	RR: 1,92 [0,35; 10,42] p = 0,687 RR: 0,50 [0,09; 2,75] p = 0,687	RR: 0,98 [0,29; 3,28]
Therapieabbrüche wegen UE	RR: 13,41 [1,77; 101,66] p < 0,001 RR: 10,10 [1,30; 78,65] p = 0,006	RR: 11,66 [2,76; 49,27]
Lokale Verträglichkeit	*nur deskriptiv dargestellt*	

(1) Aufgrund der vorliegenden Heterogenität ist die Bildung eines gemeinsamen Effektschätzers nicht sinnvoll.

CDLQI: *Children´s Dermatology Life Quality Index*; DLQI: *Dermatology Life Quality Index*; IGA: *Investigator´s Global Assessment*; KI: Konfidenzintervall; LSMD: *least squares mean difference*; PGA: *Physician´s Global Assessment*; RR: Relatives Risiko; SUE: schwerwiegende unerwünschte Ereignisse; UE: unerwünschte Ereignisse

Morbidität

Die Ergebnisse der Phase-III-Studien PERFECT-1 und -2 belegen die Wirksamkeit der Behandlung mit Trifaroten. Für alle untersuchten Morbiditätsendpunkte zeigt sich in beiden Studien eine statistisch signifikante Überlegenheit von Trifaroten gegenüber Vehikel im Hinblick auf die Verringerung der Krankheitssymptome im Gesicht und am Rumpf, welche sich jeweils auch in der meta-analytischen Zusammenfassung der Ergebnisse bestätigt:

- Signifikant mehr Patienten weisen nach zwölf Wochen unter Trifaroten eine vollständig oder nahezu vollständig abgeheilte Haut im Gesicht (IGA 0/1) auf.

- Signifikant mehr Patienten weisen nach zwölf Wochen unter Trifaroten eine vollständig oder nahezu vollständig abgeheilte Haut am Rumpf (PGA 0/1) auf.

- Die Anzahl der entzündlichen und der nicht entzündlichen Läsionen im Gesicht und am Rumpf verringert sich innerhalb der zwölfwöchigen Behandlungsphase signifikant stärker unter Trifaroten.

- Signifikant mehr Patienten unter Trifaroten schätzen die Verbesserung ihrer Akne nach zwölf Wochen Behandlung als vollständige oder zumindest deutliche Verbesserung ein.

In der Langzeitstudie SATISFY zeigt sich darüber hinaus, dass der Anteil der Patienten mit einem Therapieerfolg gemäß IGA/PGA, d.h. einer vollständigen oder fast vollständigen Abheilung der Haut im Gesicht und am Rumpf, sowie der Anteil der Patienten mit einer relevanten Verbesserung der Akne bei einer Behandlung über Woche 12 hinaus weiter steigt. Die Wirksamkeit der Behandlung mit Trifaroten ist somit langanhaltend und nimmt bei den Patienten, die nach zwölf Wochen noch nicht den maximalen Therapieeffekt erreicht haben, weiter zu.

Gesundheitsbezogene Lebensqualität

Die Ergebnisse der Studien PERFECT-1 und -2 zur Veränderung der gesundheitsbezogenen Lebensqualität gemäß DLQI zeigen nach zwölf Wochen Behandlung für die Patienten ab 17 Jahren keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen. Für die jüngeren Studienteilnehmer, die den CDLQI zur Erhebung der gesundheitsbezogenen Lebensqualität nutzten, ergibt die Meta-Analyse einen statistisch signifikanten Nachteil für Trifaroten hinsichtlich des Anteils der Patienten, die nach zwölf Wochen Behandlung einen CDLQI 0/1 erreichen. Die Einzelergebnisse der Studien zeigen keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen.

Die mittlere Beeinträchtigung der Lebensqualität bei den jüngeren Studienteilnehmern ist zu Beginn der Studien gering und eine starke Verbesserung infolge der Behandlung ist nicht zu erwarten. Dies kann daran liegen, dass Kinder und Jugendliche die Belastung durch Akne anders wahrnehmen als Erwachsene. Sie sind sich nicht bewusst, dass es eine chronische Erkrankung der Haut mit dem Risiko der Entstehung bleibender Narben ist, die man medikamentös behandeln kann und sollte. Die Akne wird in diesem Lebensalter sowohl von betroffenen jungen Patienten als auch von Personen im sozialen Umfeld weitestgehend als assoziierte, temporäre Veränderung oder Unreinheit der Haut wahrgenommen und akzeptiert. Zum anderen ist die Akne am Rumpf bei Kindern und Jugendlichen im frühen Alter wenig ausgeprägt [51]. Patienten, bei denen die Akne auf das Gesicht beschränkt ist, zeigen eine geringere Beeinträchtigung der gesundheitsbezogenen Lebensqualität als Patienten mit großflächiger Akne im Gesicht und am Rumpf [48-50].

Die Behandlung der mittelschweren Akne mit einer aktiven Substanz wie Trifaroten ist im Gegensatz zur Behandlung mit Vehikel mit transienten und überwiegend milden oder moderaten Nebenwirkungen der Haut verbunden. Es ist anzunehmen, dass bei Patienten, die ihre Akne nicht als behandlungswürdige Hautveränderung wahrnehmen, das temporäre Auftreten von Nebenwirkungen die Bewertung der gesundheitsbezogenen Lebensqualität stärker beeinflusst, als das Empfinden der Verbesserung einer chronischen Erkrankung der

Haut. Daraus könnte der beobachtete Nachteil für Trifaroten in dieser Subpopulation resultieren.

Zusammenfassend zeigen die Ergebnisse beider Studien gleichwohl, dass sich infolge der Behandlung mit Trifaroten die gesundheitsbezogene Lebensqualität im Vergleich zum Studienbeginn verbessert und dass ein großer Teil der Patienten nach Abschluss der Behandlung keine Beeinträchtigung der Lebensqualität wahrnimmt.

Sicherheit und Verträglichkeit

Die Auswertungen zu den dokumentierten UE belegen die Sicherheit von Trifaroten.

Hinsichtlich jeglicher UE ergibt sich in der Studie PERFECT-1 ein statisch signifikanter Nachteil für Trifaroten gegenüber Vehikel. In der Studie PERFECT-2 zeigt sich jedoch kein Unterschied in der Häufigkeit jeglicher UE.

Die Art der auftretenden UE entspricht dem erwarteten Nebenwirkungsprofil. Die häufigsten UE sind Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“, dabei wurden insbesondere die PT „Reizung an der Applikationsstelle“ und „Pruritus an der Applikationsstelle“ dokumentiert. Diese Ereignisse treten unter Trifaroten deutlich häufiger auf als unter Vehikel. Weiterhin sind Ereignisse der SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ unter Trifaroten häufiger als unter Vehikel.

SUE sind sowohl unter Vehikel als auch unter Trifaroten sehr selten, und keines der schwerwiegenden Ereignisse wird als mit der Studienmedikation in Zusammenhang stehend eingestuft. Es besteht kein statistisch signifikanter Unterschied zwischen der Behandlung mit Trifaroten und Vehikel im Hinblick auf das Auftreten von SUE.

Therapieabbrüche aufgrund von UE treten häufiger unter Trifaroten auf, sind jedoch in beiden Studien selten. Es besteht ein statistisch signifikanter Nachteil für Trifaroten. Am häufigsten führen Ereignisse der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ und insbesondere das Ereignis „Reizung an der Applikationsstelle“ zu einem Abbruch der Behandlung mit Trifaroten.

Hautreaktionen bzw. Hautirritationen im Gesicht und am Rumpf treten häufiger bei Anwendung von Trifaroten im Vergleich zur Kontrollgruppe auf. Die Ereignisse sind jedoch in den meisten Fällen von milder oder moderater Schwere. Zudem sind die auftretenden Hautreaktionen in der überwiegenden Mehrheit der Fälle transient und klingen im Laufe der Behandlung schnell ab.

Die Daten der Langzeitstudie SATISFY bestätigen das gute Sicherheitsprofil von Trifaroten. Die Mehrheit der aufgetretenen UE wurde in den ersten drei Monaten der Behandlung dokumentiert. Insbesondere die Therapieabbrüche aufgrund von UE treten fast alle in den ersten drei Behandlungsmonaten auf. Bei langfristiger Anwendung von Trifaroten kommt es folglich nicht zu einer Zunahme der UE und nicht zu einem gehäuften Auftreten von UE, die einen

Abbruch der Anwendung von Trifaroten erfordern. Somit ist die langfristige Behandlung mit Trifaroten und insbesondere die großflächige Anwendung am Rumpf über einen Zeitraum von 52 Wochen nachgewiesenermaßen gut verträglich und sicher.

Zusammenfassung der Aussagen zum therapeutisch relevanten Nutzen

Die Studien PERFECT-1 und -2 belegen mit hoher Aussagekraft die signifikante Überlegenheit von Trifaroten gegenüber einer Behandlung mit Vehikel im Hinblick auf die Verringerung der Krankheitslast durch eine ausgeprägte und klinisch relevante Verringerung der Symptome, die mit der Erkrankung *Acne vulgaris* einhergehen. Die Wirksamkeit der Behandlung mit Trifaroten ist über den Zeitraum von zwölf Wochen hinaus anhaltend bzw. nimmt bei den Patienten, die nach zwölf Wochen noch nicht den maximalen Therapieeffekt erreicht haben, weiter zu.

Die unter Trifaroten auftretenden UE entsprechen dem erwarteten Nebenwirkungsprofil. Die häufigsten UE sind Reaktionen der Haut, die überwiegend von milder bis moderater Schwere sind und nur transient auftreten. Bei der Anwendung von Trifaroten über 52 Wochen kommt es nicht zu einer relevanten Zunahme von Nebenwirkungen. Somit ist die Sicherheit und Verträglichkeit einer langfristigen Anwendung von Trifaroten und insbesondere einer großflächigen Anwendung am Rumpf bestätigt.

Schlussfolgerungen zum Zusatznutzen von Trifaroten

Auf Basis der derzeit verfügbaren Evidenz ist der Nachweis eines Zusatznutzens von Trifaroten gemäß § 35a SGB V und dessen Quantifizierung formal nicht möglich. Dennoch gibt es Anhaltspunkte für einen Zusatznutzen von Trifaroten gegenüber der zweckmäßigen Vergleichstherapie, der zum einen in den vorliegenden Daten und zum anderen in der zweckmäßigen Vergleichstherapie und den weiteren Therapieoptionen begründet ist.

Die Evidenz zur Wirksamkeit und Sicherheit von Trifaroten aus den beiden Phase-III-Zulassungsstudien belegt die klinische Wirksamkeit von Trifaroten hinsichtlich der Verringerung der Akne-Symptomatik bei gleichzeitig guter Verträglichkeit. Insbesondere liefern die Studiendaten erstmalig hochwertige Evidenz für die Wirksamkeit eines topischen Retinoids bei der Behandlung der Akne am Rumpf. Die Daten zeigen die schnell einsetzende klinische Wirkung im Hinblick auf die Verringerung der Läsionen am Rumpf. Gleichzeitig belegen die Daten aus der Studie SATISFY die gute Verträglichkeit der Behandlung mit Trifaroten auch bei täglicher großflächiger Anwendung am Rumpf über einen Zeitraum von 52 Wochen.

Die Datenlage zur Behandlung der Akne am Rumpf ist im Hinblick auf die Wirksamkeit und Sicherheit bei großflächiger Anwendung für die bisher zugelassenen Therapieoptionen sehr begrenzt [52]. Für keine der in Deutschland zugelassenen Therapien liegen Daten aus randomisierten klinischen Studien zur Behandlung der Akne am Rumpf vor. Die Zulassung der bisher verfügbaren Wirkstoffe ist jeweils auf Basis klinischer Studien an Patienten mit Akne im Gesicht erfolgt. Dies spiegelt sich auch in den mangelnden Leitlinien-Empfehlungen zur Behandlung der Akne am Rumpf wider [7]. Trifaroten ist das einzige topische Retinoid, welches aufgrund seiner nachgewiesenen Wirksamkeit und Sicherheit explizit zur

großflächigen Behandlung der Akne im Gesicht und am Rumpf zugelassen ist. Dies stellt einen Zusatznutzen gegenüber anderen Therapieoptionen inklusive der zweckmäßigen Vergleichstherapie dar.

Die Behandlung der Akne am Rumpf erfolgt häufig mittels einer systemischen anstelle einer topischen Therapie [52], obwohl dies bei Vorliegen einer mittelschweren Akne als eine Übertherapie anzusehen ist, welche im Falle des systemischen Einsatzes von Isotretinoin sowie der meisten Antibiotika keine zulassungskonforme Arzneimittelanwendung darstellt. Insbesondere die systemische Behandlung mit Isotretinoin ist mit dem Risiko schwerwiegender therapiebedingter Nebenwirkungen verbunden und bedarf einer strikten Überwachung durch den behandelnden Arzt [53]. Das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) hat daher im September 2019 auf der Grundlage eines Gutachtens des Ausschusses für Humanarzneimittel (CHMP) der EMA aktualisierte Maßnahmen zur Schwangerschaftsverhütung sowie Warnhinweise zu neuropsychiatrischen Erkrankungen bei oraler Anwendung eines Retinoids angeordnet [54]. Beim systemischen Einsatz von Antibiotika ist grundsätzlich das Risiko der Entwicklung bakterieller Resistenzen zu beachten [7]. Durch die Verfügbarkeit eines neuen Wirkstoffes zur Behandlung der mittelschweren Akne kann – insbesondere bei der Behandlung von Akne am Rumpf – potentiell der systemische Einsatz von Isotretinoin und Antibiotika reduziert werden.

Ein weiterer Aspekt, der einen grundsätzlichen Zusatznutzen von Trifaroten gegenüber der zweckmäßigen Vergleichstherapie Clindamycin und Benzoylperoxid darstellt, ist die Vermeidung einer topischen antibiotikahaltigen Therapie durch Trifaroten. Der Einsatz von topischen Antibiotika wie Clindamycin ist mit dem Risiko der bakteriellen Resistenzentwicklung verbunden. Das potentielle Risiko korreliert mit der Dauer der Anwendung sowie der Größe der behandelten Hautoberfläche. Aus diesem Grund ist die Anwendung von Clindamycin und Benzoylperoxid laut Fachinformation auf zwölf Wochen zu beschränken und die Richtlinien zur sachgemäßen Anwendung antimikrobieller Arzneimittel sowie verfügbare lokale Informationen zur Prävalenz von Resistenzen müssen beachtet werden [5, 6]. Auch die europäische S3-Leitlinie zur Behandlung der Akne weist explizit auf das potentielle Risiko von Antibiotika-Resistenzen hin [7]. Folglich stellt die Verfügbarkeit einer wirksamen und gut verträglichen Therapiealternative zu einer antibiotikahaltigen Therapie per se einen Zusatznutzen dar.

Über die genannten Vorteile hinaus sind weitere positive Aspekte zu nennen: Durch den spezifischeren Wirkmechanismus von Trifaroten im Vergleich zu anderen Retinoiden ist eine Verringerung des Risikos für Nebenwirkungen zu erwarten. Des Weiteren hat Trifaroten im Gegensatz zu den Benzoylperoxid-haltigen Arzneimitteln der zweckmäßigen Vergleichstherapie keine bleichende Wirkung auf Haare und Textilien, was insbesondere für Patienten mit Akne am Rumpf einen Vorteil in der Anwendung darstellt, der potentiell zu einer besseren Therapieadhärenz führen kann, da es im Therapiealltag nicht zu einer Entfärbung von Kleidung oder Bettzeug kommt.

Zusammenfassend besteht für Trifaroten ein **Anhaltspunkt für einen nicht quantifizierbaren Zusatznutzen**. Dieser basiert zum einen auf dem patientenrelevanten therapeutische Nutzen,

insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, bei Patienten mit mittelschwerer *Acne vulgaris* im Gesicht und/oder am Rumpf, welcher für die Anwendung am Rumpf erstmalig für ein Retinoid im Rahmen von randomisierten, kontrollierten klinischen Studien belegt wurde. Zum anderen ermöglicht Trifaroten als wirksame und gut verträgliche antibiotikafreie, topische Therapiealternative potentiell die Vermeidung einer topischen oder systemischen Antibiotika-Therapie sowie die Vermeidung der systemischen Anwendung von Isotretinoin.

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 4-68: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten ab zwölf Jahren mit mittelschwerer *Acne vulgaris* im Gesicht und/oder am Rumpf	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[17], Molenberghs 2010[18]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende 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[19]) 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[20]) soll ausreichend begründet

[17] Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.
[18] 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.
[19] Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.
[20] Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

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.

4.6 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. Galderma Laboratorium GmbH (2020): Selgamis 50 Mikrogramm/g Creme; Fachinformation. Stand: April 2020 [Zugriff: 13.08.2020]. URL: http://www.fachinfo.de.
2. Gemeinsamer Bundesausschuss (G-BA) (2019): Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV; Beratungsanforderung 2019-B-086.
3. Galderma Laboratorium GmbH (2008): Epiduo 0,1% / 2,5% Gel; Fachinformation. Stand: Juni 2019 [Zugriff: 29.06.2020]. URL: http://www.fachinfo.de.
4. Galderma Laboratorium GmbH (2016): Epiduo Forte; Fachinformation. Stand: März 2020 [Zugriff: 29.06.2020]. URL: http://www.fachinfo.de.
5. Stiefel GmbH & Co. KG (2004): DUAC Akne Gel 10 mg/g/50 mg/g; Fachinformation. Stand: August 2019 [Zugriff: 29.06.2020]. URL: http://www.fachinfo.de.
6. Stiefel GmbH & Co. KG (2013): Duac® 10 mg/g + 30 mg/g Gel; Fachinformation. Stand: Oktober 2014 [Zugriff: 29.06.2020]. URL: http://www.fachinfo.de.

7. Nast A, Dreno B, Bettoli V, Bukvic Mokos Z, Degitz K, Dressler C, et al. (2016): European evidence-based (S3) guideline for the treatment of acne - update 2016 - long version. [Zugriff: 20.02.2020]. URL: https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fjdv.13776&file=jdv13776-sup-0004-DocumentS1.pdf.
8. Gollnick HP, Bettoli V, Lambert J, Araviiskaia E, Binic I, Dessinioti C, et al. (2016): A consensus-based practical and daily guide for the treatment of acne patients. Journal of the European Academy of Dermatology and Venereology : JEADV; 30(9):1480-90.
9. European Medicines Agency (EMA) (2001): ICH Topic E 11 Clinical Investigation of Medicinal Products in the Paediatric Population, Step 5 - NOTE FOR GUIDANCE ON CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS IN THE PAEDIATRIC POPULATION (CPMP/ICH/2711/99). [Zugriff: 11.04.2020]. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-1.pdf.
10. European Medicines Agency (EMA) (2017): ICH E11(R1) guideline on clinical investigation of medicinal products in the pediatric population, Step 5. [Zugriff: 11.04.2020]. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e11r1-guideline-clinical-investigation-medicinal-products-pediatric-population-revision-1_en.pdf.
11. Europäische Union (EU) (2006): VERORDNUNG (EG) Nr. 1901/2006 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 12. Dezember 2006 über Kinderarzneimittel und zur Änderung der Verordnung (EWG) Nr. 1768/92, der Richtlinien 2001/20/EG und 2001/83/EG sowie der Verordnung (EG) Nr. 726/2004. [Zugriff: 27.03.2020]. URL: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:378:0001:0019:DE:PDF.
12. Europäische Union (EU) (2001): RICHTLINIE 2001/20/EG DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 4. April 2001 zur Angleichung der Rechts- und Verwaltungsvorschriften der Mitgliedstaaten über die Anwendung der guten klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Humanarzneimitteln. [Zugriff: 27.03.2020]. URL: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2001L0020:20090807:DE:PDF.
13. European Medicines Agency (EMA) (2017): European Medicines Agency decision P/0099/2017 of 11 April 2017 on the acceptance of a modification of an agreed paediatric investigation plan for trifarotene (EMEA-001492-PIP01-13-M01) in accordance with Regulation (EC) No 1901/2006 of the European Parliament and of the Council [Zugriff: 07.04.2020]. URL: https://www.ema.europa.eu/en/documents/pip-decision/p/0099/2017-ema-decision-11-april-2017-acceptance-modification-agreed-paediatric-investigation-plan_en.pdf.
14. Tan JK, Fung K, Bulger L (2006): Reliability of dermatologists in acne lesion counts and global assessments. J Cutan Med Surg; 10(4):160-5.
15. Tan JK, Tang J, Fung K, Gupta AK, Thomas DR, Sapra S, et al. (2007): Development and validation of a comprehensive acne severity scale. J Cutan Med Surg; 11(6):211-6.

