

Tabelle 4-226: Ergebnisse des ITC für FF/UMEC/VI vs. BDP/FF/GB; Studienabbruch wegen unerwünschter Ereignisse – ITT-Population (Beobachtungszeitraum 24 Wochen in der CTT116853 (FULFIL) und 52 Wochen in CTT116855 (IMPACT) und Singh 2016 (TRILOGY))

Ergebnisse für den indirekten Vergleich FF/UMEC/VI vs. BDP/FF/GB nach der Bucher-Methode			
	Odds Ratio	Relatives Risiko	Absolute Risikoreduktion
Effektgröße (95% Konfidenzintervall)	0,79 (0,44; 1,43)	0,80 (0,46; 1,40)	-0,011 (-0,043; 0,021)
p-Wert	0,4410	0,4334	0,5053
Die Werte der CTT116853 (FULFIL) wurden in Woche 24 erhoben, die Werte der CTT116855 (IMPACT) und Singh 2016 (TRILOGY) in Woche 52.			

In der ITC auf Basis der EXT-Population der CTT116853 (FULFIL) und der ITT-Population der CTT116855 (IMPACT) und Singh 2016 (TRILOGY) zeigen das Relative Risiko, das Odds Ratio sowie die absolute Risikoreduktion einen geringeren Anteil an Studienabbrüchen aufgrund von unerwünschten Ereignissen in der Studienmedikation FF/UMEC/VI im Vergleich zu BDP/FF/GB. Der Behandlungsunterschied zwischen FF/UMEC/VI und BDP/FF/GB ist statistisch nicht signifikant ($p \geq 0.05$).

Tabelle 4-227: Ergebnisse des ITC für FF/UMEC/VI vs. BDP/FF/GB; Unerwünschte Ereignisse, die zu einem Studienabbruch führten – EXT-Population in der CTT116853 (FULFIL) und ITT-Population in den CTT116855 (IMPACT) und Singh 2016 (TRILOGY)

Ergebnisse für den indirekten Vergleich FF/UMEC/VI vs. BDP/FF/GB nach der Bucher-Methode			
	Odds Ratio	Relatives Risiko	Absolute Risikoreduktion
Effektgröße (95% Konfidenzintervall)	0,73 (0,43; 1,22)	0,74 (0,45; 1,21)	-0,020 (-0,045; 0,006)
p-Wert	0,2227	0,2296	0,1306
Die Werte aller drei Studien wurden in Woche 52 erhoben.			

4.3.2.1.3.10 FEV₁ (Talspiegel)– indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

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

Tabelle 4-228: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden - FEV₁ (Talspiegel)

Anzahl Studien	Studie	FF/UMEC/VI	ICS/LABA	BDP/FF/GB
1	Singh 2016 (TRILOGY)		•	•
2	CTT116853 (FULFIL)	•	•	
	CTT116855 (IMPACT)	•	•	

In der Studie Singh 2016 (TRILOGY) betrug die Differenz in der Abweichung vom Baselinewert in der für den indirekten Vergleich herangezogenen ITT-Population 81 und 63 für die Beobachtungszeiträume von 26 und 52 Wochen. Ein höherer FEV₁-Wert steht für größere Atemkapazität eines COPD-Patienten.

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

Tabelle 4-229: Operationalisierung für FEV₁ (Talspiegel)

Studie	Operationalisierung
CTT116853 (FULFIL)	Der FEV ₁ Talspiegel gemessen vor der Medikamenteneinnahme war einer der präspezifizierten ko-primären Endpunkte der CTT116853 (FULFIL). Der Endpunkt wurde im Rahmen der ITT-Analyse für Woche 12 und 24, sowie für die EXT-Population an Woche 52, ausgewertet. Die Auswertung des Endpunktes erfolgt mittels eines „Mixed Model for Repeated Measures“ (MMRM) unter Berücksichtigung von Kovariablen.
CTT116855 (IMPACT)	Der FEV ₁ Talspiegel gemessen vor der Medikamenteneinnahme wurde im Rahmen der ITT-Analyse für Woche 16, 28 und 52 ausgewertet. Die Auswertung des Endpunktes erfolgt mittels eines „Mixed Model for Repeated Measures“ (MMRM) unter Berücksichtigung von Kovariablen.
Singh 2016 (TRILOGY)	Der FEV ₁ gemessen vor der Medikamenteneinnahme war einer der präspezifizierten ko-primären Endpunkte der Singh 2016 (TRILOGY). Der Endpunkt wurde im Rahmen der ITT-Analyse für Woche 12, 26 und 52 ausgewertet. Die Auswertung des Endpunktes erfolgt mittels eines „Mixed Model for Repeated Measures“ (MMRM) unter Berücksichtigung von Kovariablen.