16. U.S. Department of Health and Human Services - Food and Drug Administration (FDA) - Center for Drug Evaluation and Research (2018): Acne Vulgaris: Establishing Effectiveness of Drugs Intended for Treatment, Guidance for Industry. [Zugriff: 20.02.2020]. URL: https://www.fda.gov/media/71152/download.
17. Swedish Medical Products Agency (2019): Public Assessment Report - Scientific discussion: Aklief (trifarotene), SE/H/1863/01/DC.
18. Gemeinsamer Bundesausschuss (G-BA) (2017): 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 – Opicapon. [Zugriff: 07.04.2020]. URL: https://www.g-ba.de/downloads/40-268-4247/2017-03-16_AM-RL-XII_Opicapon_D-258_TrG.pdf.
19. Gemeinsamer Bundesausschuss (G-BA) (2018): 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 – Extrakt aus Cannabis sativa L., folium cum flore (Wirkstoffkombination aus Delta-9-Tetrahydrocannabinol und Cannabidiol) (Neubewertung nach Fristablauf). [Zugriff: 07.04.2020]. URL: https://www.g-ba.de/downloads/40-268-5386/2018-11-01_AM-RL-XII_Extrakt-Cannabis-sativa_D-358_TrG.pdf.
20. Gemeinsamer Bundesausschuss (G-BA) (2019): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Nintedanib (Neubewertung eines Orphan-Drugs nach Überschreitung der 50 Mio. Euro Grenze). [Zugriff: 07.04.2020]. URL: https://www.g-ba.de/downloads/40-268-6058/2019-10-17_AM-RL-XII_Nintedanib_D-450_TrG.pdf.
21. Stein Gold L, Weiss J, Rueda MJ, Liu H, Tanghetti E (2016): Moderate and Severe Inflammatory Acne Vulgaris Effectively Treated with Single-Agent Therapy by a New Fixed-Dose Combination Adapalene 0.3 %/Benzoyl Peroxide 2.5 % Gel: A Randomized, Double-Blind, Parallel-Group, Controlled Study. American Journal of Clinical Dermatology; 17(3):293-303.
22. Finlay AY, Khan GK (1994): Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol; 19(3):210-6.
23. Lewis-Jones MS, Finlay AY (1995): The Children's Dermatology Life Quality Index (CDLQI): initial validation and practical use. The British journal of dermatology; 132(6):942-9.
24. Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY (2005): Translating the science of quality of life into practice: What do dermatology life quality index scores mean? J Invest Dermatol; 125(4):659-64.
25. Marron SE, Chernyshov PV, Tomas-Aragones L (2019): Quality-of-Life Research in Acne Vulgaris: Current Status and Future Directions. Am J Clin Dermatol; 20(4):527-38.
26. Olsen JR, Gallacher J, Finlay AY, Piguet V, Francis NA (2016): Quality of life impact of childhood skin conditions measured using the Children's Dermatology Life Quality Index (CDLQI): a meta-analysis. The British journal of dermatology; 174(4):853-61.

27. Gemeinsamer Bundesausschuss (G-BA) (2017): 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 – Secukinumab (Neubewertung aufgrund neuer wissenschaftlicher Erkenntnisse). [Zugriff: 30.06.2020]. URL: https://www.g-ba.de/downloads/39-261-3037/2017-08-17_AM-RL-XII_Secukinumab_D-276_BAnz.pdf.
28. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2017): Allgemeine Methoden, Version 5.0 vom 10.07.2017. [Zugriff: 31.03.2020]. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.
29. Hoffman LK, Bhatia N, Zeichner J, Kircik LH (2018): Topical Vehicle Formulations in the Treatment of Acne. Journal of drugs in dermatology : JDD; 17(6):s6-s10.
30. Kircik LH (2011): Importance of vehicles in acne therapy. Journal of drugs in dermatology : JDD; 10(6):s17-23.
31. Chiou WL (2012): Low intrinsic drug activity and dominant vehicle (placebo) effect in the topical treatment of acne vulgaris. Int J Clin Pharmacol Ther; 50(6):434-7.
32. Tan JK, Bhate K (2015): A global perspective on the epidemiology of acne. The British journal of dermatology; 172 Suppl 1:3-12.
33. Barratt H, Hamilton F, Car J, Lyons C, Layton A, Majeed A (2009): Outcome measures in acne vulgaris: systematic review. The British journal of dermatology; 160(1):132-6.
34. Galderma R&D (2018): Clinical Study Report – RD.03.SRE.18251; A multicenter, randomized, double-blind, parallel-group vehicle-controlled study to compare the efficacy and safety of CD5789 50µg/g cream versus vehicle cream in subjects with acne vulgaris.
35. Galderma (2015): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - ClinicalTrials.gov (NCT02566369). Stand des Eintrags: 12.11.2019. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/show/NCT02566369
36. Galderma (2015): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - WHO ICTRP (NCT02566369). Stand des Eintrags: 26.08.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02566369
37. Galderma (2016): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - EU-CTR (2016-002860-15). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-002860-15
38. Tan J, Thiboutot D, Popp G, Gooderham M, Lynde C, Del Rosso J, et al. (2019): Randomized phase 3 evaluation of trifarotene 50 mug/g cream treatment of moderate facial and truncal acne. J Am Acad Dermatol; 80(6):1691-9.
39. Galderma R&D (2018): Clinical Study Report – RD.03.SRE.18252; A multicenter, randomized, double-blind, parallel-group vehicle-controlled study to compare the

efficacy and safety of CD5789 50 µg/g cream versus vehicle cream in subjects with acne vulgaris.

40. Galderma (2015): RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - ClinicalTrials.gov (NCT02556788). Stand des Eintrags: 12.11.2019. [Zugriff: 29.06.2020]. URL: https://ClinicalTrials.gov/show/NCT02556788
41. Galderma (2015): RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - EU-CTR (2015-002540-13). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-002540-13
42. Galderma (2015): RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - WHO ICTRP (NCT02556788). Stand des Eintrags: 09.09.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02556788
43. Galderma R&D (2018): Clinical Study Report – RD.06.SRE.18250; A long-term safety and efficacy study of CD5789 50µg/g cream in subjects with acne vulgaris.
44. Galderma (2014): RD.06.SPR.18250 - A Long-Term Safety and Efficacy Study of CD5789 50µg/g (Trifarotene) Cream in Subjects with Acne Vulgaris - ClinicalTrials.gov (NCT02189629). Stand des Eintrags: 14.11.2019. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/show/NCT02189629
45. Galderma (2014): RD.06.SPR.18250 - A Long-Term Safety and Efficacy Study of CD5789 50µg/g Cream in Subjects with Acne Vulgaris - EU-CTR (2014-001755-23). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001755-23
46. Galderma (2014): RD.06.SPR.18250 - A Long-Term Safety and Efficacy Study of CD5789 (Trifarotene) 50µg/g Cream in Subjects with Acne Vulgaris - WHO ICTRP (NCT02189629). Stand des Eintrags: 02.11.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02189629
47. Blume-Peytavi U, Fowler J, Kemeny L, Draelos Z, Cook-Bolden F, Dirschka T, et al. (2019): Long-term safety and efficacy of trifarotene 50 mug/g cream, a first-in-class RAR-gamma selective topical retinoid, in patients with moderate facial and truncal acne. Journal of the European Academy of Dermatology and Venereology : JEADV; 34(1):166-73.
48. Tan J, et al. (2020): The burden of combined facial and truncal acne (submitted abstract).
49. Tan J, Chavda R, Le Calve P, Guillaume X, Tsankova A, Dreno B (2020): Patient perspective on burden of truncal acne and impact on quality of life (submitted abstract).
50. Kantar, Galderma (2020): Burden of combined facial and truncal acne (final post-board report) - facial/truncal acne: quantitative survey

51. Eichenfield LF, Krakowski AC, Piggott C, Del Rosso J, Baldwin H, Friedlander SF, et al. (2013): Evidence-based recommendations for the diagnosis and treatment of pediatric acne. Pediatrics; 131 Suppl 3:S163-86.
52. Poli F, Auffret N, Leccia M-T, Claudel J-P, Dréno B (2020): Truncal acne, what do we know? Journal of the European Academy of Dermatology and Venereology; doi:101111/jdv16634;
53. Almirall Hermal GmbH (2002): Aknenormin® 10 mg / 20 mg Weichkapseln; Fachinformation. Stand: März 2019 [Zugriff: 24.04.2020]. URL: http://www.fachinfo.de.
54. Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) (2019): Retinoide: Aktualisierte Maßnahmen zur Schwangerschaftsverhütung sowie Warnhinweise zu neuropsychiatrischen Erkrankungen bei oraler Anwendung. Ergänzungsbescheid vom 08.08.2019. [Zugriff: 20.05.2020]. URL: https://www.bfarm.de/SharedDocs/Risikoinformationen/Pharmakovigilanz/DE/RV_STP/m-r/retinoide.html.
55. 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	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[21]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

[21] 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 4-69: Bibliografische Literaturrecherche in EMBASE – Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	16.06.2020	
Zeitsegment	1974 to 2020 June 15	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [55]	
#	**Suchbegriffe**	**Ergebnis**
1	trifaroten*.mp.	20
2	exp trifarotene/	19
3	CD-5789.mp.	1
4	CD5789.mp.	1
5	CD 5789.mp.	1
6	aklief.mp.	4
7	selgamis.mp.	0
8	1 or 2 or 3 or 4 or 5 or 6 or 7	20
9	random*.tw.	1.540.309
10	placebo*.mp.	454.001
11	double-blind*.tw.	209.692
12	9 or 10 or 11	1.794.480
13	8 and 12	4

Tabelle 4-70: Bibliografische Literaturrecherche in MEDLINE – Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	MEDLINE	
Suchoberfläche	Ovid	
Datum der Suche	16.06.2020	
Zeitsegment	Ovid MEDLINE® 1946 to June Week 1 2020, Ovid MEDLINE® Epub Ahead of Print June 15, 2020, Ovid MEDLINE® In-Process & Other Non-Indexed Citations 1946 to June 15, 2020, Ovid MEDLINE® Daily Update June 15, 2020	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [55]	
#	**Suchbegriffe**	**Ergebnis**
1	trifaroten*.mp.	13
2	CD-5789.mp.	0
3	CD5789.mp.	1
4	CD 5789.mp.	0
5	aklief.mp.	3
6	selgamis.mp.	0
7	1 or 2 or 3 or 4 or 5 or 6	13
8	randomized.mp.	843.950
9	placebo.mp.	215.411
10	randomized controlled trial.pt.	507.515
11	8 or 9 or 10	907.307
12	7 and 11	1

Tabelle 4-71: Bibliografische Literaturrecherche in der Cochrane Library – Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	Cochrane Database	
Suchoberfläche	The Cochrane Library	
Datum der Suche	16.06.2020	
Zeitsegment	keine Einschränkung	
Suchfilter	keine Einschränkung	
#	**Suchbegriffe**	**Ergebnis**
1	trifaroten*	5
2	CD-5789	2
3	CD 5789	3
4	CD5789	11
5	aklief	0
6	selgamis	0
7	#1 or #2 or #3 or #4 or #5 or #6	13
8	Einschränkung: trials	12

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken 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 jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank 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/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-72: Suchstrategie der Recherche in Studienregistern – Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	ClinicalTrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	17.06.2020
Suchstrategie	trifarotene OR CD-5789 OR CD5789 OR "CD 5789" OR aklief OR selgamis [Other Terms]
Treffer	**7**
Studienregister	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/search
Datum der Suche	17.06.2020
Suchstrategie	trifarotene OR CD-5789 OR CD5789 OR "CD 5789" OR aklief OR selgamis
Treffer	**8**
Studienregister	International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal)
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	17.06.2020
Suchstrategie	trifarotene OR CD-5789 OR CD5789 OR CD 5789 OR aklief OR selgamis
Treffer	**18**

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

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

Es wurden keine Dokumente im Volltext gesichtet.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Entfällt, es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Entfällt, es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Entfällt, es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträgen 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 4-73: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – ClinicalTrials.gov

Nr.	Studie	Ausschlussgrund[(1)]
1.	Galderma (2012): RD.06.SPR18223 - A Randomized, Multi-center, Investigator-blind, Vehicle- and Active-controlled, Phase 2 Study to Assess the Efficacy and Safety of Different Concentrations of CD5789 Cream Applied Once Daily in Subjects With Moderate to Severe Acne Vulgaris - ClinicalTrials.gov (NCT01616654). Stand des Eintrags: 11.02.2014. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT01616654	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
2.	Galderma (2014): RD.06.SPR.18250 - A LONG-TERM SAFETY AND EFFICACY STUDY OF CD5789 (Trifarotene) 50 µg/g CREAM IN SUBJECTS WITH ACNE VULGARIS - ClinicalTrials.gov (NCT02189629). Stand des Eintrags: 14.11.2019. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT02189629	A5: Die Studie ist keine RCT.
3.	Galderma (2015): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - ClinicalTrials.gov (NCT02566369). Stand des Eintrags: 12.11.2019. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT02566369	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
4.	Galderma (2015): RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - ClinicalTrials.gov (NCT02556788). Stand des Eintrags: 12.11.2019. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT02556788	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
5.	Galderma (2019): RD.06.SPR.118295 - A Multi-Center Study to Evaluate Subject Reported Outcomes With Use of Trifarotene 50 µg/g Cream in the Treatment of Moderate Facial and Truncal Acne Vulgaris - ClinicalTrials.gov (NCT03915860). Stand des Eintrags:	A5: Die Studie ist keine RCT.

Nr.	Studie	Ausschlussgrund(1)
	11.05.2020. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT03915860	
6.	M.D. Anderson Cancer Center (2013): 2012-0710 NCI-2014-01377 (Registry Identifier: NCI CTRP) - Exploratory Study to Evaluate the Safety and Efficacy of CD5789 in Subjects With Early Stage Cutaneous T-Cell Lymphoma - ClinicalTrials.gov (NCT01804335). Stand des Eintrags: 13.04.2016. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT01804335	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
7.	Mayne Pharma (2018): 18-ICH-001 - A Phase 2 Randomized, Multicenter, Doubleblind, Vehicle Controlled, 12 Week, Safety, Efficacy & Systemic Exposure Study Followed by a 12 Week Open-label Extension of CD5789 in Adults and Adolescents With Autosomal Recessive Ichthyosis With Lamellar Scale - ClinicalTrials.gov (NCT03738800). Stand des Eintrags: 17.03.2020. [Zugriff: 17.06.2020]. URL: https://ClinicalTrials.gov/ct2/show/record/NCT03738800	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.

(1) Definition der Ausschlussgründe siehe Tabelle 4-2. Bei Zutreffen mehrerer Ausschlussgründe ist jeweils nur ein Ausschlussgrund angegeben.

Tabelle 4-74: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel – EU-CTR

Nr.	Studie	Ausschlussgrund(1)
1.	Galderma (2009): RD.03.SPR.40076E - Exploratory Study to Evaluate the Efficacy and Safety of Cd5789 in Subjects with Acne - EU-CTR (2008-007981-44). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-007981-44	A6: Studie ist kürzer als 12 Wochen.
2.	Galderma (2010): RD.03.SPR.40126 - Pilot Study to Explore Efficacy and Safety of Different Dose Regimens of Cd5789 in Subjects with Acne Vulgaris - EU-CTR (2010-022211-19). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-022211-19	A6: Studie ist kürzer als 12 Wochen.
3.	Galderma (2011): RD.03.SPR.40129 - Exploratory study to assess facial tolerability after daily application of several concentrations and formulations containing CD5789 in acne subjects - EU-CTR (2010-024214-54). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2010-024214-54	A6: Studie ist kürzer als 12 Wochen.
4.	Galderma (2013): RD.03.SPR.40214E - Efficacy, safety and pharmacokinetics of 2 concentrations and 2 dosage regimens of CD5789 in subjects with Lamellar Ichthyosis - EU-CTR (2013-001796-20). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-001796-20	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.

Nr.	Studie	Ausschlussgrund(1)
5.	Galderma (2014): RD.06.SPR.18250 - A LONG-TERM SAFETY AND EFFICACY STUDY OF CD5789 50µg/g CREAM IN SUBJECTS WITH ACNE VULGARIS - EU-CTR (2014-001755-23). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001755-23	A5: Die Studie ist keine RCT.
6.	Galderma (2015): RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - EU-CTR (2015-002540-13). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-002540-13	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
7.	Galderma (2016): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - EU-CTR (2016-002860-15). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2016-002860-15	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
8.	Mayne Pharma (2019): 18-ICH-001 - A Phase 2 Randomized, Multi-center, Double-blind, Vehicle-controlled, 12-Week, Safety, Efficacy, and Systemic Exposure Study followed by a 12-Week Open-label Extension of Trifarotene (CD5789) Cream HE1 in Adults and Adolescents with Autosomal Recessive Ichthyosis with Lamellar Scale - EU-CTR (2018-003272-12). [Zugriff: 17.06.2020]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-003272-12	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.

(1) Definition der Ausschlussgründe siehe Tabelle 4-2. Bei Zutreffen mehrerer Ausschlussgründe ist jeweils nur ein Ausschlussgrund angegeben.

Tabelle 4-75: Liste der ausgeschlossenen Studien bei der Suche nach RCT mit dem zu bewertenden Arzneimittel - ICTRP

Nr.	Studie	Ausschlussgrund(1)
1.	Galderma (2009): RD.03.SPR.40076E - Exploratory Study to Evaluate the Efficacy and Safety of CD5789 in Subjects with Acne - WHO-ICTRP (EUCTR2008-007981-44-FR). Stand des Eintrags: 01.03.2012. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-007981-44-FR	A6: Studie ist kürzer als 12 Wochen.
2.	Galderma (2010): RD.03.SPR.40126 - Pilot Study to Explore Efficacy and Safety of Different Dose Regimens of CD5789 in Subjects with Acne Vulgaris - WHO-ICTRP (EUCTR2010-022211-19-DE). Stand des Eintrags: 02.10.2012. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-022211-19-DE	A6: Studie ist kürzer als 12 Wochen.
3.	Galderma (2011): RD.03.SPR.40129 - Exploratory study to assess facial tolerability after daily application of several concentrations and formulations containing CD5789 in acne subjects - WHO-ICTRP (EUCTR2010-024214-54-BE). Stand des Eintrags: 02.08.2017.	A6: Studie ist kürzer als 12 Wochen.

Nr.	Studie	Ausschlussgrund[(1)]
	[Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2010-024214-54-BE	
4.	Galderma (2012): RD.06.SPR18223 - A Randomized, Multi-center, Investigator-blind, Vehicle- and Active-controlled, Phase 2 Study to Assess the Efficacy and Safety of Different Concentrations of CD5789 Cream Applied Once Daily in Subjects With Moderate to Severe Acne Vulgaris - WHO-ICTRP (NCT01616654). Stand des Eintrags: 01.02.2015. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01616654	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
5.	Galderma (2013): RD.03.SPR.40214E - Efficacy, safety and pharmacokinetics of 2 concentrations and 2 dosage regimens of CD5789 in subjects with Lamellar Ichthyosis - WHO-ICTRP (EUCTR2013-001796-20-SE). Stand des Eintrags: 01.03.2014. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-001796-20-SE	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
6.	Galderma (2014): RD.06.SPR.18250 - A LONG-TERM SAFETY AND EFFICACY STUDY OF CD5789 50µg/g CREAM IN SUBJECTS WITH ACNE VULGARIS - WHO-ICTRP (EUCTR2014-001755-23-HU). Stand des Eintrags: 03.04.2017. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001755-23-HU	A5: Die Studie ist keine RCT.
7.	Galderma (2014): RD.06.SPR.18250 - A LONG-TERM SAFETY AND EFFICACY STUDY OF CD5789 (Trifarotene) 50 µg/g CREAM IN SUBJECTS WITH ACNE VULGARIS - WHO-ICTRP (NCT02189629). Stand des Eintrags: 02.11.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02189629	A5: Die Studie ist keine RCT.
8.	Galderma (2015): 2015-002540-13-HU RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - Double-Blind Efficacy and Safety of CD5789 50µg/g Cream versus Vehicle Cream in Acne Vulgaris - WHO-ICTRP (EUCTR2015-002540-13-ES). Stand des Eintrags: 02.05.2017. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-002540-13-ES	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
9.	Galderma (2015): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - WHO-ICTRP (NCT02566369). Stand des Eintrags: 02.08.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02566369	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
10.	Galderma (2015): RD.06.SPR.18252 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 (Trifarotene) 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - WHO-ICTRP (NCT02556788). Stand des Eintrags: 09.09.2019. [Zugriff:	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.

Nr.	Studie	Ausschlussgrund[1]
	17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02556788	
11.	Galderma (2016): RD.06.SPR.18251 - A Multi-Center, Randomized, Double-Blind, Parallel-Group Vehicle Controlled Study To Compare The Efficacy And Safety Of CD5789 50µg/g Cream Versus Vehicle Cream In Subjects With Acne Vulgaris - Double-Blind Efficacy and Safety of CD5789 50µg/g Cream versus Vehicle cream in acne vulgaris - WHO-ICTRP (EUCTR2016-002860-15-DE). Stand des Eintrags: 01.12.2017. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2016-002860-15-DE	A3: Der Komparator entspricht nicht der zweckmäßigen Vergleichstherapie.
12.	Galderma (2019): RD.06.SPR.118295 - A Multi-Center Study to Evaluate Subject Reported Outcomes With Use of Trifarotene 50 µg/g Cream in the Treatment of Moderate Facial and Truncal Acne Vulgaris - WHO-ICTRP (NCT03915860). Stand des Eintrags: 02.05.2020. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03915860	A5: Die Studie ist keine RCT.
13.	M.D. Anderson Cancer Center (2013): 2012-0710 NCI-2014-01377 - Exploratory Study to Evaluate the Safety and Efficacy of CD5789 in Subjects With Early Stage Cutaneous T-Cell Lymphoma - WHO-ICTRP (NCT01804335). Stand des Eintrags: 02.04.2016. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01804335	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
14.	Mayne Pharma (2018): 18-ICH-001 - A Phase 2 Randomized, Multicenter, Doubleblind, Vehicle Controlled, 12 Week, Safety, Efficacy & Systemic Exposure Study Followed by a 12 Week Open-label Extension of CD5789 in Adults and Adolescents With Autosomal Recessive Ichthyosis With Lamellar Scale - WHO-ICTRP (NCT03738800). Stand des Eintrags: 03.03.2020. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03738800	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
15.	Mayne Pharma (2019): 18-ICH-001 2018-003272-12-FR NCT03738800 - A Phase 2 Randomized, Multi-center, Double-blind, Vehicle-controlled, 12-Week, Safety, Efficacy, and Systemic Exposure Study followed by a 12-Week Open-label Extension of Trifarotene (CD5789) Cream HE1 in Adults and Adolescents with Autosomal Recessive Ichthyosis with Lamellar Scale - WHO-ICTRP (EUCTR2018-003272-12-ES). Stand des Eintrags: 03.06.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-003272-12-ES	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
16.	Mayne Pharma (2019): 18-ICH-001 2018-003272-12-GB NCT03738800 - A Phase 2 Randomized, Multi-center, Double-blind, Vehicle-controlled, 12-Week, Safety, Efficacy, and Systemic Exposure Study followed by a 12-Week Open-label Extension of Trifarotene (CD5789) Cream HE1 in Adults and Adolescents with Autosomal Recessive Ichthyosis with Lamellar Scale - WHO-ICTRP (EUCTR2018-003272-12-DE). Stand des Eintrags: 02.11.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-003272-12-DE	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.
17.	Mayne Pharma (2019): 18-ICH-001 2018-003272-12-GB NCT03738800 - A Phase 2 Randomized, Multi-center, Double-blind,	A1: Die Patientenpopulation

Nr.	Studie	Ausschlussgrund(1)
	Vehicle-controlled, 12-Week, Safety, Efficacy, and Systemic Exposure Study followed by a 12-Week Open-label Extension of Trifarotene (CD5789) Cream HE1 in Adults and Adolescents with Autosomal Recessive Ichthyosis with Lamellar Scale - WHO-ICTRP (EUCTR2018-003272-12-FR). Stand des Eintrags: 01.12.2019. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-003272-12-FR	entspricht nicht dem zugelassenen Anwendungsgebiet.
18.	Mayne Pharma (2019): 18-ICH-001 NCT03738800 - A Phase 2 Randomized, Multi-center, Double-blind, Vehicle-controlled, 12-Week, Safety, Efficacy, and Systemic Exposure Study followed by a 12-Week Open-label Extension of Trifarotene (CD5789) Cream HE1 in Adults and Adolescents with Autosomal Recessive Ichthyosis with Lamellar Scale - WHO-ICTRP (EUCTR2018-003272-12-GB). Stand des Eintrags: 02.05.2020. [Zugriff: 17.06.2020]. URL: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-003272-12-GB	A1: Die Patientenpopulation entspricht nicht dem zugelassenen Anwendungsgebiet.

(1) Definition der Ausschlussgründe siehe Tabelle 4-2. Bei Zutreffen mehrerer Ausschlussgründe ist jeweils nur ein Ausschlussgrund angegeben.

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Entfällt, es wurde keine Suche nach RCT für indirekte Vergleiche durchgeführt.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Entfällt, es wurde keine Suche nach nicht randomisierten vergleichenden Studien durchgeführt.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Entfällt, es wurde keine Suche nach weiteren Untersuchungen durchgeführt.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-76 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 4-76 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-76 (Anhang): Studiendesign und -methodik für Studie PERFECT-1 (18251)

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Ziel der Studie war die Untersuchung der Wirksamkeit und Sicherheit von CD5789 (Trifaroten) 50 µg/g Creme bei täglicher Anwendung für 12 Wochen bei Patienten mit mittelschwerer *Acne vulgaris*. **Primäre Ziele:** • Bestimmung der Therapieerfolgsrate auf Basis der IGA-Skala (IGA 0 oder 1 und Verbesserung um mindestens zwei Punkte) unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl entzündlicher Läsionen im Gesicht unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung **Sekundäre Ziele:** • Bestimmung der Therapieerfolgsrate auf Basis der PGA-Skala (PGA 0 oder 1 und Verbesserung um mindestens zwei Punkte) unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl entzündlicher Läsionen am Rumpf unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung **Hypothese:** Die Behandlung mit Trifaroten ist der Behandlung mit Vehikel hinsichtlich der Wirksamkeit bei mittelschwerer *Acne vulgaris* überlegen.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, kontrollierte, doppelblinde, multizentrische Phase-III-Studie im Parallelgruppendesign, Randomisierung im Verhältnis 1:1 in die Behandlungsgruppen mit Trifaroten und Vehikel
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	**Protokoll-Amendment 1 (20. September 2016):** • Erhöhung der Anzahl der Studienzentren auf 120 und Streichung der geplanten Zahl an Patienten pro Zentrum • Aufnahme von europäischen Ländern als potentielle Orte der Studiendurchführung • Diverse weitere Ergänzungen zur Klarstellung und sprachliche Anpassungen

Item[a]	Charakteristikum	Studieninformation
		Protokoll-Amendment 2 (06. März 2017): • Erhöhung der Anzahl der Studienzentren auf 160 • Aufnahme von weiteren europäischen Ländern und Russland als potentielle Orte der Studiendurchführung
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** 1. Männliche oder weibliche Teilnehmer im Alter von 9 Jahren oder älter 2. Akne im Gesicht vom Schweregrad 3 (mittelschwer) auf der IGA-Skala zum Zeitpunkt des Screenings und zur Baseline-Visite 3. Mindestens 20 entzündliche und 25 nicht entzündliche Läsionen im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 4. Akne am Rumpf (Schultern, oberer Rücken und oberer Brustkorb – erreichbar zur Selbstanwendung der Studienmedikation durch den Teilnehmer) vom Schweregrad 3 (mittelschwer) auf der PGA-Skala zum Zeitpunkt des Screenings und zur Baseline-Visite; optionales Kriterium für Teilnehmer im Alter von 9 bis 11 Jahren 5. Mindestens 20 entzündliche und mindestens 20, aber nicht mehr als 100, nicht entzündliche Läsionen am Rumpf (Schultern, oberer Rücken und oberer Brustkorb – erreichbar zur Selbstanwendung der Studienmedikation durch den Teilnehmer) zum Zeitpunkt des Screenings und zur Baseline-Visite, optionales Kriterium für Teilnehmer im Alter von 9 bis 11 Jahren 6. Frauen außerhalb des gebärfähigen Alters 7. Frauen im gebärfähigen Alter: a. mit negativem Urin-Schwangerschaftstest zum Zeitpunkt des Screenings und bei Behandlungsbeginn b. strikt abstinent im Zeitraum von einem Monat vor Behandlungsbeginn, während der Studie und mindestens einen Monat nach der letzten Anwendung der Studienmedikation <u>oder</u> Anwendung einer geeigneten Verhütungsmethode für die Dauer der Studie bis mindestens einen Monat nach der letzten Anwendung der Studienmedikation 8. Frauen im gebärfähigen Alter mit oralen Kontrazeptiva, die zur Akne-Behandlung zugelassen sind: stabile Dosierung für mindestens 6 Monate vor der Baseline-Visite 9. Mädchen, bei denen im Studienverlauf erstmalig die Menstruation einsetzt:

Item[a]	Charakteristikum	Studieninformation
		a. strikt abstinent im Zeitraum von einem Monat vor Behandlungsbeginn, während der Studie und mindestens einen Monat nach der letzten Anwendung der Studienmedikation oder Anwendung einer geeigneten Verhütungsmethode für die Dauer der Studie bis mindestens einen Monat nach der letzten Anwendung der Studienmedikation 10. Teilnehmer, die gewillt und in der Lage sind, die Anforderungen des Studienprotokolls zu erfüllen 11. Vorliegen einer unterschriebenen Einverständniserklärung; bei minderjährigen Teilnehmern Vorliegen einer durch die Eltern oder Erziehungsberechtigten unterschriebenen Einverständniserklärung und Unterzeichnung einer Willenserklärung durch den Patienten **Ausschlusskriterien:** 1. Schwere Akne-Formen (z. B. *Acne conglobata*, *Acne fulminans*) oder sekundäre Akne-Formen (z. B. Chlor-Akne, medikamenteninduzierte Akne) 2. Mehr als ein Knoten im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 3. Mehr als ein Knoten am Rumpf zum Zeitpunkt des Screenings und zur Baseline-Visite 4. Vorliegen von einer oder mehreren Akne-Zysten im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 5. Vorliegen von einer oder mehreren Akne-Zysten am Rumpf zum Zeitpunkt des Screenings und zur Baseline-Visite 6. Bart oder Gesichtsbehaarung, die die Untersuchungen in der Studie beeinträchtigt 7. Tattoos am Rumpf, die die Untersuchungen in der Studie beeinträchtigen 8. Vorliegen einer unkontrollierten oder schwerwiegenden Erkrankung oder eines medizinischen Zustandes, welche die Interpretation der Studienergebnisse beeinflussen oder den Patienten bei Studienteilnahme gefährden 9. Klinisch relevante Laborwerte außerhalb des Normbereichs 10. Bekannte oder vermutete Allergien gegen oder Empfindlichkeit für einen Bestandteil der Studienmedikation 11. Stillende Frauen 12. Kinderwunsch im Zeitraum der Studie 13. Aktuelle Teilnahme an einer anderen Studie oder Teilnahme innerhalb von 30 Tagen vor Baseline-Visite 14. Teilnahme an der Langzeit-Studie RD.03.SPR.18250 15. Einnahme oder Anwendung folgender Wirkstoffe:

Item[a]	Charakteristikum	Studieninformation
		- Topische Anwendung von Antibiotika, Benzoylperoxid, Azelainsäure, Zink und Kortikosteroiden innerhalb von 2 Wochen vor Baseline-Visite oder Retinoiden innerhalb von 4 Wochen vor Baseline-Visite - Phototherapie und Anwendung anderer Verfahren im Gesicht innerhalb von einer Woche vor Baseline-Visite - Systemische Anwendung von NSAID innerhalb von 2 Wochen vor Baseline-Visite, Kortikosteroiden (außer inhalative oder intrathekale Kortikosteroide) und Antibiotika (außer Penicillin) innerhalb von 4 Wochen vor Baseline-Visite, Retinoide und Immunmodulatoren innerhalb von 6 Monaten vor Baseline-Visite 16. Exzessive UV-Strahlung innerhalb eines Monats vor Baseline-Visite oder während der Studie 17. Keine Bereitschaft, auf die Anwendung von Präparaten mit α-Hydroxycarbonsäuren, medizinische Rasiercremes, Aftershaves, antibakterielle Seifen oder nicht zur Behandlung der Akne zugelassenen Präparaten im Zeitraum der Studie zu verzichten 18. Vulnerable Personen gemäß Abschnitt 1.61 ICH-GCP
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie in insgesamt 119 Zentren in 5 Ländern: Deutschland, Kanada, Puerto Rico, Ungarn, USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	**Trifaroten-Arm:** • Creme mit 50 µg/g Trifaroten einmal täglich für 12 Wochen **Vehikel-Arm:** • Vehikel-Creme einmal täglich für 12 Wochen Alle Teilnehmer erhielten mündliche und schriftliche Anweisungen zur richtigen Dosierung und Anwendung der Studienmedikation. Die Studienmedikation wurde von den Teilnehmern wie folgt angewendet: • Anwendung abends nach Reinigung der Haut • Anwendung im gesamten Gesicht und am gesamten oberen Rumpf, auch in Bereichen ohne Läsionen • Dosierung: ein Pumpenhub für das Gesicht und zwei Pumpenhübe für den Rumpf Die Teilnehmer wurden ermutigt, eine nicht komedogene und hypoallergene Feuchtigkeitscreme eine Stunde vor und nach der Studienmedikation anzuwenden, um Hautirritationen zu verringern.
6	Zielkriterien	

Item[a]	Charakteristikum	Studieninformation
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	Es fand ein Training der Prüfer hinsichtlich der Erhebung des Akne-Schweregrads im Gesicht (*Investigator´s Global Assessment*, IGA) und am Rumpf (*Physician´s Global Assessment*, PGA) sowie der Bestimmung der Zahl der Läsionen statt. Die Auswertung sollte für jeden Patienten wenn möglich durch den gleichen Prüfarzt erfolgen. **Primäre Endpunkte:** • Therapieerfolgsrate in Woche 12 (definiert als Anteil der Patienten mit IGA 0 oder 1 und einer Verbesserung um mindestens 2 Punkte im Vergleich zur Baseline-Visite) • Absolute Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung • Absolute Änderung der Anzahl entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung **Sekundäre Endpunkte:** • Anteil der Patienten mit PGA 0 oder 1 in Woche 12 und einer Verbesserung um mindestens 2 Punkte im Vergleich zur Baseline-Visite • Absolute Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung • Absolute Änderung der Anzahl entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung **Supportive Wirksamkeits-Endpunkte:** • Prozentuale Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung • Prozentuale Änderung der Anzahl entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung • Prozentuale Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung • Prozentuale Änderung der Anzahl entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung • Selbsteinschätzung der Patienten zur Verbesserung der Akne im Gesicht nach 12 Wochen Behandlung **Sicherheitsendpunkte:** • Unerwünschte Ereignisse • Lokale Verträglichkeit im Gesicht • Lokale Verträglichkeit am Rumpf • Laborwertänderungen **Weitere Endpunkte:** • *Dermatology Life Quality Index* (DLQI) / *Children´s Dermatology Life Quality Index* (CDLQI) nach 12 Wochen Behandlung • *Skin Oiliness Scale Questionnaire* nach 12 Wochen Behandlung

Item[a]	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	**Protokoll-Amendment 1 (20. September 2016):** • Aufnahme des *Skin Oiliness Scale Questionnaire* als explorativer Endpunkt **SAP-Addendum (23. Oktober 2017):** • keine Änderungen der geplanten Zielkriterien • Änderungen bei den geplanten Subgruppenanalysen (neue zusätzliche Alterskategorisierung [< 14, 14 - 17, ≥ 18 Jahre], Wegfall der Subgruppenanalyse nach Land bzw. Region, Subgruppenauswertungen für sekundäre Endpunkte) • Spezifizierung der statistischen Methoden
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlbestimmung basierte auf der Phase-II-Studie SPR18223. Der Anteil der Patienten mit Behandlungserfolg gemäß IGA-Skala wurde mit 23 % (Trifaroten) bzw. 15 % (Vehikel) angenommen. Um mit 90%iger Power in einem zweiseitigen Signifikanztest mit einem Typ-I-Fehler von 0,05 einen signifikanten Unterschied zwischen Trifaroten und der Vehikel-Creme nachzuweisen, waren 504 Teilnehmer pro Behandlungsarm erforderlich. Um Drop-outs zu kompensieren, wurde die Fallzahl auf insgesamt 1200 Teilnehmer (600 pro Behandlungsarm) erhöht. Auf Basis der Phase-II-Studie wurde ein mittlerer Unterschied zwischen den Behandlungsarmen von 3 entzündlichen bzw. 6 nicht entzündlichen Läsionen angenommen, mit einer Standardabweichung von 12 bzw. 20 Läsionen. Basierend auf diesen Annahmen und unter Zugrundelegung einer Fallzahl von 504 Teilnehmern pro Behandlungsarm wird eine über 95%ige Power für beide Auswertungen erreicht. Für die Endpunkte zur Akne am Rumpf wurden keine abweichenden Annahmen getroffen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es war keine Interimsanalyse geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung erfolgte durch ein interaktives System (IRT, *Interactive Response Technology*).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockrandomisierung im Verhältnis 1:1 in die parallelen Behandlungsgruppen mit Stratifizierung nach Zentrum (Aufgrund einer fehlerhaften Implementierung der Stratifizierung im IRT-System wurde ein Teil der Patienten randomisiert, ohne die Stratifizierung nach Zentrum zu berücksichtigen. Der Fehler wurde am 22.12.2016 behoben und alle nachfolgend rekrutierten Teilnehmer wurden stratifiziert randomisiert. Die Zuteilung der Teilnehmer zu den

Item[a]	Charakteristikum	Studieninformation
		beiden Behandlungsgruppen wurde durch eine externe Person unverblindet überprüft. Es wurden keine relevanten Imbalancen festgestellt. Die Randomisierung wurde für alle Patienten als valide erachtet.)
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	Zuteilung per Telefon/Internet mittels IRT. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet. Nur das für die Verpackung und Kennzeichnung der Studienmedikation zuständige Personal hatte Zugang zur Randomisierungsliste.
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 Randomisierungsliste wurde durch einen Statistiker von Galderma erstellt. Die Zuteilung der Teilnehmer zur Randomisierungsliste erfolgte per Telefon/Internet mittels IRT.
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?	a) ja b) ja c) ja Die Verblindung erfolgte durch die Zuteilung der Behandlung mittels IRT.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Cremes mit Trifaroten und Vehikel waren in identischen Pumpflaschen abgefüllt.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	**Populationen zur Auswertung der Wirksamkeitsendpunkte:** Alle Wirksamkeitsendpunkte zur Akne im Gesicht wurden auf Basis der ITT(*intention to treat*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten. Die Patienten wurden ausgewertet wie randomisiert. Alle Wirksamkeitsendpunkte zur Akne am Rumpf wurden auf Basis der ITTT(*intention to treat on the trunk*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten mit Akne am Rumpf bei der Baseline-Visite. Die Patienten wurden ausgewertet wie randomisiert. **Populationen zur Auswertung der Sicherheitsendpunkte:** Alle Sicherheitsendpunkte mit Ausnahme der lokalen Verträglichkeit am Rumpf wurden auf Basis der SAF(*safety*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten, die mindestens einmal die Studienmedikation

Item[a]	Charakteristikum	Studieninformation
		angewendet haben. Die Patienten wurden ausgewertet wie behandelt. Die lokale Verträglichkeit am Rumpf wurde auf Basis der SAFT(*safety on the trunk*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten, die mindestens einmal die Studienmedikation am Rumpf angewendet haben. Die Patienten wurden ausgewertet wie behandelt. **Statistische Auswertung der primären Endpunkte:** <u>Therapieerfolgsrate (IGA 0/1)</u> Die Auswertung erfolgte für die ITT-Population mittels Cochran-Mantel-Haenszel(CMH)-Test stratifiziert nach Zentrum. Fehlende Werte wurde mittels multipler Imputation unter MAR-Annahme (*missing at random*) ersetzt. Der p-Wert basiert auf dem stratifizierten CMH-Test. Der Gruppenunterschied sowie das 95 %-Konfidenzintervall basiert auf einem Approximationsverfahren für binäre Daten innerhalb einer großen Stichprobe. <u>Absolute Anzahl der Läsionen</u> Die Auswertung erfolgte für die ITT-Population mittels eines *Analysis-of-Covariance*(ANCOVA)-Modells mit dem Baseline-Wert, Zentrum und Behandlung als Effekte. Fehlende Werte wurde mittels multipler Imputation unter MAR-Annahme ersetzt. Aus dem ANCOVA-Modell ergeben sich p-Werte für den Behandlungsunterschied sowie Schätzer für die *Least Square Means Difference* (LSMD) und die zugehörigen 95 %-Konfidenzintervalle. **Statistische Auswertung der sekundären Endpunkte:** Die sekundären Endpunkte wurden auf Basis der ITTT-Population mit den gleichen Methodiken wie die primären Endpunkte ausgewertet. **Statistische Auswertung der weiteren Endpunkte:** Die supportiven Wirksamkeitsendpunkte zur relativen Änderung der Läsionen wurden mittels eines CMH-Tests mit *Ridit-Scoring*, stratifiziert nach Zentrum, basierend auf der ITT-Population ausgewertet. Fehlende Werte wurde mittels multipler Imputation unter MAR-Annahme ersetzt. Die Selbsteinschätzung der Patienten wurde für die ITT-Population auf Basis der beobachteten Werte ausgewertet. Der *Skin Oiliness Scale Questionnaire* sowie DLQI und CDLQI wurden deskriptiv auf Basis der beobachteten Werte ausgewertet. Sicherheitsendpunkte wurden deskriptiv ausgewertet.