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 4-230: Bewertung des Verzerrungspotenzials für FEV₁ (Talspiegel) in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
CTT116853 (FULFIL)	niedrig	ja	ja	ja	ja	niedrig
CTT116855 (IMPACT)	niedrig	ja	unklar	ja	ja	hoch
Singh 2016 (TRILOGY)	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei den Studien CTT116855 (IMPACT) und CTT116853 (FULFIL) handelte es sich um doppelblinde, aktiv kontrollierte, parallele multizentrische klinische Studien. Prüfärzte, Studienpersonal und der Sponsor verblieben über den gesamten Zeitraum verblindet. Die Analyse erfolgte auf Basis der ITT-Population. In der CTT116855 (IMPACT) wurden allerdings nur 96,8% (FF/UMEC/VI) bzw. 94,2% (UMEC/VI) der Patienten der ITT-Population berücksichtigt.

In der Singh 2016 (TRILOGY) wurde der Endpunkt verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt und der Endpunkt wurde ergebnisunabhängig berichtet. Es gibt keine weiteren Aspekte, die zu Verzerrungen führen könnten.

Für die Analyse des Endpunktes FEV₁ (Talspiegel) ergibt sich daher für die Studien CTT116853 (FULFIL) und Singh 2016 (TRILOGY) kein Hinweis auf ein mögliches Verzerrungspotenzial. Für die Studie CTT116855 (IMPACT) wird das Verzerrungspotenzial aufgrund der nicht berücksichtigten Patienten als hoch eingestuft.

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 4-231: Ergebnisse FEV₁ (Talspiegel) für indirekte Vergleiche

Studie	Ergebnisse für FEV ₁ (Talspiegel)	
	FF/UMEC/VI	ICS/LABA
CTT116853 (FULFIL)		
Population	ITT	
N	911	899

Studie	Ergebnisse für FEV ₁ (Talspiegel)	
	Erhebungszeitraum	
	Baseline bis Woche 12	
	<u>Anzahl der Patienten, n (%)</u>	<u>867 (95)</u> <u>826 (92)</u>
	DFCB (ml)	148,00
	95% Untere Konfidenzschranke	125,66
	95% Obere Konfidenzschranke	170,34
	p-Wert	<0,0001
	Erhebungszeitraum	
	Baseline bis Woche 24	
	<u>Anzahl der Patienten, n (%)</u>	<u>836 (92)</u> <u>781 (87)</u>
	DFCB (ml)	171,00
	95% Untere Konfidenzschranke	147,87
	95% Obere Konfidenzschranke	194,13
	p-Wert	<0,0001
	Population	EXT
	N	210 220
	Erhebungszeitraum	
	Baseline bis Woche 52	
	<u>Anzahl der Patienten, n (%)</u>	<u>183 (87)</u> <u>171 (78)</u>
	DFCB (ml)	179,00
	95% Untere Konfidenzschranke	131,57
	95% Obere Konfidenzschranke	226,43
	p-Wert	<0,0001
CTT116855 (IMPACT)	FF/UMEC/VI	ICS/LABA
	ITT	
	Population	
	N	4151 4134
	Erhebungszeitraum	
	Baseline bis Woche 16	
	<u>Anzahl der Patienten, n (%)</u>	<u>3760 (91)</u> <u>3539 (86)</u>
	DFCB (ml)	97,00
	95% Untere Konfidenzschranke	86,42
	95% Obere Konfidenzschranke	107,58
	p-Wert	<0,0001
	Erhebungszeitraum	
	Baseline bis Woche 28	
	<u>Anzahl der Patienten, n (%)</u>	<u>3609 (87)</u> <u>3322 (80)</u>
	DFCB (ml)	101,00
	95% Untere Konfidenzschranke	89,83
	95% Obere Konfidenzschranke	112,17
	p-Wert	<0,0001