Item[a]	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	**Sensitivitätsanalysen:** <u>Primäre Endpunkte</u> Die primären Endpunkte wurden auf Basis der PP(*Per-Protocol)*-Population analog zur Hauptanalyse ausgewertet. Die PP-Population umfasst alle Patienten der ITT-Population, bei denen keine wesentliche Protokollverletzung vorlag. Des Weiteren wurden die primären Endpunkte auf Basis der ITT-Population mit alternativen Ersetzungsstrategien ausgewertet: • Multiple Imputation basierend auf dem *Pattern-Mixture-Model* unter MNAR-Annahme (*missing not at random*) • LOCF (*last observation carried forward*) • Patienten mit fehlenden IGA-Werten in Woche 12 wurden als Non-Responder gewertet; fehlende Werte für die Anzahl der Läsionen in Woche 12 wurden mit der mittleren Anzahl der Läsionen bei Patienten ohne Therapieerfolg (IGA 0/1) im selben Behandlungsarm ersetzt <u>Sekundäre Endpunkte</u> Für die sekundären Endpunkte wurden Sensitivitätsanalysen analog zu den primären Endpunkten, jeweils auf Basis der PPT- und ITTT-Population, durchgeführt. **Subgruppenanalysen:** Subgruppenanalysen wurden explorativ durchgeführt. Für die primären und sekundären Endpunkte wurde der Behandlungseffekt und das zugehörige 95 %-Konfidenzintervall für die folgenden Subgruppen ausgewertet: • Alterskategorien (9-11, 12-17, ≥ 18 Jahre) • Geschlecht (männlich, weiblich) • Abstammung (weiß, schwarz/afroamerikanisch, asiatisch, andere) • Ethnie (hispanisch, nicht hispanisch) • Hauttyp (I-III, IV-VI)
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) Total: 1208; Trifaroten: 612; Vehikel: 596 b) Total: 1201; Trifaroten: 610; Vehikel: 591 (Alle Patienten erhielten eine Behandlung. Aufgrund einer fälschlichen Zuteilung der Studienmedikation erhielten insgesamt 7 Patienten nicht die ihnen zugeteilte Behandlung gemäß Randomisierung: 5 Patienten im Vehikel-Arm erhielten die Creme mit Trifaroten und 2 Patienten im Interventionsarm erhielten die Vehikel-Creme.) c) Total: 1208; Trifaroten: 612; Vehikel: 596
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Es wurden 1524 Patienten gescreent. Davon erfüllten 316 die Einschlusskriterien nicht, und 1208 Patienten wurden randomisiert. Alle 1208 Patienten erhielten mindestens einmal die Studienmedikation. Aufgrund einer fälschlichen Zuteilung der Studienmedikation erhielten insgesamt 7 Patienten nicht die ihnen zugeteilte Behandlung gemäß Randomisierung: 5 Patienten im Vehikel-Arm erhielten die Creme mit Trifaroten und 2 Patienten im Interventionsarm erhielten die Vehikel-Creme. Die Anzahl der Patienten, die die Behandlung vorzeitig abgebrochen haben, und die Gründe für diese Therapieabbrüche sind im nachfolgenden Flussdiagramm angegeben.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn: 30. November 2015 (Screening des ersten Patienten) Studienende: 17. November 2017 (letzte Visite des letzten Patienten)
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete planmäßig.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

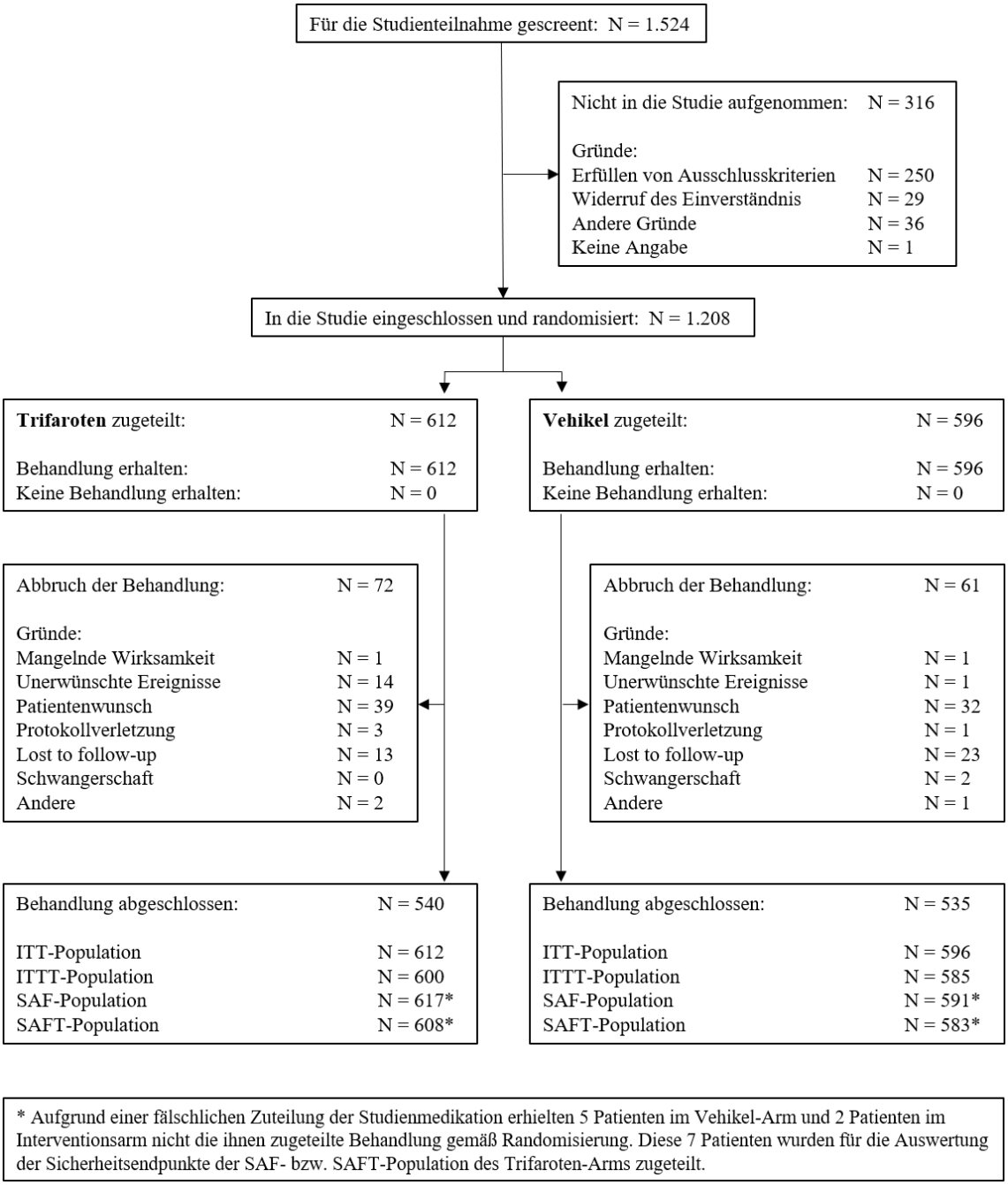


Abbildung 4-24: Patientenfluss der Studie PERFECT-1

Tabelle 4-77 (Anhang): Studiendesign und -methodik für Studie PERFECT-2 (18252)

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Ziel der Studie war die Untersuchung der Wirksamkeit und Sicherheit von CD5789 (Trifaroten) 50 µg/g Creme bei täglicher Anwendung für 12 Wochen bei Patienten mit mittelschwerer *Acne vulgaris*. **Primäre Ziele:** • Bestimmung der Therapieerfolgsrate auf Basis der IGA-Skala (IGA 0 oder 1 und Verbesserung um mindestens zwei Punkte) unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl entzündlicher Läsionen im Gesicht unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung **Sekundäre Ziele:** • Bestimmung der Therapieerfolgsrate auf Basis der PGA-Skala (PGA 0 oder 1 und Verbesserung um mindestens zwei Punkte) unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung • Bestimmung der absoluten Änderung der Anzahl entzündlicher Läsionen am Rumpf unter Trifaroten im Vergleich zum Vehikel nach 12 Wochen Behandlung **Hypothese:** Die Behandlung mit Trifaroten ist der Behandlung mit Vehikel hinsichtlich der Wirksamkeit bei mittelschwerer *Acne vulgaris* überlegen.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, kontrollierte, doppelblinde, multizentrische Phase-III-Studie im Parallelgruppendesign, Randomisierung im Verhältnis 1:1 in die Behandlungsgruppen mit Trifaroten und Vehikel
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	**Protokoll-Amendment 1 (19. Januar 2017):** • Erhöhung der Anzahl der Studienzentren auf 85 und Streichung der geplanten Zahl an Patienten pro Zentrum • Diverse weitere Ergänzungen zur Klarstellung und sprachliche Anpassungen
4	Probanden / Patienten	

Item[a]	Charakteristikum	Studieninformation
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien:** 1. Männliche oder weibliche Teilnehmer im Alter von 9 Jahren oder älter 2. Akne im Gesicht vom Schweregrad 3 (mittelschwer) auf der IGA-Skala zum Zeitpunkt des Screenings und zur Baseline-Visite 3. Mindestens 20 entzündliche und 25 nicht entzündliche Läsionen im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 4. Akne am Rumpf (Schultern, oberer Rücken und oberer Brustkorb – erreichbar zur Selbstanwendung der Studienmedikation durch den Teilnehmer) vom Schweregrad 3 (mittelschwer) auf der PGA-Skala zum Zeitpunkt des Screenings und zur Baseline-Visite; optionales Kriterium für Teilnehmer im Alter von 9 bis 11 Jahren 5. Mindestens 20 entzündliche und mindestens 20, aber nicht mehr als 100, nicht entzündliche Läsionen am Rumpf (Schultern, oberer Rücken und oberer Brustkorb – erreichbar zur Selbstanwendung der Studienmedikation durch den Teilnehmer) zum Zeitpunkt des Screenings und zur Baseline-Visite, optionales Kriterium für Teilnehmer im Alter von 9 bis 11 Jahren 6. Frauen außerhalb des gebärfähigen Alters 7. Frauen im gebärfähigen Alter: a. mit negativem Urin-Schwangerschaftstest zum Zeitpunkt des Screenings und bei Behandlungsbeginn b. strikt abstinent im Zeitraum von einem Monat vor Behandlungsbeginn, während der Studie und mindestens einen Monat nach der letzten Anwendung der Studienmedikation oder Anwendung einer geeigneten Verhütungsmethode für die Dauer der Studie bis mindestens einen Monat nach der letzten Anwendung der Studienmedikation 8. Frauen im gebärfähigen Alter mit oralen Kontrazeptiva, die zur Akne-Behandlung zugelassen sind: stabile Dosierung für mindestens 6 Monate vor der Baseline-Visite 9. Mädchen, bei denen im Studienverlauf erstmalig die Menstruation einsetzt: a. strikt abstinent im Zeitraum von einem Monat vor Behandlungsbeginn, während der Studie und mindestens einen Monat nach der letzten Anwendung der Studienmedikation oder Anwendung einer geeigneten Verhütungsmethode für die Dauer der Studie bis mindestens einen Monat nach der letzten Anwendung der Studienmedikation

Item[a]	Charakteristikum	Studieninformation
		10. Teilnehmer, die gewillt und in der Lage sind, die Anforderungen des Studienprotokolls zu erfüllen 11. Vorliegen einer unterschriebenen Einverständniserklärung; bei minderjährigen Teilnehmern Vorliegen einer durch die Eltern oder Erziehungsberechtigten unterschriebenen Einverständniserklärung und Unterzeichnung einer Willenserklärung durch den Patienten **Ausschlusskriterien:** 1. Schwere Akne-Formen (z. B. *Acne conglobata*, *Acne fulminans*) oder sekundäre Akne-Formen (z. B. Chlor-Akne, medikamenteninduzierte Akne) 2. Mehr als ein Knoten im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 3. Mehr als ein Knoten am Rumpf zum Zeitpunkt des Screenings und zur Baseline-Visite 4. Vorliegen von einer oder mehreren Akne-Zysten im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 5. Vorliegen von einer oder mehreren Akne-Zysten am Rumpf zum Zeitpunkt des Screenings und zur Baseline-Visite 6. Bart oder Gesichtsbehaarung, die die Untersuchungen in der Studie beeinträchtigt 7. Tattoos am Rumpf, die die Untersuchungen in der Studie beeinträchtigen 8. Vorliegen einer unkontrollierten oder schwerwiegenden Erkrankung oder eines medizinischen Zustandes, welche die Interpretation der Studienergebnisse beeinflussen oder den Patienten bei Studienteilnahme gefährden 9. Klinisch relevante Laborwerte außerhalb des Normbereichs 10. Bekannte oder vermutete Allergien gegen oder Empfindlichkeit für einen Bestandteil der Studienmedikation 11. Stillende Frauen 12. Kinderwunsch im Zeitraum der Studie 13. Aktuelle Teilnahme an einer anderen Studie oder Teilnahme innerhalb von 30 Tagen vor Baseline-Visite 14. Teilnahme an der Langzeit-Studie RD.03.SPR.18250 15. Einnahme oder Anwendung folgender Wirkstoffe: - Topische Anwendung von Antibiotika, Benzoylperoxid, Azelainsäure, Zink und Kortikosteroiden innerhalb von 2 Wochen vor Baseline-Visite oder Retinoiden innerhalb von 4 Wochen vor Baseline-Visite - Phototherapie und Anwendung anderer Verfahren im Gesicht innerhalb von einer Woche vor Baseline-Visite

Item[a]	Charakteristikum	Studieninformation
		- Systemische Anwendung von NSAID innerhalb von 2 Wochen vor Baseline-Visite, Kortikosteroiden (außer inhalative oder intrathekale Kortikosteroide) und Antibiotika (außer Penicillin) innerhalb von 4 Wochen vor Baseline-Visite, Retinoide und Immunmodulatoren innerhalb von 6 Monaten vor Baseline-Visite 16. Exzessive UV-Strahlung innerhalb eines Monats vor Baseline-Visite oder während der Studie 17. Keine Bereitschaft, auf die Anwendung von Präparaten mit α-Hydroxycarbonsäuren, medizinische Rasiercremes, Aftershaves, antibakterielle Seifen oder nicht zur Behandlung der Akne zugelassenen Präparaten im Zeitraum der Studie zu verzichten 18. Vulnerable Personen gemäß Abschnitt 1.61 ICH-GCP
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrische Studie in insgesamt 82 Zentren in 8 Ländern: Polen, Rumänien, Russland, Spanien, Tschechien, Ukraine, Ungarn, USA
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	**Trifaroten-Arm:** • Creme mit 50 µg/g Trifaroten einmal täglich für 12 Wochen **Vehikel-Arm:** • Vehikel-Creme einmal täglich für 12 Wochen Alle Teilnehmer erhielten mündliche und schriftliche Anweisungen zur richtigen Dosierung und Anwendung der Studienmedikation. Die Studienmedikation wurde von den Teilnehmern wie folgt angewendet: • Anwendung abends nach Reinigung der Haut • Anwendung im gesamten Gesicht und am gesamten oberen Rumpf, auch in Bereichen ohne Läsionen • Dosierung: ein Pumpenhub für das Gesicht und zwei Pumpenhübe für den Rumpf Die Teilnehmer wurden ermutigt, eine nicht komedogene und hypoallergene Feuchtigkeitscreme eine Stunde vor und nach der Studienmedikation anzuwenden, um Hautirritationen zu verringern.
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	Es fand ein Training der Prüfer hinsichtlich der Erhebung des Akne-Schweregrads im Gesicht (*Investigator´s Global Assessment*, IGA) und am Rumpf (*Physician´s Global Assessment*, PGA) sowie der Bestimmung der Zahl der Läsionen statt. Die Auswertung sollte für jeden Patienten wenn möglich durch den gleichen Prüfarzt erfolgen. **Primäre Endpunkte:**

Item[a]	Charakteristikum	Studieninformation
		• Therapieerfolgsrate in Woche 12 (definiert als Anteil der Patienten mit IGA 0 oder 1 und einer Verbesserung um mindestens 2 Punkte im Vergleich zur Baseline-Visite) • Absolute Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung • Absolute Änderung der Anzahl entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung **Sekundäre Endpunkte:** • Anteil der Patienten mit PGA 0 oder 1 in Woche 12 und einer Verbesserung um mindestens 2 Punkte im Vergleich zur Baseline-Visite • Absolute Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung • Absolute Änderung der Anzahl entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung **Supportive Wirksamkeits-Endpunkte:** • Prozentuale Änderung der Anzahl nicht entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung • Prozentuale Änderung der Anzahl entzündlicher Läsionen im Gesicht nach 12 Wochen Behandlung • Prozentuale Änderung der Anzahl nicht entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung • Prozentuale Änderung der Anzahl entzündlicher Läsionen am Rumpf nach 12 Wochen Behandlung • Selbsteinschätzung der Patienten zur Verbesserung der Akne im Gesicht nach 12 Wochen Behandlung **Sicherheitsendpunkte:** • Unerwünschte Ereignisse • Lokale Verträglichkeit im Gesicht • Lokale Verträglichkeit am Rumpf • Laborwertänderungen **Weitere Endpunkte:** • *Dermatology Life Quality Index* (DLQI) / *Children´s Dermatology Life Quality Index* (CDLQI) nach 12 Wochen Behandlung
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	**SAP-Addendum (23. Oktober 2017):** • keine Änderungen der geplanten Zielkriterien • Änderungen bei den geplanten Subgruppenanalysen (neue zusätzliche Alterskategorisierung [< 14, 14 - 17, ≥ 18 Jahre], Wegfall der Subgruppenanalyse nach Land bzw. Region, Subgruppenauswertungen für sekundäre Endpunkte) • Spezifizierung der statistischen Methoden
7	Fallzahl	

Item[a]	Charakteristikum	Studieninformation
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlbestimmung basierte auf der Phase-II-Studie SPR18223. Der Anteil der Patienten mit Behandlungserfolg gemäß IGA-Skala wurde mit 23 % (Trifaroten) bzw. 15 % (Vehikel) angenommen. Um mit 90%iger Power in einem zweiseitigen Signifikanztest mit einem Typ-I-Fehler von 0,05 einen signifikanten Unterschied zwischen Trifaroten und der Vehikel-Creme nachzuweisen, waren 504 Teilnehmer pro Behandlungsarm erforderlich. Um Drop-outs zu kompensieren, wurde die Fallzahl auf insgesamt 1200 Teilnehmer (600 pro Behandlungsarm) erhöht. Auf Basis der Phase-II-Studie wurde ein mittlerer Unterschied zwischen den Behandlungsarmen von 3 entzündlichen bzw. 6 nicht entzündlichen Läsionen angenommen, mit einer Standardabweichung von 12 bzw. 20 Läsionen. Basierend auf diesen Annahmen und unter Zugrundelegung einer Fallzahl von 504 Teilnehmern pro Behandlungsarm wird eine über 95%ige Power für beide Auswertungen erreicht. Für die Endpunkte zur Akne am Rumpf wurden keine abweichenden Annahmen getroffen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es war keine Interimsanalyse geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung erfolgte durch ein interaktives System (IRT, *Interactive Response Technology*).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockrandomisierung im Verhältnis 1:1 in die parallelen Behandlungsgruppen mit Stratifizierung nach Zentrum (Aufgrund einer fehlerhaften Implementierung der Stratifizierung im IRT-System wurde ein Teil der Patienten randomisiert, ohne die Stratifizierung nach Zentrum zu berücksichtigen. Der Fehler wurde am 22.12.2016 behoben und alle nachfolgend rekrutierten Teilnehmer wurden stratifiziert randomisiert. Die Zuteilung der Teilnehmer zu den beiden Behandlungsgruppen wurde durch eine externe Person unverblindet überprüft. Es wurden keine relevanten Imbalancen festgestellt. Die Randomisierung wurde für alle Patienten als valide erachtet.)
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	Zuteilung per Telefon/Internet mittels IRT. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet. Nur das für die Verpackung und Kennzeichnung der Studienmedikation zuständige Personal hatte Zugang zur Randomisierungsliste.

Item[a]	Charakteristikum	Studieninformation
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 Randomisierungsliste wurde durch einen Statistiker von Galderma erstellt. Die Zuteilung der Teilnehmer zur Randomisierungsliste erfolgte per Telefon/Internet mittels IRT.
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?	a) ja b) ja c) ja Die Verblindung erfolgte durch die Zuteilung der Behandlung mittels IRT.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Cremes mit Trifaroten und Vehikel waren in identischen Pumpflaschen abgefüllt.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	**Populationen zur Auswertung der Wirksamkeitsendpunkte:** Alle Wirksamkeitsendpunkte zur Akne im Gesicht wurden auf Basis der ITT(*intention to treat*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten. Die Patienten wurden ausgewertet wie randomisiert. Alle Wirksamkeitsendpunkte zur Akne am Rumpf wurden auf Basis der ITTT(*intention to treat on the trunk*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten mit Akne am Rumpf bei der Baseline-Visite. Die Patienten wurden ausgewertet wie randomisiert. **Populationen zur Auswertung der Sicherheitsendpunkte:** Alle Sicherheitsendpunkte mit Ausnahme der lokalen Verträglichkeit am Rumpf wurden auf Basis der SAF(*safety*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten, die mindestens einmal die Studienmedikation angewendet haben. Die Patienten wurden ausgewertet wie behandelt. Die lokale Verträglichkeit am Rumpf wurde auf Basis der SAFT(*safety on the trunk*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten, die mindestens einmal die Studienmedikation am Rumpf angewendet haben. Die Patienten wurden ausgewertet wie behandelt. **Statistische Auswertung der primären Endpunkte:** <u>Therapieerfolgsrate (IGA 0/1)</u> Die Auswertung erfolgte für die ITT-Population mittels Cochran-Mantel-Haenszel(CMH)-Test stratifiziert nach

Item[a]	Charakteristikum	Studieninformation
		Zentrum. Fehlende Werte wurde mittels multipler Imputation unter MAR-Annahme (*missing at random*) ersetzt. Der p-Wert basiert auf dem stratifizierten CMH-Test. Der Gruppenunterschied sowie das 95 %-Konfidenzintervall basiert auf einem Approximationsverfahren für binäre Daten innerhalb einer großen Stichprobe. <u>Absolute Anzahl der Läsionen</u> Die Auswertung erfolgte für die ITT-Population mittels eines *Analysis-of-Covariance*(ANCOVA)-Modells mit dem Baseline-Wert, Zentrum und Behandlung als Effekte. Fehlende Werte wurde mittels multipler Imputation unter MAR-Annahme ersetzt. Aus dem ANCOVA-Modell ergeben sich p-Werte für den Behandlungsunterschied sowie Schätzer für die *Least Square Means Difference* (LSMD) und die zugehörigen 95 %-Konfidenzintervalle. **Statistische Auswertung der sekundären Endpunkte:** Die sekundären Endpunkte wurden auf Basis der ITTT-Population mit den gleichen Methodiken wie die primären Endpunkte ausgewertet. **Statistische Auswertung der weiteren Endpunkte:** Die supportiven Wirksamkeitsendpunkte zur relativen Änderung der Läsionen wurden mittels eines CMH-Tests mit *Ridit-Scoring*, stratifiziert nach Zentrum, basierend auf der ITT-Population ausgewertet. Fehlende Werte wurde mittels multipler Imputation unter MAR-Annahme ersetzt. Die Selbsteinschätzung der Patienten wurde für die ITT-Population auf Basis der beobachteten Werte ausgewertet. Der DLQI und CDLQI wurden deskriptiv auf Basis der beobachteten Werte ausgewertet. Sicherheitsendpunkte wurden deskriptiv ausgewertet.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	**Sensitivitätsanalysen:** <u>Primäre Endpunkte</u> Die primären Endpunkte wurden auf Basis der PP(*Per-Protocol)*-Population analog zur Hauptanalyse ausgewertet. Die PP-Population umfasst alle Patienten der ITT-Population, bei denen keine wesentliche Protokollverletzung vorlag. Des Weiteren wurden die primären Endpunkte auf Basis der ITT-Population mit alternativen Ersetzungsstrategien ausgewertet: • Multiple Imputation basierend auf dem *Pattern-Mixture-Model* unter MNAR-Annahme (*missing not at random*) • LOCF (*last observation carried forward*) • Patienten mit fehlenden IGA-Werten in Woche 12 wurden als Non-Responder gewertet; fehlende Werte für die Anzahl der Läsionen in Woche 12 wurden mit der mittleren Anzahl der Läsionen bei Patienten ohne

Item[a]	Charakteristikum	Studieninformation
		Therapieerfolg (IGA 0/1) im selben Behandlungsarm ersetzt <u>Sekundäre Endpunkte</u> Für die sekundären Endpunkte wurden Sensitivitätsanalysen analog zu den primären Endpunkten, jeweils auf Basis der PPT- und ITTT-Population, durchgeführt. **Subgruppenanalysen:** Subgruppenanalysen wurden explorativ durchgeführt. Für die primären und sekundären Endpunkte wurde der Behandlungseffekt und das zugehörige 95 %-Konfidenzintervall für die folgenden Subgruppen ausgewertet: • Alterskategorien (9-11, 12-17, ≥ 18 Jahre) • Geschlecht (männlich, weiblich) • Abstammung (weiß, schwarz/afroamerikanisch, asiatisch, andere) • Ethnie (hispanisch, nicht hispanisch) • Hauttyp (I-III, IV-VI)
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
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) Total: 1212; Trifaroten: 602; Vehikel: 610 b) Total: 1212; Trifaroten: 602; Vehikel: 610 c) Total: 1212; Trifaroten: 602; Vehikel: 610
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Es wurden 1293 Patienten gescreent. Davon erfüllten 81 die Einschlusskriterien nicht, und 1212 Patienten wurden randomisiert. Alle 1212 Patienten erhielten mindestens einmal die Studienmedikation. Aufgrund einer fälschlichen Zuteilung der Studienmedikation erhielten insgesamt 3 Patienten nicht die ihnen zugeteilte Behandlung gemäß Randomisierung: ein Patient im Vehikel-Arm erhielt die Creme mit Trifaroten und 2 Patienten im Interventionsarm erhielten die Vehikel-Creme. Die Anzahl der Patienten, die die Behandlung vorzeitig abgebrochen haben, und die Gründe für diese Therapieabbrüche sind im nachfolgenden Flussdiagramm angegeben.
14	Aufnahme / Rekrutierung	

Item[a]	Charakteristikum	Studieninformation
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Studienbeginn: 23. November 2015 (Screening des ersten Patienten) Studienende: 12. Mai 2017 (letzte Visite des letzten Patienten)
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endete planmäßig.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

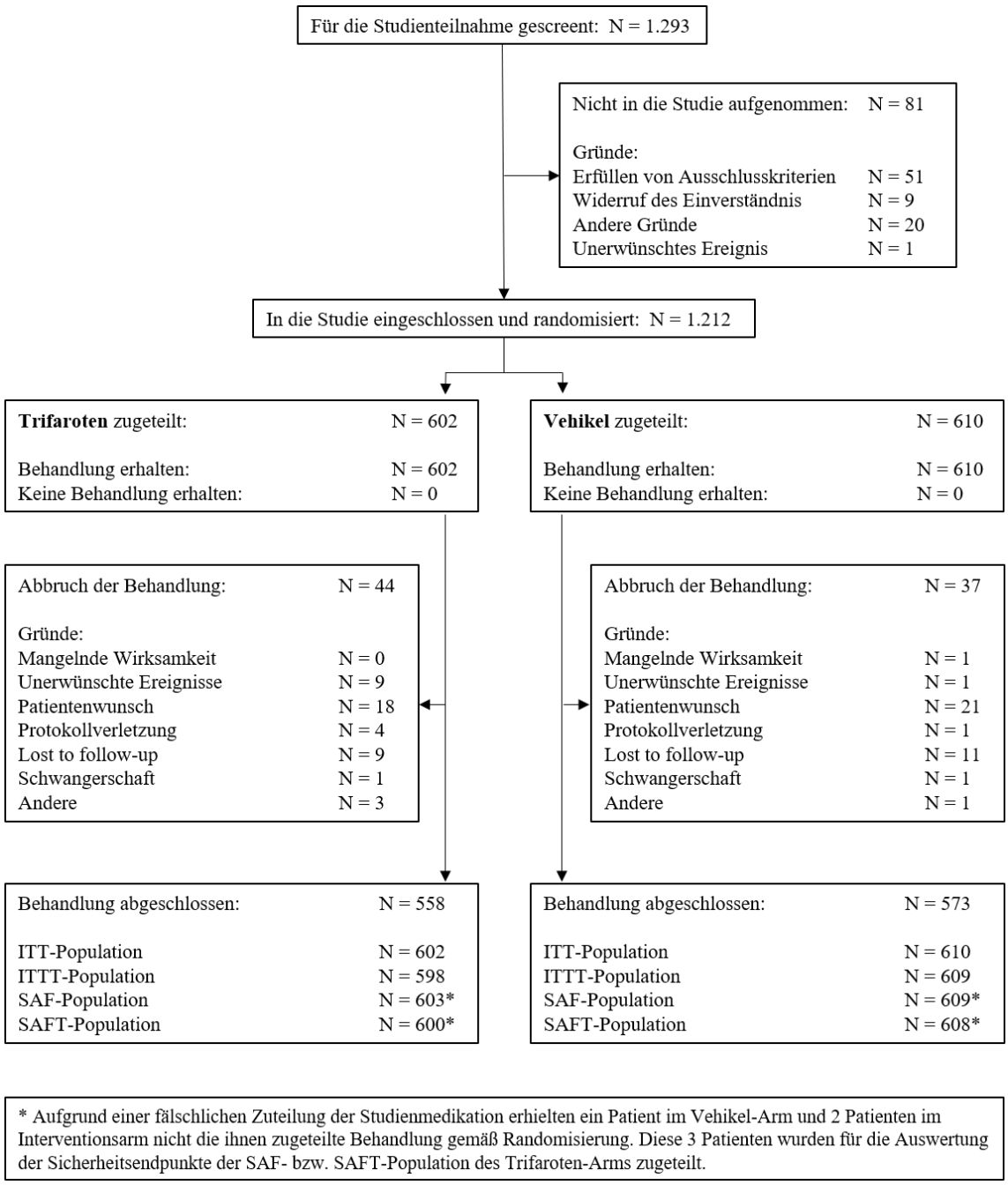


Abbildung 4-25: Patientenfluss der Studie PERFECT-2

Tabelle 4-78 (Anhang): Studiendesign und -methodik für Studie SATISFY (18250)

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2	Hintergrund - Wissenschaftlicher Hintergrund und Erläuterung der Rationale - Theorien bei der Konzeption von Verhaltensinterventionen	Primäres Ziel der Studie war die Untersuchung der Langzeitsicherheit von CD5789 (Trifaroten) 50 µg/g Creme bei täglicher Anwendung für bis zu 52 Wochen bei Patienten mit *Acne vulgaris*. Sekundäres Ziel war die Evaluation der Wirksamkeit.
Methoden		
3	Probanden - Eignungskriterien für Studienteilnehmer, einschließlich der Kriterien auf verschiedenen Ebenen des Rekrutierungsplans und Plans zur Stichprobennahme (z.B. Städte, Kliniken, Patienten) - Rekrutierungsmethode (z.B. Überweisung, Selbstauswahl), einschließlich der Methode zur Stichprobennahme, falls ein systematischer Plan dazu verwendet wurde - Vorgehensweise bei der Rekrutierung - Umfeld der Datenerhebung und Orte, an denen die Daten erhoben wurden	**Einschlusskriterien:** 1. Männliche oder weibliche Teilnehmer im Alter von 9 Jahren oder älter 2. Akne im Gesicht vom Schweregrad 3 (mittelschwer) auf der IGA-Skala zum Zeitpunkt des Screenings und zur Baseline-Visite 3. Mindestens 20 entzündliche und 25 nicht entzündliche Läsionen im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 4. Akne am Rumpf (Schultern, oberer Rücken und oberer Brustkorb – erreichbar zur Selbstanwendung der Studienmedikation durch den Teilnehmer) vom Schweregrad 3 (mittelschwer) auf der PGA-Skala zum Zeitpunkt des Screenings und zur Baseline-Visite; optionales Kriterium für Teilnehmer im Alter von 9 bis 11 Jahren 5. Mindestens 20 entzündliche und 20 nicht entzündliche Läsionen am Rumpf (Schultern, oberer Rücken und oberer Brustkorb – erreichbar zur Selbstanwendung der Studienmedikation durch den Teilnehmer) zum Zeitpunkt des Screenings und zur Baseline-Visite, optionales Kriterium für Teilnehmer im Alter von 9 bis 11 Jahren 6. Frauen außerhalb des gebärfähigen Alters 7. Frauen im gebärfähigen Alter: a. mit negativem Urin-Schwangerschaftstest zum Zeitpunkt des Screenings und bei Behandlungsbeginn b. strikt abstinent im Zeitraum von einem Monat vor Behandlungsbeginn, während der Studie und mindestens einen Monat nach der letzten Anwendung der Studienmedikation <u>oder</u> Anwendung einer geeigneten Verhütungsmethode für die Dauer der Studie bis mindestens einen Monat nach der letzten Anwendung der Studienmedikation 8. Frauen im gebärfähigen Alter mit oralen Kontrazeptiva, die zur Akne-Behandlung zugelassen sind: stabile

Item[a]	Charakteristikum	Studieninformation
		Dosierung für mindestens 6 Monate vor der Baseline-Visite 9. Mädchen, bei denen im Studienverlauf erstmalig die Menstruation einsetzt: a. strikt abstinent im Zeitraum von einem Monat vor Behandlungsbeginn, während der Studie und mindestens einen Monat nach der letzten Anwendung der Studienmedikation oder Anwendung einer geeigneten Verhütungsmethode für die Dauer der Studie bis mindestens einen Monat nach der letzten Anwendung der Studienmedikation 10. Teilnehmer, die gewillt und in der Lage sind, die Anforderungen des Studienprotokolls zu erfüllen 11. Vorliegen einer unterschriebenen Einverständniserklärung; bei minderjährigen Teilnehmern Vorliegen einer durch die Eltern oder Erziehungsberechtigten unterschriebenen Einverständniserklärung und Unterzeichnung einer Willenserklärung durch den Patienten **Ausschlusskriterien:** 1. Schwere Akne-Formen (z. B. *Acne conglobata*, *Acne fulminans*) oder sekundäre Akne-Formen (z. B. Chlor-Akne, medikamenteninduzierte Akne) 2. Mehr als ein Knoten im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 3. Mehr als ein Knoten am Rumpf zum Zeitpunkt des Screenings und zur Baseline-Visite 4. Vorliegen von einer oder mehreren Akne-Zysten im Gesicht zum Zeitpunkt des Screenings und zur Baseline-Visite 5. Vorliegen von einer oder mehreren Akne-Zysten am Rumpf zum Zeitpunkt des Screenings und zur Baseline-Visite 6. Bart oder Gesichts- bzw. Körperbehaarung, die die Untersuchungen in der Studie beeinträchtigt 7. Tattoos am Rumpf, die die Untersuchungen in der Studie beeinträchtigen 8. Vorliegen einer unkontrollierten oder schwerwiegenden Erkrankung oder eines medizinischen Zustandes, welche die Interpretation der Studienergebnisse beeinflussen oder den Patienten bei Studienteilnahme gefährden 9. Klinisch relevante Laborwerte außerhalb des Normbereichs 10. Bekannte oder vermutete Allergien gegen oder Empfindlichkeit für einen Bestandteil der Studienmedikation 11. Stillende Frauen 12. Kinderwunsch im Zeitraum der Studie

Item[a]	Charakteristikum	Studieninformation
		13. Aktuelle Teilnahme an einer anderen Studie oder Teilnahme innerhalb von 30 Tagen vor Baseline-Visite 14. Teilnahme an der Langzeit-Studie RD.03.SPR.18250 15. Einnahme oder Anwendung folgender Wirkstoffe: - Topische Anwendung von Antibiotika, Benzoylperoxid, Azelainsäure, Zink und Kortikosteroiden innerhalb von 2 Wochen vor Baseline-Visite oder Retinoiden innerhalb von 4 Wochen vor Baseline-Visite - Phototherapie und Anwendung anderer Verfahren im Gesicht innerhalb von einer Woche vor Baseline-Visite - Systemische Anwendung von NSAID innerhalb von 2 Wochen vor Baseline-Visite, Kortikosteroiden (außer inhalative oder intrathekale Kortikosteroide) und Antibiotika (außer Penicillin) innerhalb von 4 Wochen vor Baseline-Visite, Retinoide und Immunmodulatoren innerhalb von 6 Monaten vor Baseline-Visite 16. Exzessive UV-Strahlung innerhalb eines Monats vor Baseline-Visite oder während der Studie 17. Keine Bereitschaft, auf die Anwendung von Präparaten mit α-Hydroxycarbonsäuren, medizinische Rasiercremes, Aftershaves, antibakterielle Seifen oder nicht zur Behandlung der Akne zugelassenen Präparaten im Zeitraum der Studie zu verzichten 18. Vulnerable Personen gemäß Abschnitt 1.61 ICH-GCP
4	Interventionen Details der vorgesehenen Interventionen für jede Studiensituation, wie und wann sie verabreicht wurden, und zwar einschließlich: - Inhalt: Was wurde gegeben? - Verabreichungsmethode: Wie wurde der Inhalt gegeben? - Verabreichungseinheit: Wie wurden die Patienten während der Verabreichung eingeteilt? - Verabreicher: Wer verabreichte die Intervention? - Umfeld: Wo wurde die Intervention verabreicht? - Expositionsmenge und-dauer: Wie viele Sitzungen oder Episoden oder Ereignisse sollten verabreicht werden? Wie lange sollten sie dauern?	**Intervention:** • Creme mit 50 µg/g Trifaroten einmal täglich für bis zu 52 Wochen Alle Teilnehmer erhielten mündliche und schriftliche Anweisungen zur richtigen Dosierung und Anwendung der Studienmedikation. Die Studienmedikation wurde von den Teilnehmern wie folgt angewendet: • Anwendung abends nach Reinigung der Haut • Anwendung im gesamten Gesicht und am gesamten oberen Rumpf, auch in Bereichen ohne Läsionen • Dosierung: ein Pumpenhub für das Gesicht und zwei Pumpenhübe für den Rumpf Die Teilnehmer wurden ermutigt, eine nicht komedogene und hypoallergene Feuchtigkeitscreme eine Stunde vor und nach der Studienmedikation anzuwenden, um Hautirritationen zu verringern. **Dauer der Anwendung:** • **Bis Woche 12:** tägliche Anwendung im Gesicht und am Rumpf • **Woche 12 bis Woche 52:**