Studie	Ergebnisse für FEV ₁ (Talspiegel)	
Erhebungszeitraum	Baseline bis Woche 52	
<u>Anzahl der Patienten, n (%)</u>	<u>3366 (81)</u>	<u>3060 (74)</u>
DFCB (ml)	97,00	
95% Untere Konfidenzschranke	85,04	
95% Obere Konfidenzschranke	108,96	
p-Wert	<0,0001	
Singh 2016 (TRILOGY)	BDP/FF/GB	ICS/LABA
Population	ITT	
N	687	680
Erhebungszeitraum	Baseline bis Woche 12	
<u>Anzahl der Patienten, n (%)</u>	<u>660 (96)NB</u>	<u>654 (96)NB</u>
DFCB (ml)	<u>77,88</u>	<u>68,70</u>
95% Untere Konfidenzschranke	<u>52,40</u>	<u>43,22</u>
95% Obere Konfidenzschranke	<u>103,36</u>	<u>94,18</u>
p-Wert	<0,0001	
Erhebungszeitraum	Baseline bis Woche 26	
<u>Anzahl der Patienten, n (%)</u>	<u>NB642 (93)</u>	<u>NB616 (91)</u>
DFCB (ml)	81,00	
95% Untere Konfidenzschranke	52,50	
95% Obere Konfidenzschranke	109,50	
p-Wert	<0,0001	
Erhebungszeitraum	Baseline bis Woche 52	
<u>Anzahl der Patienten, n (%)</u>	<u>NB606 (88)</u>	<u>NB578 (85)</u>
DFCB (ml)	63,00	
95% Untere Konfidenzschranke	32,00	
95% Obere Konfidenzschranke	94,00	
p-Wert	<0,0001	
Statistische Methode	Die Änderungen vom Baselinewert wurden anhand der Ergebnisse eines Mixed Models für Messwiederholungen unter Berücksichtigung von Kovariablen (CTT116853 (FULFIL) und Singh 2016 (TRILOGY)) sowie eines GLM basierend auf der negativen Binomialverteilung (CTT116855 (IMPACT)) berechnet.	
DFCB=Differenz der Veränderung im Vergleich zu Baseline, NB=Nicht berichtet		
Quellen: (³² GSK, 2018; ⁷⁰ GSK, 2016; ⁷² Singh, et al., 2016; ⁷³ Clinicaltrials.gov, 2017)		

Die Differenz in der Änderung vom Baselinewert wurde für die Zeiträume 12+4, 24+4 und 52 Wochen berechnet. In der Singh 2016 (TRILOGY) betrug die Differenz in der Änderung vom Baselinewert zwischen BDP/FF/GB vs. ICS/LABA für die drei genannten Zeiträume

~~77,88~~68,70, 81,00 und 63,00 ml. In der CTT116853 (FULFIL) war diese für den Vergleich von FF/UMEC/VI mit ICS/LABA höher und betrug 148, 171 und 179 ml für 16, 28 und 52 Wochen. In der CTT116855 (IMPACT) war die Differenz in der Änderung des FEV₁ vom Baselinewert wiederum niedriger und betrug 97, 101 und 97ml für die drei Zeiträume. Alle Behandlungsgruppenunterschiede waren statistisch signifikant.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Für den Vergleich von FF/UMEC/VI vs. BDP/FF/GB wurden drei Meta-Analysen durchgeführt. Da der FEV₁ in den drei Studien sowohl für kürzere Beobachtungszeiträume (12+4, 24+4 Wochen) als auch für längere (52 Wochen) berichtet wird, wurde pro Zeitraum eine Meta-Analyse durchgeführt. In allen Meta-Analysen konnte statistische Heterogenität nachgewiesen werden; daher werden Random-Effects-Modelle verwendet. Die Differenz in den Änderungen vom Baselinewert wird als normalverteilt angenommen und der dazugehörige SE aus dem Konfidenzintervall geschätzt, falls er nicht berichtet wurde.