Item[a]	Charakteristikum	Studieninformation
	- Zeitspanne: Wie lange sollte die Verabreichung der Intervention bei jeder Einheit dauern? - Aktivitäten zur Erhöhung der Compliance oder Adhärenz (z.B. Anreize)	- Patienten mit IGA 0 und PGA 0 → keine Anwendung von Trifaroten - Patienten mit IGA 0 und PGA ≥ 1 → Anwendung von Trifaroten nur am Rumpf - Patienten mit IGA ≥ 1und PGA 0 → Anwendung von Trifaroten nur im Gesicht - Patienten mit wiederauftretender Akne im Gesicht und am Rumpf nehmen die Behandlung im Gesicht bzw. am Rumpf wieder auf
5	Zielsetzungen - Spezifische Studienziele und Hypothesen	**Studienziel:** Die Wirksamkeit und Sicherheit von CD5789 (Trifaroten) 50 µg/g Creme bei täglicher Anwendung für bis zu 52 Wochen bei Patienten mit *Acne vulgaris* sollte untersucht werden. **Hypothese:** Die tägliche topische Anwendung von Trifaroten 50 µg/g Creme für bis zu 52 Wochen ist sicher und wirksam bei Patienten mit *Acne vulgaris*.
6	Zielkriterien - Klar definierte primäre und sekundäre Wirksamkeitsmessgrößen - Methoden zur Datenerhebung und Methoden zur Verbesserung der Qualität von Messungen - Informationen über validierte Instrumente wie beispielsweise psychometrische und biometrische Eigenschaften	**Primäre Zielkriterien:** • Sicherheitsparameter (Unerwünschte Ereignisse, lokale Verträglichkeit im Gesicht und am Rumpf, Laborwertänderungen) **Sekundäre und weitere Zielkriterien:** • Therapieerfolgsrate definiert als Anteil der Patienten mit IGA 0 oder 1 • Therapieerfolgsrate definiert als Anteil der Patienten mit IGA 0 oder 1 • Änderung des IGA im Vergleich zum Baseline-Wert • Änderung des PGA im Vergleich zum Baseline-Wert • Selbsteinschätzung der Patienten zur Verbesserung der Akne im Gesicht • *Dermatology Life Quality Index* (DLQI) / *Children´s Dermatology Life Quality Index* (CDLQI) • PK-Parameter
7	Fallzahl - Wie die Stichprobengröße bestimmt wurde und, falls zutreffend, Erläuterung etwaiger Zwischenanalysen und Abbruchregeln	Es sollten etwa 600 Patienten gescreent werden, um 450 Studienteilnehmer weltweit einschließen zu können. Die Fallzahl sollte sicherstellen, dass mindestens 300 Teilnehmer über 6 Monate und mindestens 100 Teilnehmer über ein Jahr an der Studie teilnehmen. Dies entspricht den Vorgaben der ICH-E1a-Leitlinie zur Bewertung der Sicherheit von Arzneimitteln für die chronische Anwendung.
8	Zuweisungsmethode - Zuweisungseinheit (die Einheit, die der Studienbedingung zugewiesen wird, wie z.B. Person, Gruppe, Gemeinde)	Eine Randomisierung wurde nicht durchgeführt, da es sich um eine offene einarmige Phase III-Studie handelt. Alle in die Studie eingeschlossenen Patienten erhielten die Intervention Trifaroten.

Item[a]	Charakteristikum	Studieninformation
	- Methode der Zuweisung von Einheiten zu Studienbedingungen, einschließlich Details über etwaige Restriktion (z.B. Blockbildung, Stratifizierung, Minimierung) - Einbeziehung von Aspekten, um potenzielle Verzerrung durch Nichtrandomisierung zu minimieren (z.B. Matching)	
9	Verblindung Unabhängig davon, ob sie Teilnehmer waren: - Wurden diejenigen, die die Interventionen durchführten, und diejenigen, die die Ergebnisse beurteilten, gegenüber der Zuweisung der Studienbedingung verblindet? - Wenn dies der Fall ist, Erklärung, wie die Verblindung erfolgte und wie sie bewertet wurde	Es handelt sich um eine offene Studie, eine Verblindung lag nicht vor.
10	Analyseeinheit - Bezeichnung der kleinsten Einheit, die analysiert wird, um Wirkungen der Interventionen zu beurteilen (z.B. Person, Gruppe, Gemeinde) - Wenn sich die Analyseeinheit von der Zuweisungseinheit unterscheidet, ist die verwendete Analysemethode anzugeben, die dies berücksichtigt (z.B. Adjustierung der Standardfehlerschätzungen durch den Designeffekt oder durch Verwendung einer Multilevel-Analyse)	Kleinste analysierte Einheit: Patient
11	Statistische Methoden - Verwendete statistische Methoden zum Vergleich der Studiengruppen in Bezug auf	**Populationen zur Auswertung:** Alle Sicherheitsendpunkte mit Ausnahme der lokalen Verträglichkeit am Rumpf sowie die Auswertungen zum IGA wurden auf Basis der SAF(*safety*)-Population ausgewertet.

Item[a]	Charakteristikum	Studieninformation
	die primär(en) Wirksamkeitsvariable(n), einschließlich komplexer Methoden für korrelierte Daten - Verwendete statistische Methoden für zusätzliche Analysen wie z.B. Subgruppenanalysen und adjustierte Analysen - Methoden zur Imputation fehlender Werte - Verwendete statistische Software oder Programme	Diese umfasste alle randomisierten Patienten, die mindestens einmal die Studienmedikation angewendet haben. Die lokale Verträglichkeit am Rumpf wurden auf Basis der SAFT(*safety on the trunk*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten, die mindestens einmal die Studienmedikation am Rumpf angewendet haben. Auswertungen zum PGA wurden auf Basis der SAFP(*safety for the PGA*)-Population ausgewertet. Diese umfasste alle randomisierten Patienten, die bei Studienbeginn eine mittelschwere Akne am Rumpf aufwiesen. **Statistische Auswertungen:** • kategorielle Variablen: Deskriptive Analysen inklusive Häufigkeiten (Patientenanzahl) und Prozentzahlen (Anteile) • kontinuierliche Variablen: Deskriptive Analysen inklusive Mittelwert, Median, Standardabweichung, Bereich (Minimum; Maximum) • keine Ersetzung fehlender Daten (Auswertung *as observed*)
Resultate		
12	Patientenfluss Durchlauf der Studienteilnehmer durch jede Studienphase: Rekrutierung, Zuweisung, Allokation und Interventions-exposition, Nachbeobachtung, Analyse	siehe Flow Chart
13	Rekrutierung Daten, die den Rekrutierungs- und den Nachbeobachtungszeitraum eingrenzen	Studienbeginn: 23. Februar 2015 (Screening des ersten Patienten) Studienende: 23. Februar 2017 (letzte Visite des letzten Patienten) Die Studie endete planmäßig.
a: nach TREND 2004		

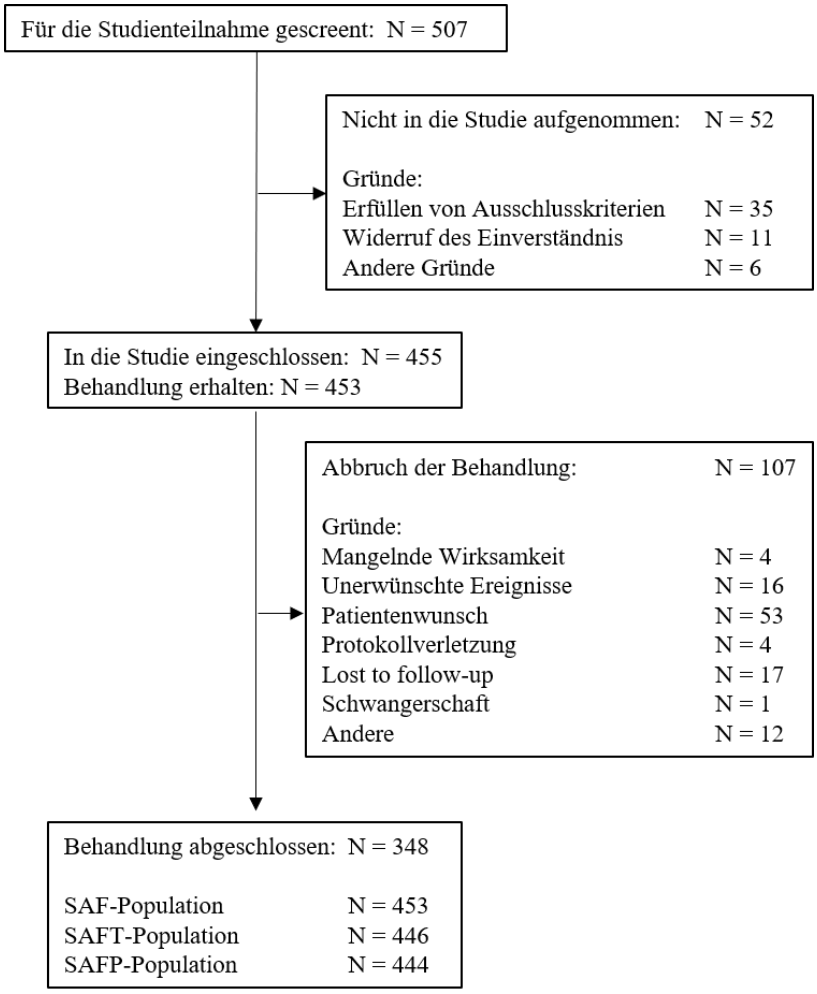


Abbildung 4-26: Patientenfluss der Studie SATISFY

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 4-79 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PERFECT-1

Studie: ____PERFECT-1_____

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht: Galderma R&D (2018): Clinical Study Report – RD.03.SRE.18251; A multicenter, randomized, double-blind, parallel-group vehicle-controlled study to compare the efficacy and safety of CD5789 50µg/g cream versus vehicle cream in subjects with acne vulgaris.	

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:

Die Patienten wurden mittels Blockrandomisierung mit Stratifizierung nach Zentrum im Verhältnis 1:1 den Behandlungsgruppen zugeteilt. Die Randomisierung erfolgte zentral über ein interaktives System (IRT, Interactive Response Technology).

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate 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 Patienten waren verblindet. Die Cremes mit Trifaroten und Vehikel waren in identischen Pumpflaschen abgefüllt. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:

Die behandelnden Personen waren verblindet. Die Cremes mit Trifaroten und Vehikel waren in identischen Pumpflaschen abgefüllt. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> 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:

Bei der Studie PERFECT-1 handelt es sich um eine randomisierte, kontrollierte, doppelblinde Phase-III-Studie. Die Randomisierung auf die Behandlungsgruppen wurde mittels eines Interactive-Response-Technology-Systems realisiert. Patienten und behandelnde Personen waren verblindet. Die Erhebung und Auswertung der Endpunkte erfolgte verblindet. Eine ergebnisunabhängige Berichterstattung liegt vor. Abschließend wird das Verzerrungspotential als niedrig eingestuft.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: __Therapieerfolgsrate im Gesicht gemäß IGA_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA ist somit als niedrig einzustufen.

Endpunkt: __Therapieerfolgsrate am Rumpf gemäß PGA_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Die ITTT-Population umfasst mindestens 98 % der Patienten der ITT-Population der jeweiligen Behandlungsarme.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA ist somit als niedrig einzustufen.

Endpunkt: __ Änderung der Anzahl der Läsionen im Gesicht _____

1. **Verblindung der Endpunkterheber**

   ☒ **ja** ☐ **unklar** ☐ **nein**

   Angaben zum Kriterium; obligate Begründung für die Einstufung:

   Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Änderung der Anzahl der Läsionen im Gesicht ist somit als niedrig einzustufen.

Endpunkt: __ Änderung der Anzahl der Läsionen am Rumpf _____

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Die ITTT-Population umfasst mindestens 98 % der Patienten der ITT-Population der jeweiligen Behandlungsarme.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Änderung der Anzahl der Läsionen am Rumpf somit als niedrig einzustufen.

Endpunkt: __ Selbsteinschätzung der Verbesserung der Akne_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Selbsteinschätzung der Verbesserung der Akne ist somit als niedrig einzustufen.

Endpunkt: __ Gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI _____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt ist patientenberichtet. Die Patienten waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. In allen Behandlungsgruppen liegt die Rücklaufquote für den DLQI bzw. den CDLQI in Woche 12 bei über 80 %. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI ist somit als niedrig einzustufen.

Endpunkt: __ Sicherheit_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Die Sicherheitsendpunkte wurden verblindet erhoben und auf Basis der SAF Population adäquat ausgewertet. Diese umfasst alle randomisierten Patienten, die mindestens einmal die Studienmedikation angewendet haben. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für die Endpunkte zur Sicherheit ist somit als niedrig einzustufen.

Endpunkt: __ Lokale Verträglichkeit_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-1 als niedrig einzuschätzen. Die Sicherheitsendpunkte wurden verblindet erhoben und auf Basis der SAF- bzw. SAFT-Population adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für die Endpunkte zur lokalen Verträglichkeit ist somit als niedrig einzustufen.

Tabelle 4-80 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PERFECT-2

Studie: ____PERFECT-2_____

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht: Galderma R&D (2018): Clinical Study Report – RD.03.SRE.18252; A multicenter, randomized, double-blind, parallel-group vehicle-controlled study to compare the efficacy and safety of CD5789 50 µg/g cream versus vehicle cream in subjects with acne vulgaris.	

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:

Die Patienten wurden mittels Blockrandomisierung mit Stratifizierung nach Zentrum im Verhältnis 1:1 den Behandlungsgruppen zugeteilt. Die Randomisierung erfolgte zentral über ein interaktives System (IRT, Interactive Response Technology).

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate 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 Patienten waren verblindet. Die Cremes mit Trifaroten und Vehikel waren in identischen Pumpflaschen abgefüllt. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; <u>obligate</u> Begründung für die Einstufung:

Die behandelnden Personen waren verblindet. Die Cremes mit Trifaroten und Vehikel waren in identischen Pumpflaschen abgefüllt. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> 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:

Bei der Studie PERFECT-2 handelt es sich um eine randomisierte, kontrollierte, doppelblinde Phase-III-Studie. Die Randomisierung auf die Behandlungsgruppen wurde mittels eines Interactive-Response-Technology-Systems realisiert. Patienten und behandelnde Personen waren verblindet. Die Erhebung und Auswertung der Endpunkte erfolgte verblindet. Eine ergebnisunabhängige Berichterstattung liegt vor. Abschließend wird das Verzerrungspotential als niedrig eingestuft.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: __Therapieerfolgsrate im Gesicht gemäß IGA_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Therapieerfolgsrate im Gesicht gemäß IGA ist somit als niedrig einzustufen.

Endpunkt: __Therapieerfolgsrate am Rumpf gemäß PGA_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Die ITTT-Population umfasst mindestens 98 % der Patienten der ITT-Population der jeweiligen Behandlungsarme.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Therapieerfolgsrate am Rumpf gemäß PGA ist somit als niedrig einzustufen.

Endpunkt: __ Änderung der Anzahl der Läsionen im Gesicht _____

1. **Verblindung der Endpunkterheber**

   ☒ **ja** ☐ **unklar** ☐ **nein**

   Angaben zum Kriterium; obligate Begründung für die Einstufung:

   Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Änderung der Anzahl der Läsionen im Gesicht ist somit als niedrig einzustufen.

Endpunkt: __ Änderung der Anzahl der Läsionen am Rumpf _____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Die ITTT-Population umfasst mindestens 98 % der Patienten der ITT-Population der jeweiligen Behandlungsarme.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITTT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Änderung der Anzahl der Läsionen am Rumpf somit als niedrig einzustufen.

Endpunkt: __ Selbsteinschätzung der Verbesserung der Akne_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt Selbsteinschätzung der Verbesserung der Akne ist somit als niedrig einzustufen.

Endpunkt: __ Gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI _____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Endpunkt ist patientenberichtet. Die Patienten waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Der Endpunkt wurde verblindet erhoben und auf Basis der ITT-Population unter Ersetzung fehlender Werte adäquat ausgewertet. In allen Behandlungsgruppen liegt die Rücklaufquote für den DLQI bzw. den CDLQI in Woche 12 bei über 80 %. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für den Endpunkt gesundheitsbezogene Lebensqualität gemäß DLQI / CDLQI ist somit als niedrig einzustufen.

Endpunkt: __ Sicherheit_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Die Sicherheitsendpunkte wurden verblindet erhoben und auf Basis der SAF Population adäquat ausgewertet. Diese umfasst alle randomisierten Patienten, die mindestens einmal die Studienmedikation angewendet haben. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für die Endpunkte zur Sicherheit ist somit als niedrig einzustufen.

Endpunkt: __ Lokale Verträglichkeit_____

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Endpunkterheber waren verblindet. Die Geheimhaltung der Zuteilung während der Studie war gewährleistet.

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:

Das Verzerrungspotential auf Studienebene ist für die Studie PERFECT-2 als niedrig einzuschätzen. Die Sicherheitsendpunkte wurden verblindet erhoben und auf Basis der SAF- bzw. SAFT-Population adäquat ausgewertet. Es gibt keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Sonstige verzerrende Aspekt liegen nicht vor. Das Verzerrungspotential für die Endpunkte zur lokalen Verträglichkeit ist somit als niedrig einzustufen.

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.

<u>für randomisierte Studien:</u>

Verdeckung der Gruppenzuteilung („allocation concealment“)

☐ **ja**: <u>Eines der folgenden Merkmale trifft zu:</u>
- 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, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u>

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

☐ **ja**: <u>Eines der folgenden Merkmale trifft zu:</u>
- 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, <u>obligate</u> 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.*
  - *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:

Anhang 4-G: Ergänzende Daten

Auswertungen zum DLQI/CDLQI aus PERFECT-1 (18251) und PERFECT-2 (18252) 236
Unerwünschte Ereignisse aus PERFECT-1 (18251) und PERFECT-2 (18252)..... 244
Subgruppenanalysen aus PERFECT-1 (18251) und PERFECT-2 (18252)..... 266

Auswertungen zum DLQI/CDLQI aus PERFECT-1 (18251) und PERFECT-2 (18252)

Galderma R&D
RD.06.SPR.18251

Page 1 of 2

Table 14.2.11.1 Summary of Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Baseline, n	351	364
Mean (SD)	5.9 (5.31)	6.0 (5.27)
Median	4.0	5.0
Min, Max	0, 27	0, 25
No Effect, n (%)	72 (20.5)	80 (22.0)
Small Effect, n (%)	138 (39.3)	121 (33.2)
Moderate Effect, n (%)	79 (22.5)	92 (25.3)
Very Large Effect, n (%)	53 (15.1)	65 (17.9)
Extremely Large Effect, n (%)	9 (2.6)	6 (1.6)

DLQI = dermatology life quality index; SD = standard deviation.
Note: The DLQI is designed for use in adults over the age of 16. Counts and percentages on this table are limited to subjects with DLQI data.
Note: The DLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on DLQI Total Score is calculated as 0-1: No Effect, 2-5: Small Effect, 6-10: Moderate Effect, 11-20: Very Large Effect, 21-30: Extremely Large Effect.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_qs200s1.sas Executed: 13MAR2018

Page 418 of 1289

Table 14.2.11.1 Summary of Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Week 12, n	306	321
Mean (SD)	3.0 (3.58)	2.7 (3.28)
Median	2.0	2.0
Min, Max	0, 22	0, 23
No Effect, n (%)	133 (43.5)	144 (44.9)
Small Effect, n (%)	130 (42.5)	133 (41.4)
Moderate Effect, n (%)	25 (8.2)	32 (10.0)
Very Large Effect, n (%)	16 (5.2)	11 (3.4)
Extremely Large Effect, n (%)	2 (0.7)	1 (0.3)
Week 12 Change from Baseline, n	301	319
Mean (SD)	-2.6 (4.57)	-3.1 (4.08)
Median	-2.0	-2.0
Min, Max	-22, 10	-18, 8

DLQI = dermatology life quality index; SD = standard deviation.
Note: The DLQI is designed for use in adults over the age of 16. Counts and percentages on this table are limited to subjects with DLQI data.
Note: The DLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on DLQI Total Score is calculated as 0-1: No Effect, 2-5: Small Effect, 6-10: Moderate Effect, 11-20: Very Large Effect, 21-30: Extremely Large Effect.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_qs200s1.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18251

Page 1 of 2

Table 14.2.11.2 Summary of Children's Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Baseline, n	254	229
Mean (SD)	3.9 (3.93)	3.9 (3.77)
Median	3.0	3.0
Min, Max	0, 22	0, 23
No Effect, n (%)	75 (29.5)	68 (29.7)
Small Effect, n (%)	137 (53.9)	125 (54.6)
Moderate Effect, n (%)	32 (12.6)	27 (11.8)
Very Large Effect, n (%)	6 (2.4)	7 (3.1)
Extremely Large Effect, n (%)	4 (1.6)	2 (0.9)

CDLQI = children's dermatology life quality index; SD = standard deviation.
Note: The CDLQI is designed for use in children at or under the age of 16. Counts and percentages on this table are limited to subjects with CDLQI data.
Note: The CDLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on CDLQI Total Score is calculated as 0-1: No Effect, 2-6: Small Effect, 7-12: Moderate Effect, 13-18: Very Large Effect, 19-30: Extremely Large Effect.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_qs200s2.sas Executed: 13MAR2018

Page 420 of 1289

Table 14.2.11.2 Summary of Children's Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Week 12, n	234	215
Mean (SD)	2.9 (3.08)	2.2 (2.65)
Median	2.0	1.0
Min, Max	0, 17	0, 13
No Effect, n (%)	96 (41.0)	109 (50.7)
Small Effect, n (%)	111 (47.4)	88 (40.9)
Moderate Effect, n (%)	20 (8.5)	16 (7.4)
Very Large Effect, n (%)	7 (3.0)	2 (0.9)
Extremely Large Effect, n (%)	0	0
Week 12 Change from Baseline, n	234	214
Mean (SD)	-1.0 (3.43)	-1.6 (3.54)
Median	-1.0	-1.0
Min, Max	-17, 12	-19, 9

CDLQI = children's dermatology life quality index; SD = standard deviation.
Note: The CDLQI is designed for use in children at or under the age of 16. Counts and percentages on this table are limited to subjects with CDLQI data.
Note: The CDLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on CDLQI Total Score is calculated as 0-1: No Effect, 2-6: Small Effect, 7-12: Moderate Effect, 13-18: Very Large Effect, 19-30: Extremely Large Effect.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_qs200s2.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18252

Table 14.2.11.1 Summary of Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Baseline, n	382	373
Mean (SD)	6.9 (5.22)	7.0 (5.45)
Median	6.0	6.0
Min, Max	0, 25	0, 29
No Effect, n (%)	57 (14.9)	53 (14.2)
Small Effect, n (%)	124 (32.5)	131 (35.1)
Moderate Effect, n (%)	108 (28.3)	102 (27.3)
Very Large Effect, n (%)	86 (22.5)	78 (20.9)
Extremely Large Effect, n (%)	7 (1.8)	9 (2.4)

DLQI = dermatology life quality index; SD = standard deviation.
Note: The DLQI is designed for use in adults over the age of 16. Counts and percentages on this table are limited to subjects with DLQI data.
Note: The DLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on DLQI Total Score is calculated as 0-1: No Effect, 2-5: Small Effect, 6-10: Moderate Effect, 11-20: Very Large Effect, 21-30: Extremely Large Effect.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_qs200s1.sas Executed: 13MAR2018

Table 14.2.11.1 Summary of Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Week 12, n	360	347
Mean (SD)	3.7 (4.34)	3.7 (4.38)
Median	2.0	2.0
Min, Max	0, 30	0, 25
No Effect, n (%)	129 (35.8)	139 (40.1)
Small Effect, n (%)	145 (40.3)	125 (36.0)
Moderate Effect, n (%)	64 (17.8)	57 (16.4)
Very Large Effect, n (%)	18 (5.0)	24 (6.9)
Extremely Large Effect, n (%)	4 (1.1)	2 (0.6)
Week 12 Change from Baseline, n	358	347
Mean (SD)	-3.3 (4.82)	-3.3 (4.64)
Median	-3.0	-2.0
Min, Max	-17, 22	-24, 13

DLQI = dermatology life quality index; SD = standard deviation.
Note: The DLQI is designed for use in adults over the age of 16. Counts and percentages on this table are limited to subjects with DLQI data.
Note: The DLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on DLQI Total Score is calculated as 0-1: No Effect, 2-5: Small Effect, 6-10: Moderate Effect, 11-20: Very Large Effect, 21-30: Extremely Large Effect.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_qs200s1.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18252

Table 14.2.11.2 Summary of Children's Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Baseline, n	217	236
Mean (SD)	4.5 (3.94)	4.6 (4.41)
Median	3.0	3.0
Min, Max	0, 21	0, 20
No Effect, n (%)	55 (25.3)	70 (29.7)
Small Effect, n (%)	104 (47.9)	105 (44.5)
Moderate Effect, n (%)	49 (22.6)	43 (18.2)
Very Large Effect, n (%)	8 (3.7)	15 (6.4)
Extremely Large Effect, n (%)	1 (0.5)	3 (1.3)

CDLQI = children's dermatology life quality index; SD = standard deviation.
Note: The CDLQI is designed for use in children at or under the age of 16. Counts and percentages on this table are limited to subjects with CDLQI data.
Note: The CDLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on CDLQI Total Score is calculated as 0-1: No Effect, 2-6: Small Effect, 7-12: Moderate Effect, 13-18: Very Large Effect, 19-30: Extremely Large Effect.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_qs200s2.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18252

Table 14.2.11.2 Summary of Children's Dermatology Life Quality Index Total Score
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Week 12, n	203	225
Mean (SD)	3.1 (3.29)	2.7 (3.39)
Median	2.0	2.0
Min, Max	0, 18	0, 23
No Effect, n (%)	77 (37.9)	104 (46.2)
Small Effect, n (%)	103 (50.7)	97 (43.1)
Moderate Effect, n (%)	18 (8.9)	20 (8.9)
Very Large Effect, n (%)	5 (2.5)	2 (0.9)
Extremely Large Effect, n (%)	0	2 (0.9)
Week 12 Change from Baseline, n	202	225
Mean (SD)	-1.4 (3.76)	-1.9 (3.73)
Median	-1.0	-1.0
Min, Max	-15, 12	-15, 7

CDLQI = children's dermatology life quality index; SD = standard deviation.
Note: The CDLQI is designed for use in children at or under the age of 16. Counts and percentages on this table are limited to subjects with CDLQI data.
Note: The CDLQI total score ranges from 0 to 30 with higher scores representing greater impairment of quality of life due to acne.
Note: The effect based on CDLQI Total Score is calculated as 0-1: No Effect, 2-6: Small Effect, 7-12: Moderate Effect, 13-18: Very Large Effect, 19-30: Extremely Large Effect.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_qs200s2.sas Executed: 13MAR2018

Unerwünschte Ereignisse aus PERFECT-1 (18251) und PERFECT-2 (18252)

Galderma R&D
RD.06.SPR.18251

Page 1 of 10

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Number of TEAEs	402	192
Subjects with any TEAE, n (%)	209 (33.9)	123 (20.8)
General disorders and administration site conditions	92 (14.9)	15 (2.5)
Application site irritation	66 (10.7)	4 (0.7)
Application site pruritus	24 (3.9)	8 (1.4)
Application site pain	4 (0.6)	0
Application site discolouration	2 (0.3)	0
Application site dryness	2 (0.3)	0
Application site rash	2 (0.3)	0
Application site swelling	2 (0.3)	0
Application site erosion	1 (0.2)	0
Application site reaction	1 (0.2)	0
Application site urticaria	1 (0.2)	0
Application site vesicles	1 (0.2)	0
Chest pain	1 (0.2)	0
Influenza like illness	0	2 (0.3)
Application site exfoliation	0	1 (0.2)
Pyrexia	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Page 594 of 1289

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Infections and infestations	70 (11.3)	63 (10.7)
Nasopharyngitis	24 (3.9)	27 (4.6)
Upper respiratory tract infection	10 (1.6)	8 (1.4)
Influenza	6 (1.0)	9 (1.5)
Sinusitis	6 (1.0)	1 (0.2)
Pharyngitis streptococcal	4 (0.6)	3 (0.5)
Bronchitis	3 (0.5)	1 (0.2)
Gastroenteritis viral	3 (0.5)	0
Gastroenteritis	2 (0.3)	1 (0.2)
Oral herpes	2 (0.3)	1 (0.2)
Application site infection	2 (0.3)	0
Laryngitis	2 (0.3)	0
Impetigo	1 (0.2)	2 (0.3)
Body tinea	1 (0.2)	1 (0.2)
Folliculitis	1 (0.2)	1 (0.2)
Hordeolum	1 (0.2)	1 (0.2)
Tooth abscess	1 (0.2)	1 (0.2)
Viral upper respiratory tract infection	1 (0.2)	1 (0.2)
Vulvovaginal mycotic infection	1 (0.2)	1 (0.2)
Cellulitis	1 (0.2)	0
Conjunctivitis	1 (0.2)	0
Gastrointestinal viral infection	1 (0.2)	0
Gonorrhoea	1 (0.2)	0
Infectious mononucleosis	1 (0.2)	0
Localised infection	1 (0.2)	0
Ophthalmic herpes simplex	1 (0.2)	0

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Pharyngitis	1 (0.2)	0
Staphylococcal skin infection	1 (0.2)	0
Tinea versicolour	1 (0.2)	0
Ear infection	0	2 (0.3)
Labyrinthitis	0	2 (0.3)
Acute haemorrhagic conjunctivitis	0	1 (0.2)
Atypical pneumonia	0	1 (0.2)
Cystitis	0	1 (0.2)
Cytomegalovirus infection	0	1 (0.2)
Epstein-Barr virus infection	0	1 (0.2)
Erythrasma	0	1 (0.2)
Lice infestation	0	1 (0.2)
Lower respiratory tract infection	0	1 (0.2)
Otitis media	0	1 (0.2)
Pneumonia	0	1 (0.2)
Rhinitis	0	1 (0.2)
Tonsillitis	0	1 (0.2)
Tooth infection	0	1 (0.2)
Viral infection	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18251

Page 4 of 10

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Injury, poisoning and procedural complications	50 (8.1)	15 (2.5)
Sunburn	27 (4.4)	5 (0.8)
Skin abrasion	4 (0.6)	1 (0.2)
Drug administered at inappropriate site	3 (0.5)	0
Ligament sprain	3 (0.5)	0
Accidental overdose	2 (0.3)	0
Incorrect dosage administered	2 (0.3)	0
Foot fracture	1 (0.2)	1 (0.2)
Laceration	1 (0.2)	1 (0.2)
Bone contusion	1 (0.2)	0
Contusion	1 (0.2)	0
Facial bones fracture	1 (0.2)	0
Joint injury	1 (0.2)	0
Medication error	1 (0.2)	0
Muscle rupture	1 (0.2)	0
Muscle strain	1 (0.2)	0
Post-traumatic neck syndrome	1 (0.2)	0
Procedural dizziness	1 (0.2)	0
Tibia fracture	1 (0.2)	0
Animal scratch	0	1 (0.2)
Concussion	0	1 (0.2)
Drug administration error	0	1 (0.2)
Face injury	0	1 (0.2)
Hand fracture	0	1 (0.2)
Procedural pain	0	1 (0.2)
Thermal burn	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Page 597 of 1289

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Skin and subcutaneous tissue disorders	24 (3.9)	6 (1.0)
Skin irritation	8 (1.3)	0
Dermatitis allergic	4 (0.6)	1 (0.2)
Acne	3 (0.5)	2 (0.3)
Erythema	2 (0.3)	0
Chronic actinic dermatitis	1 (0.2)	0
Eczema asteatotic	1 (0.2)	0
Ingrowing nail	1 (0.2)	0
Intertrigo	1 (0.2)	0
Rash	1 (0.2)	0
Seborrhoeic dermatitis	1 (0.2)	0
Skin burning sensation	1 (0.2)	0
Skin exfoliation	1 (0.2)	0
Skin hyperpigmentation	1 (0.2)	0
Mechanical urticaria	0	1 (0.2)
Miliaria	0	1 (0.2)
Psoriasis	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Nervous system disorders	12 (1.9)	12 (2.0)
Headache	6 (1.0)	12 (2.0)
Syncope	2 (0.3)	0
Hyperaesthesia	1 (0.2)	0
Nerve compression	1 (0.2)	0
Restless legs syndrome	1 (0.2)	0
Sciatica	1 (0.2)	0
Sinus headache	1 (0.2)	0
Gastrointestinal disorders	12 (1.9)	8 (1.4)
Toothache	3 (0.5)	2 (0.3)
Nausea	2 (0.3)	1 (0.2)
Vomiting	2 (0.3)	0
Abdominal pain upper	1 (0.2)	0
Anal fissure	1 (0.2)	0
Constipation	1 (0.2)	0
Diarrhoea	1 (0.2)	0
Lip blister	1 (0.2)	0
Lip swelling	1 (0.2)	0
Aphthous stomatitis	0	1 (0.2)
Colitis psychogenic	0	1 (0.2)
Dental caries	0	1 (0.2)
Food poisoning	0	1 (0.2)
Gastritis	0	1 (0.2)
Gastrooesophageal reflux disease	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Musculoskeletal and connective tissue disorders	8 (1.3)	4 (0.7)
Arthralgia	2 (0.3)	0
Muscle spasms	2 (0.3)	0
Neck pain	2 (0.3)	0
Back pain	1 (0.2)	1 (0.2)
Myositis	1 (0.2)	0
Plantar fasciitis	1 (0.2)	0
Myalgia	0	2 (0.3)
Fracture pain	0	1 (0.2)
Respiratory, thoracic and mediastinal disorders	7 (1.1)	14 (2.4)
Nasal congestion	3 (0.5)	1 (0.2)
Rhinitis allergic	2 (0.3)	3 (0.5)
Sinus congestion	1 (0.2)	2 (0.3)
Rhinorrhoea	1 (0.2)	0
Oropharyngeal pain	0	7 (1.2)
Cough	0	2 (0.3)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Investigations	5 (0.8)	6 (1.0)
Aspartate aminotransferase increased	3 (0.5)	0
Alanine aminotransferase increased	2 (0.3)	0
Blood creatinine increased	1 (0.2)	2 (0.3)
Blood bilirubin increased	1 (0.2)	1 (0.2)
Haemoglobin increased	1 (0.2)	0
Lymphocyte count decreased	1 (0.2)	0
White blood cell count decreased	1 (0.2)	0
Liver function test abnormal	0	2 (0.3)
Blood cholesterol increased	0	1 (0.2)
Blood glucose increased	0	1 (0.2)
Hepatic enzyme increased	0	1 (0.2)
Reproductive system and breast disorders	5 (0.8)	6 (1.0)
Dysmenorrhoea	5 (0.8)	3 (0.5)
Bartholin's cyst	0	1 (0.2)
Endometriosis	0	1 (0.2)
Polycystic ovaries	0	1 (0.2)
Psychiatric disorders	4 (0.6)	1 (0.2)
Anxiety	2 (0.3)	0
Aggression	1 (0.2)	0
Sleep disorder	1 (0.2)	0
Mood swings	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18251

Page 9 of 10

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Renal and urinary disorders	3 (0.5)	4 (0.7)
Leukocyturia	2 (0.3)	0
Proteinuria	1 (0.2)	0
Urinary tract infection	0	3 (0.5)
Urinary incontinence	0	1 (0.2)
Blood and lymphatic system disorders	3 (0.5)	1 (0.2)
Iron deficiency anaemia	1 (0.2)	1 (0.2)
Eosinophilia	1 (0.2)	0
Lymphadenitis	1 (0.2)	0
Immune system disorders	3 (0.5)	1 (0.2)
Seasonal allergy	2 (0.3)	0
Drug hypersensitivity	1 (0.2)	0
Hypersensitivity	0	1 (0.2)
Eye disorders	2 (0.3)	2 (0.3)
Eyelid oedema	2 (0.3)	0
Eye pain	0	1 (0.2)
Giant papillary conjunctivitis	0	1 (0.2)
Vascular disorders	2 (0.3)	0
Flushing	1 (0.2)	0
Hypertension	1 (0.2)	0

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Page 602 of 1289

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Hepatobiliary disorders	1 (0.2)	1 (0.2)
Hepatitis alcoholic	1 (0.2)	0
Hyperbilirubinaemia	0	1 (0.2)
Ear and labyrinth disorders	1 (0.2)	0
Tympanosclerosis	1 (0.2)	0
Congenital, familial and genetic disorders	0	1 (0.2)
Hereditary angioedema	0	1 (0.2)
Metabolism and nutrition disorders	0	1 (0.2)
Hyperuricaemia	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s1.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18252

Page 1 of 8

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Number of TEAEs	185	146
Subjects with any TEAE, n (%)	122 (20.2)	117 (19.2)
Infections and infestations	56 (9.3)	73 (12.0)
Nasopharyngitis	26 (4.3)	29 (4.8)
Upper respiratory tract infection	9 (1.5)	8 (1.3)
Influenza	5 (0.8)	9 (1.5)
Respiratory tract infection viral	3 (0.5)	4 (0.7)
Sinusitis	2 (0.3)	4 (0.7)
Bronchitis	2 (0.3)	2 (0.3)
Tonsillitis	2 (0.3)	1 (0.2)
Tinea versicolour	2 (0.3)	0
Urinary tract infection	2 (0.3)	0
Viral infection	2 (0.3)	0
Gastrointestinal viral infection	1 (0.2)	0
Gingivitis	1 (0.2)	0
Herpes simplex	1 (0.2)	0
Pharyngitis streptococcal	1 (0.2)	0
Subcutaneous abscess	1 (0.2)	0
Urethritis	1 (0.2)	0
Gastroenteritis	0	3 (0.5)
Oral herpes	0	3 (0.5)
Rhinitis	0	3 (0.5)
Cystitis	0	2 (0.3)
Laryngitis	0	2 (0.3)
Appendicitis	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Page 549 of 1151

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Conjunctivitis	0	1 (0.2)
Gastroenteritis viral	0	1 (0.2)
Impetigo	0	1 (0.2)
Labyrinthitis	0	1 (0.2)
Otitis externa	0	1 (0.2)
Pharyngitis	0	1 (0.2)
General disorders and administration site conditions	29 (4.8)	4 (0.7)
Application site irritation	18 (3.0)	0
Application site pruritus	5 (0.8)	2 (0.3)
Application site pain	4 (0.7)	0
Application site dryness	3 (0.5)	0
Application site erosion	1 (0.2)	0
Application site erythema	1 (0.2)	0
Asthenia	1 (0.2)	0
Cyst rupture	1 (0.2)	0
Application site hypersensitivity	0	1 (0.2)
Pyrexia	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Injury, poisoning and procedural complications	13 (2.2)	4 (0.7)
Sunburn	6 (1.0)	1 (0.2)
Ligament sprain	2 (0.3)	1 (0.2)
Hand fracture	2 (0.3)	0
Thermal burn	2 (0.3)	0
Drug administered at inappropriate site	1 (0.2)	0
Incorrect dosage administered	1 (0.2)	0
Arthropod bite	0	1 (0.2)
Foot fracture	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Gastrointestinal disorders	12 (2.0)	14 (2.3)
Diarrhoea	3 (0.5)	4 (0.7)
Toothache	3 (0.5)	4 (0.7)
Abdominal pain upper	2 (0.3)	1 (0.2)
Nausea	2 (0.3)	0
Vomiting	1 (0.2)	2 (0.3)
Food poisoning	1 (0.2)	1 (0.2)
Abdominal discomfort	1 (0.2)	0
Cheilitis	1 (0.2)	0
Gastrooesophageal reflux disease	0	2 (0.3)
Abdominal pain	0	1 (0.2)
Nervous system disorders	12 (2.0)	4 (0.7)
Headache	10 (1.7)	4 (0.7)
Dizziness	2 (0.3)	0
Migraine	1 (0.2)	0
Seizure	1 (0.2)	0

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Skin and subcutaneous tissue disorders	6 (1.0)	5 (0.8)
Seborrhoeic dermatitis	2 (0.3)	0
Skin irritation	2 (0.3)	0
Intertrigo	1 (0.2)	0
Skin fissures	1 (0.2)	0
Dermatitis	0	1 (0.2)
Dermatitis contact	0	1 (0.2)
Drug eruption	0	1 (0.2)
Rash	0	1 (0.2)
Skin exfoliation	0	1 (0.2)
Reproductive system and breast disorders	6 (1.0)	0
Dysmenorrhoea	6 (1.0)	0
Musculoskeletal and connective tissue disorders	5 (0.8)	0
Arthralgia	2 (0.3)	0
Back pain	2 (0.3)	0
Neck pain	1 (0.2)	0
Pain in extremity	1 (0.2)	0