Tabelle 4-232: Ergebnisse der Meta-Analyse für FF/UMEC/VI; FEV₁ (Talspiegel) – Kürzere Beobachtungszeiträume von 12+4 und 24+4 Wochen, ITT-Populationen

Ergebnisse der Meta-Analyse für den indirekten Vergleich				
	Studie zu BDP/FF/GB vs ICS/LABA [a]		Studien zu FF/UMEC/VI vs. ICS/LABA [b]	
	DFCB		DFCB	
	12 Wochen	24 Wochen	12 Wochen	24 Wochen
Effektgröße (95% Konfidenzintervall)	77,88 <u>68,70</u> (52,40 <u>43,22</u> ; 103,36 <u>94,18</u>)	81,00 (52,50; 109,50)	122 (72; 171)	135,24 (66,66; 203,82)
p-Wert	<0,0001	<0,0001	<0,0001	0,0001
Varianz zwischen den Studien (τ^2)	NBB	NBB	1220,9	2364,1

Ergebnisse der Meta-Analyse für den indirekten Vergleich				
I ² (%)	NBB	NBB	93,9	96,5
p-Wert für Test auf Heterogenität	NBB	NBB	<0.0001	<0.0001
Statistische Methode	<p>ITT-Population für die Singh 2016 (TRILOGY) zu BDP/FF/GB vs ICS/LABA (12 und 26 Wochen), ITT-Population für die CTT116853 (FULFIL) zu FF/UMEC/VI vs. ICS/LABA (12 und 24 Wochen) und ITT-Population der CTT116855 (IMPACT) zu FF/UMEC/VI vs. ICS/LABA (12 und 28 Wochen).</p> <p>Die Meta-Analyse basiert auf einem Random-Effects-Model. Da Heterogenität beobachtet wurde, unterscheidet sich dies von einem Fixed-Effects-Model. Die Differenz in den Änderungen vom Baselinewert wird als normalverteilt angenommen und der dazugehörige SE aus dem Konfidenzintervall geschätzt, falls er nicht berichtet wurde.</p>			
<p>DFCB=Differenz der Veränderung im Vergleich zu Baseline, NB=Nicht berichtet</p> <p>[a] Für den Vergleich BDP/FF/GB vs. ICS/LABA wurde nur die Singh 2016 (TRILOGY) herangezogen. Die Daten der Singh 2016 (TRILOGY) wurden in Woche 12 und 26 erhoben.</p> <p>[b] Für den Vergleich FF/UMEC/VI vs. ICS/LABA wurden die Studien CTT116853 (FULFIL) und CTT116855 (IMPACT) herangezogen. Der Wert der CTT116853 (FULFIL) wurde in Woche 12 und 24 erhoben, für die CTT116855 (IMPACT) in Woche 16 und 28.</p>				

Tabelle 4-233: Ergebnisse der Meta-Analyse für FF/UMEC/VI; FEV₁ (Talspiegel) –EXT-Population CTT116853 (FULFIL) und ITT-Populationen CTT116855 (IMPACT) und Singh 2016 (TRILOGY)

Ergebnisse der Meta-Analyse für den indirekten Vergleich		
	Studie zu BDP/FF/GB vs ICS/LABA [a]	Studien zu FF/UMEC/VI vs. ICS/LABA [b]
	DFCB	DFCB
Effektgröße (95% Konfidenzintervall)	63,00 (32,00; 94,00)	134,66 (54,57; 214,75)
p-Wert	<0,0001	0,0010
Varianz zwischen den Studien (τ^2)	NBB	3050,6
I ² (%)	NBB	90,7
p-Wert für Test auf Heterogenität	NBB	0,0010
Statistische Methode	<p>ITT-Population für die Singh 2016 (TRILOGY) zu BDP/FF/GB vs ICS/LABA, EXT-Population für die CTT116853 (FULFIL) zu FF/UMEC/VI vs. ICS/LABA und ITT-Population für die CTT116855 (IMPACT) zu FF/UMEC/VI vs. ICS/LABA (Beobachtungszeiträume 52 Wochen).</p> <p>Die Meta-Analyse basiert auf einem Random-Effects-Model. Da Heterogenität beobachtet wurde, unterscheidet sich dies von einem Fixed-Effects-Model. Die Differenz in den Änderungen vom Baselinewert wird als normalverteilt angenommen und der dazugehörige SE aus dem Konfidenzintervall geschätzt, falls er nicht berichtet wurde.</p>	
DFCB=Differenz der Veränderung im Vergleich zu Baseline, NBB=Nicht berechenbar		

[a] Für den Vergleich BDP/FF/GB vs. ICS/LABA wurde nur die Singh 2016 (TRILOGY) herangezogen. Die Daten der Singh 2016 (TRILOGY) wurden in Woche 52 erhoben.

[b] Für den Vergleich FF/UMEC/VI vs. ICS/LABA wurden die Studien CTT116853 (FULFIL) und CTT116855 (IMPACT) herangezogen. Die Daten der CTT116853 (FULFIL) und CTT116855 (IMPACT) wurden in Woche 52 erhoben.

In der Meta-Analyse für die ITT-Populationen der drei Studien (kürzere Beobachtungszeiträume von 12+4 und 24+4 Wochen) zeigt die Veränderung des FEV₁ von ~~43,63~~52,81 ml und 54,24 ml eine erhöhte Atemkapazität der COPD-Patienten im Vergleich zu Baseline für FF/UMEC/VI vs. BDP/FF/GB. Der Behandlungsunterschied zwischen FF/UMEC/VI und BDP/FF/GB ist statistisch nicht signifikant ($p \geq 0.05$).

Tabelle 4-234: Ergebnisse des ITC für FF/UMEC/VI vs. BDP/FF/GB; FEV₁ (Talspiegel) – ITT-Population

Ergebnisse für den indirekten Vergleich FF/UMEC/VI vs. BDP/FF/GB nach der Bucher-Methode		
	DFCB	
	Woche 12	Woche 24
Effektgröße (95% Konfidenzintervall)	43,63 <u>52,81</u> (- 12,44 <u>3,26</u> ; 99,70 <u>108,88</u>)	54,24 (-20,03; 128,51)
p-Wert	0,127 <u>0,649</u>	0,1523
DFCB=Differenz der Veränderung im Vergleich zu Baseline Die Daten der Singh 2016 (TRILOGY) wurden in Woche 12 und 26 erhoben, die Daten der CTT116853 (FULFIL) in Woche 12 und 24 und die Daten der CTT116855 (IMPACT) in Woche 16 und 28.		

In der ITC basieren auf der EXT-Population der CTT116853 (FULFIL) und den ITT-Populationen der CTT116855 (IMPACT) und Singh 2016 (TRILOGY) (längere Beobachtungszeiträume von 52 Wochen) zeigt die Veränderung des FEV₁ von 71,66 eine erhöhte Atemkapazität der COPD-Patienten im Vergleich zu Baseline für FF/UMEC/VI vs. BDP/FF/GB. Der Behandlungsunterschied zwischen FF/UMEC/VI und BDP/FF/GB ist statistisch nicht signifikant ($p \geq 0.05$).

Tabelle 4-235: Ergebnisse des ITC für FF/UMEC/VI vs. BDP/FF/GB; FEV₁ (Talspiegel) – EXT-Population

Ergebnisse für den indirekten Vergleich FF/UMEC/VI vs. BDP/FF/GB nach der Bucher-Methode	
	DFCB
Effektgröße (95% Konfidenzintervall)	71,66 (-14,23; 157,54)
p-Wert	0,1020
DFCB=Differenz der Veränderung im Vergleich zu Baseline Die Daten der Studien wurden in Woche 52 erhoben.	