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Investigations	4 (0.7)	9 (1.5)
Alanine aminotransferase increased	2 (0.3)	2 (0.3)
Blood creatinine increased	2 (0.3)	1 (0.2)
Aspartate aminotransferase increased	1 (0.2)	2 (0.3)
Blood alkaline phosphatase increased	1 (0.2)	0
Liver function test abnormal	0	3 (0.5)
Blood bilirubin increased	0	2 (0.3)
Platelet count decreased	0	1 (0.2)
Renal and urinary disorders	3 (0.5)	3 (0.5)
Dysuria	2 (0.3)	0
Urinary tract infection	1 (0.2)	2 (0.3)
Leukocyturia	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Respiratory, thoracic and mediastinal disorders	2 (0.3)	6 (1.0)
Cough	1 (0.2)	0
Pulmonary mass	1 (0.2)	0
Oropharyngeal pain	0	2 (0.3)
Asthma	0	1 (0.2)
Nasal congestion	0	1 (0.2)
Rhinorrhoea	0	1 (0.2)
Sinus congestion	0	1 (0.2)
Upper respiratory tract congestion	0	1 (0.2)
Eye disorders	2 (0.3)	3 (0.5)
Eyelid oedema	1 (0.2)	1 (0.2)
Eyelid exfoliation	1 (0.2)	0
Photophobia	1 (0.2)	0
Blepharitis	0	1 (0.2)
Chalazion	0	1 (0.2)
Conjunctivitis allergic	0	1 (0.2)
Blood and lymphatic system disorders	2 (0.3)	2 (0.3)
Anaemia	1 (0.2)	1 (0.2)
Lymphadenopathy	1 (0.2)	0
Neutropenia	0	1 (0.2)
Psychiatric disorders	2 (0.3)	2 (0.3)
Major depression	1 (0.2)	1 (0.2)
Suicide attempt	1 (0.2)	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Table 14.3.3.1 Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Insomnia	1 (0.2)	0
Anxiety	0	1 (0.2)
Immune system disorders	1 (0.2)	2 (0.3)
Drug hypersensitivity	1 (0.2)	0
Seasonal allergy	0	2 (0.3)
Hepatobiliary disorders	1 (0.2)	0
Hyperbilirubinaemia	1 (0.2)	0
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	1 (0.2)	0
Skin papilloma	1 (0.2)	0
Metabolism and nutrition disorders	0	1 (0.2)
Hypercholesterolaemia	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s1.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18251

Table 14.3.4.1 Summary of Serious Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Number of serious TEAEs	4	3
Subjects with any serious TEAE, n (%)	4 (0.6)	2 (0.3)
Infections and infestations	2 (0.3)	1 (0.2)
Cellulitis	1 (0.2)	0
Infectious mononucleosis	1 (0.2)	0
Atypical pneumonia	0	1 (0.2)
Injury, poisoning and procedural complications	2 (0.3)	0
Facial bones fracture	1 (0.2)	0
Procedural dizziness	1 (0.2)	0
Congenital, familial and genetic disorders	0	1 (0.2)
Hereditary angioedema	0	1 (0.2)
Renal and urinary disorders	0	1 (0.2)
Urinary tract infection	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s3.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18252

Page 1 of 1

Table 14.3.4.1 Summary of Serious Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Number of serious TEAEs	3	4
Subjects with any serious TEAE, n (%)	2 (0.3)	4 (0.7)
Psychiatric disorders	1 (0.2)	1 (0.2)
Suicide attempt	1 (0.2)	1 (0.2)
Major depression	1 (0.2)	0
Injury, poisoning and procedural complications	1 (0.2)	0
Ligament sprain	1 (0.2)	0
Infections and infestations	0	2 (0.3)
Appendicitis	0	1 (0.2)
Sinusitis	0	1 (0.2)
Respiratory, thoracic and mediastinal disorders	0	1 (0.2)
Asthma	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s3.sas Executed: 13MAR2018

Page 563 of 1151

Galderma R&D
RD.06.SPR.18251

Table 14.3.8.1 Summary of Treatment-Emergent Adverse Events Leading to Discontinuation by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=617)	Vehicle (N=591)
Subjects with any TEAE leading to discontinuation, n (%)	14 (2.3)	1 (0.2)
General disorders and administration site conditions	8 (1.3)	0
Application site irritation	8 (1.3)	0
Skin and subcutaneous tissue disorders	5 (0.8)	0
Acne	2 (0.3)	0
Skin irritation	2 (0.3)	0
Dermatitis allergic	1 (0.2)	0
Injury, poisoning and procedural complications	1 (0.2)	0
Muscle rupture	1 (0.2)	0
Tibia fracture	1 (0.2)	0
Infections and infestations	0	1 (0.2)
Bronchitis	0	1 (0.2)

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5812-007 and 8135-005 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. These subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.251\Statistics\ROUND3\Output\Prog\251_ae201s5.sas Executed: 13MAR2018

Galderma R&D
RD.06.SPR.18252

Table 14.3.8.1 Summary of Treatment-Emergent Adverse Events Leading to Discontinuation by System Organ Class and Preferred Term
Safety Population

System Organ Class/ Preferred Term	CD5789 50 µg/g (N=603)	Vehicle (N=609)
Subjects with any TEAE leading to discontinuation, n (%)	10 (1.7)	1 (0.2)
General disorders and administration site conditions	7 (1.2)	0
Application site irritation	5 (0.8)	0
Application site erosion	1 (0.2)	0
Application site erythema	1 (0.2)	0
Application site pain	1 (0.2)	0
Infections and infestations	1 (0.2)	1 (0.2)
Bronchitis	1 (0.2)	0
Sinusitis	0	1 (0.2)
Immune system disorders	1 (0.2)	0
Drug hypersensitivity	1 (0.2)	0
Psychiatric disorders	1 (0.2)	0
Major depression	1 (0.2)	0
Suicide attempt	1 (0.2)	0

TEAE = treatment-emergent adverse event.
Note: System organ classes are sorted by frequency in the CD5789 group. Preferred terms are sorted by frequency in the CD5789 group within system organ class.
Note: Adverse events are coded using MedDRA version 18.0.
Note: Subjects 5770-011 and 5778-012 were randomized to CD5789 50 ug/g but received Vehicle at or after Baseline. Subject 8208-003, randomized to Vehicle, received CD5789 50 ug/g. These 3 subjects were treated as CD5789 50 ug/g in the SAF and SAFT populations.

Prog: \Galderma.252\Statistics\ROUND3\Output\Prog\252_ae201s5.sas Executed: 13MAR2018

Subgruppenanalysen aus PERFECT-1 (18251) und PERFECT-2 (18252)

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 15 - Analysis of Success Rate of Investigator's Global Assessment at Week 12 by age categories (MI)
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
9 to 11 Years, n	10	9
Success Rate (%)	20.4	12.9
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	7.5 (-26.9, 41.9)	
12 to 17 Years, n	304	269
Success Rate (%)	25.6	14.7
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	10.9 (4.3, 17.6)	
>=18 Years, n	298	318
Success Rate (%)	33.5	23.8
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	9.6 (2.2, 17.1)	
p-value for Heterogeneity Test	0.785	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; IGA=Investigator's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an IGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with IGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by age categories run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 16 - Analysis of Success Rate of Investigator's Global Assessment at Week 12 by gender (MI)
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Female, n	305	324
Success Rate (%)	32.8	20.9
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	11.9 (4.8, 19.0)	
Male, n	307	272
Success Rate (%)	25.9	17.9
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	8.0 (1.0, 14.9)	
p-value for Heterogeneity Test	0.407	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; IGA=Investigator's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an IGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with IGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by gender run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 17 - Analysis of Success Rate of Physician's Global Assessment at Week 12 by age categories (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=600)	Vehicle (N=585)
12 to 17 Years, n	302	269
Success Rate (%)	31.8	21.0
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	10.8 (3.5, 18.1)	
>=18 Years, n	296	314
Success Rate (%)	40.0	28.6
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	11.4 (3.6, 19.2)	
p-value for Heterogeneity Test	0.938	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; PGA=Physician's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an PGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with PGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by age categories run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 18 - Analysis of Success Rate of Physician's Global Assessment at Week 12 by gender (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=600)	Vehicle (N=585)
Female, n	298	316
Success Rate (%)	43.6	28.7
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	14.9 (7.2, 22.6)	
Male, n	302	269
Success Rate (%)	28.0	20.6
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	7.3 (0.2, 14.5)	
p-value for Heterogeneity Test	0.176	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; PGA=Physician's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an PGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with PGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by gender run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 19 - Analysis of Facial Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
9 to 11 Years		
Week 12 Change from Baseline, n	10	9
Mean (SE)	-10.4 (3.62)	-19.0 (3.80)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	8.6 (-1.7, 18.9)	
12 to 17 Years		
Week 12 Change from Baseline, n	304	269
Mean (SE)	-18.7 (0.88)	-14.8 (1.02)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.8 (-6.5, -1.2)	
>=18 Years		
Week 12 Change from Baseline, n	298	318
Mean (SE)	-19.7 (0.70)	-16.1 (0.72)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.6 (-5.6, -1.6)	
p-value for Heterogeneity Test	0.144	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error

Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data

Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 20 - Analysis of Facial Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Female		
Week 12 Change from Baseline, n	305	324
Mean (SE)	-19.5 (0.77)	-16.5 (0.77)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-2.9 (-5.1, -0.8)	
Male		
Week 12 Change from Baseline, n	307	272
Mean (SE)	-18.6 (0.81)	-14.4 (0.95)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.2 (-6.7, -1.7)	
p-value for Heterogeneity Test	0.454	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 21 - Analysis of Facial Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
9 to 11 Years		
Week 12 Change from Baseline, n	10	9
Mean (SE)	-16.2 (8.49)	-7.5 (9.35)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-8.7 (-33.4, 16.0)	
12 to 17 Years		
Week 12 Change from Baseline, n	304	269
Mean (SE)	-26.5 (1.41)	-16.8 (1.62)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-9.6 (-13.8, -5.4)	
>=18 Years		
Week 12 Change from Baseline, n	298	318
Mean (SE)	-25.9 (1.25)	-19.9 (1.27)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-6.0 (-9.5, -2.5)	
p-value for Heterogeneity Test	0.231	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error

Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data

Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 22 - Analysis of Facial Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Population

	CD5789 50 µg/g (N=612)	Vehicle (N=596)
Female		
Week 12 Change from Baseline, n	305	324
Mean (SE)	-27.3 (1.37)	-19.9 (1.40)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-7.4 (-11.2, -3.5)	
Male		
Week 12 Change from Baseline, n	307	272
Mean (SE)	-24.7 (1.28)	-16.4 (1.44)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-8.4 (-12.1, -4.6)	
p-value for Heterogeneity Test	0.740	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 23 - Analysis of Truncal Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=600)	Vehicle (N=585)
12 to 17 Years		
Week 12 Change from Baseline, n	302	269
Mean (SE)	-21.4 (1.03)	-18.0 (1.06)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.4 (-6.3, -0.5)	
>=18 Years		
Week 12 Change from Baseline, n	296	314
Mean (SE)	-21.8 (0.85)	-18.6 (0.88)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.3 (-5.7, -0.9)	
p-value for Heterogeneity Test	0.488	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error

Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 24 - Analysis of Truncal Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=600)	Vehicle (N=585)
Female		
Week 12 Change from Baseline, n	298	316
Mean (SE)	-22.2 (0.83)	-19.9 (0.97)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-2.2 (-4.7, 0.3)	
Male		
Week 12 Change from Baseline, n	302	269
Mean (SE)	-21.0 (1.05)	-16.4 (0.92)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.7 (-7.4, -2.0)	
p-value for Heterogeneity Test	0.941	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 25 - Analysis of Truncal Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=600)	Vehicle (N=585)
12 to 17 Years		
Week 12 Change from Baseline, n	302	269
Mean (SE)	-22.2 (1.30)	-17.2 (1.67)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-5.0 (-9.1, -0.8)	
>=18 Years		
Week 12 Change from Baseline, n	296	314
Mean (SE)	-23.0 (1.35)	-20.6 (1.27)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-2.4 (-6.0, 1.2)	
p-value for Heterogeneity Test	0.413	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error

Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data

Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18251 - Adhoc analysis for the German HTA process

Table 26 - Analysis of Truncal Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=600)	Vehicle (N=585)
Female		
Week 12 Change from Baseline, n	298	316
Mean (SE)	-24.8 (1.31)	-21.2 (1.32)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.6 (-7.2, 0.1)	
Male		
Week 12 Change from Baseline, n	302	269
Mean (SE)	-20.4 (1.32)	-16.1 (1.61)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.2 (-8.3, -0.2)	
p-value for Heterogeneity Test	0.816	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18251-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 10:50

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 15 - Analysis of Success Rate of Investigator's Global Assessment at Week 12 by age categories (MI)
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
9 to 11 Years, n	9	6
Success Rate (%)	33.3	33.3
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	0.0 (-48.7, 48.7)	
12 to 17 Years, n	267	288
Success Rate (%)	35.8	20.4
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	15.4 (7.9, 23.0)	
>=18 Years, n	326	316
Success Rate (%)	47.8	30.3
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	17.5 (10.0, 25.1)	
p-value for Heterogeneity Test	0.775	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; IGA=Investigator's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an IGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with IGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by age categories run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 16 - Analysis of Success Rate of Investigator's Global Assessment at Week 12 by gender (MI)
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Female, n	357	338
Success Rate (%)	43.8	27.3
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	16.4 (9.3, 23.6)	
Male, n	245	272
Success Rate (%)	40.1	23.6
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	16.5 (8.5, 24.6)	
p-value for Heterogeneity Test	0.629	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; IGA=Investigator's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an IGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with IGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by gender run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 17 - Analysis of Success Rate of Physician's Global Assessment at Week 12 by age categories (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=598)	Vehicle (N=609)
12 to 17 Years, n	267	288
Success Rate (%)	38.7	25.8
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	12.9 (5.0, 20.8)	
>=18 Years, n	326	316
Success Rate (%)	45.9	33.5
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	12.4 (4.7, 20.1)	
p-value for Heterogeneity Test	0.755	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; PGA=Physician's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an PGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with PGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by age categories run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 18 - Analysis of Success Rate of Physician's Global Assessment at Week 12 by gender (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=598)	Vehicle (N=609)
Female, n	355	337
Success Rate (%)	46.3	32.2
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	14.1 (6.8, 21.5)	
Male, n	243	272
Success Rate (%)	37.2	27.1
Difference in Success Rates of CD5789 50 µg/g vs. Vehicle (95% CI)	10.1 (1.9, 18.3)	
p-value for Heterogeneity Test	0.696	

Abbreviations: CI=Confidence Interval; MI=Multiple Imputation method for missing data; PGA=Physician's Global Assessment; ITT=Intent-to-Treat
Notes: Success is defined as an PGA score of "clear (0)" or "almost clear (1)" at Week 12 and a grade change of at least two from Baseline
Success Rates are calculated as the number subjects achieving success divided by the number of subjects with PGA data at Week 12 and combined according to Rubin's method
Difference in Success Rates and Confidence Interval are based on the large-sample approximation method for binary data without the use of a continuity correction and combined according to Rubin's method
p-value for heterogeneity test is based on the Breslow-Day test stratified by gender run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 19 - Analysis of Facial Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
9 to 11 Years		
Week 12 Change from Baseline, n	9	6
Mean (SE)	-16.3 (4.26)	-16.5 (5.02)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	0.2 (-14.3, 14.6)	
12 to 17 Years		
Week 12 Change from Baseline, n	267	288
Mean (SE)	-24.0 (0.90)	-18.7 (1.07)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-5.3 (-8.1, -2.6)	
>=18 Years		
Week 12 Change from Baseline, n	326	316
Mean (SE)	-24.5 (0.83)	-20.2 (0.86)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.3 (-6.7, -2.0)	
p-value for Heterogeneity Test	0.709	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 20 - Analysis of Facial Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Female		
Week 12 Change from Baseline, n	357	338
Mean (SE)	-25.1 (0.80)	-19.8 (0.84)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-5.3 (-7.6, -3.1)	
Male		
Week 12 Change from Baseline, n	245	272
Mean (SE)	-22.8 (0.93)	-19.0 (1.10)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.8 (-6.6, -1.0)	
p-value for Heterogeneity Test	0.377	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 21 - Analysis of Facial Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
9 to 11 Years		
Week 12 Change from Baseline, n	9	6
Mean (SE)	-14.2 (6.45)	-12.5 (16.70)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-1.7 (-44.7, 41.3)	
12 to 17 Years		
Week 12 Change from Baseline, n	267	288
Mean (SE)	-33.8 (1.76)	-22.8 (1.21)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-11.0 (-15.2, -6.8)	
>=18 Years		
Week 12 Change from Baseline, n	326	316
Mean (SE)	-27.3 (1.06)	-22.1 (1.14)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-5.2 (-8.3, -2.2)	
p-value for Heterogeneity Test	0.737	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 22 - Analysis of Facial Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Population

	CD5789 50 µg/g (N=602)	Vehicle (N=610)
Female		
Week 12 Change from Baseline, n	357	338
Mean (SE)	-32.0 (1.35)	-23.6 (1.13)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-8.4 (-11.8, -5.0)	
Male		
Week 12 Change from Baseline, n	245	272
Mean (SE)	-27.1 (1.39)	-20.8 (1.23)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-6.3 (-10.0, -2.7)	
p-value for Heterogeneity Test	0.548	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 23 - Analysis of Truncal Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=598)	Vehicle (N=609)
12 to 17 Years		
Week 12 Change from Baseline, n	267	288
Mean (SE)	-25.4 (1.03)	-19.2 (1.09)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-6.2 (-9.2, -3.3)	
>=18 Years		
Week 12 Change from Baseline, n	326	316
Mean (SE)	-26.3 (0.95)	-21.7 (0.93)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.6 (-7.2, -2.0)	
p-value for Heterogeneity Test	0.173	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error

Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method

p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data

Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 24 - Analysis of Truncal Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=598)	Vehicle (N=609)
Female		
Week 12 Change from Baseline, n	355	337
Mean (SE)	-26.7 (0.89)	-20.3 (0.91)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-6.4 (-8.9, -3.9)	
Male		
Week 12 Change from Baseline, n	243	272
Mean (SE)	-24.5 (1.12)	-20.3 (1.14)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.1 (-7.3, -1.0)	
p-value for Heterogeneity Test	0.634	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 25 - Analysis of Truncal Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by age categories (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=598)	Vehicle (N=609)
12 to 17 Years		
Week 12 Change from Baseline, n	267	288
Mean (SE)	-25.7 (1.34)	-20.1 (1.14)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-5.7 (-9.1, -2.2)	
>=18 Years		
Week 12 Change from Baseline, n	326	316
Mean (SE)	-26.1 (1.01)	-22.8 (1.10)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-3.3 (-6.2, -0.3)	
p-value for Heterogeneity Test	0.691	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error

Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method

p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, age category and treatment by age category interaction as factors run on observed data

Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09

Galderma R&D
RD.06.SPR.18252 - Adhoc analysis for the German HTA process

Table 26 - Analysis of Truncal Non-Inflammatory Lesion Count Absolute Change from Baseline to Week 12 by gender (MI)
ITT Trunk Population

	CD5789 50 µg/g (N=598)	Vehicle (N=609)
Female		
Week 12 Change from Baseline, n	355	337
Mean (SE)	-27.1 (1.08)	-22.7 (1.07)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.3 (-7.3, -1.3)	
Male		
Week 12 Change from Baseline, n	243	272
Mean (SE)	-24.1 (1.24)	-19.6 (1.18)
Means Difference of CD5789 50 µg/g vs. Vehicle (95% CI)	-4.4 (-7.8, -1.1)	
p-value for Heterogeneity Test	0.598	

Abbreviations: ANCOVA=Analysis of Covariance; CI=Confidence Interval; MI=Multiple Imputation method for missing data; ITT=Intent-to-Treat; SE=Standard Error
Notes: Mean Changes from Baseline, SE, Differences in Mean Changes from Baseline and their Confidence intervals are based on a two-sample t-test (Satterthwaite method) and combined according to Rubin's method
p-value for heterogeneity test is the p-value of the interaction factor derived from Type 3 analysis of ANCOVA model with baseline lesion count, treatment, gender and treatment by gender interaction as factors run on observed data
Program: 18252-adhoc-03-ger-hta-heterogeneity-tests 27MAR2020 11:09