Table 5: GRADE quality of evidence assessments for efficacy outcomes for the DTG + 2 NRTIs vs LPV/r + 2 NRTIs comparison in second-line patients

Outcome	No of patients		Direct Effect	Uncombined Estimates						Combined Estimates				
	DTG + 2 NRTIs	LPV/r + 2 NRTIs		Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Quality of direct evidence	Odds ratio (95% CrI)	Absolute effects	Indirect evidence precision	Network Transitivity	Overall quality of evidence
Viral supp. at 4 weeks	208/312	75/312	**6.32 (4.45, 8.97)**	0	0	0	0	0	⊕⊕⊕⊕ High	**6.36 (4.50, 9.13)**	**384 more per 1,000** (299 to 468)	0	0	⊕⊕⊕⊕ High
Viral supp. at 12 weeks	238/312	134/312	**4.27 (3.03, 6.03)**	0	0	0	0	0	⊕⊕⊕⊕ High	**4.30 (3.06, 6.09)**	**297 more per 1,000** (241 to 348)	0	0	⊕⊕⊕⊕ High
Viral supp. at 24 weeks	257/312	215/312	**2.11 (1.45, 3.07)**	0	0	0	0	0	⊕⊕⊕⊕ High	**2.12 (1.46, 3.10)**	**113 more per 1,000** (62 to 157)	0	0	⊕⊕⊕⊕ High
Viral supp. at 48 weeks	261/312	219/312	**2.17 (1.48, 3.20)**	0	0	0	0	0	⊕⊕⊕⊕ High	**2.18 (1.49, 3.22)**	**113 more per 1,000** (63 to 155)	0	0	⊕⊕⊕⊕ High
VS for >100,000 at 48 weeks	45/70	41/63	0.97 (0.47, 1.97)	0	0	0	-1	0	⊕⊕⊕ Moderate	0.96 (0.47, 1.98)	**10 fewer per 1,000** (-185 to 161)	0	0	⊕⊕⊕ Moderate
Change in CD4 at 24	312	312	2.00 (-11.37, 15.37)	0	0	0	-1	0	⊕⊕⊕ Moderate	--	**2.27 cells/ml higher** (-11.06, 15.57)	0	0	⊕⊕⊕ Moderate
Change in CD4 at 48	312	312	2.00 (-13.49, 17.49)	0	0	0	-1	0	⊕⊕⊕ Moderate	--	**2.14 cells/ml higher** (-13.4, 17.44)	0	0	⊕⊕⊕ Moderate
Change in body weight at 48 weeks	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Mortality	1/314	3/310	0.33 (0.03, 3.16)	0	0	0	-2	0	⊕⊕ Low	0.26 (0.01, 2.41)	**12 fewer per 1,000** (-23 to 19)	0	0	⊕⊕ Low
Discontinuations	34/312	52/312	**0.61 (0.38, 0.97)**	0	0	0	-1	0	⊕⊕⊕ Moderate	**0.61 (0.38, 0.96)**	**25 fewer per 1,000** (-45 to -3)	0	0	⊕⊕⊕ Moderate
Discontinuations due to AEs	7/314	17/310	**0.39 (0.16, 0.96)**	0	0	0	-1	0	⊕⊕⊕ Moderate	**0.39 (0.14, 0.91)**	**25 fewer per 1,000** (-42 to -3)	0	0	⊕⊕⊕ Moderate
Neuropsychiatric AEs (any grade)	19/312	17/312	1.13 (0.57, 2.21)	0	0	0	-2	0	⊕⊕⊕ Moderate	1.13 (0.57, 2.26)	**6 more per 1,000** (-24 to 47)	0	0	⊕⊕⊕ Moderate
Overall resistance	2/312	3/312	0.66 (0.11, 4.00)	0	0	0	-2	0	⊕⊕ Low	0.62 (0.07, 3.87)	**10 fewer per 1,000** (-35 to 56)	0	0	⊕⊕ Low
Treatment emergent SAEs	17/314	18/310	0.93 (0.47, 1.84)	0	0	0	-2	0	⊕⊕ Low	0.92 (0.45, 1.85)	**7 fewer per 1,000** (-56 to 69)	0	0	⊕⊕ Low

Treatment emergent AEs	47/314	113/310	**0.31 (0.21, 0.45)**	0	0	0	0	0	⊕⊕⊕⊕ **High**	**0.30 (0.21, 0.45)**	**47 fewer per 1,000** (-66 to -31)	0	0	⊕⊕⊕⊕ **High**
Treatment-related SAEs	2/314	2/310	0.99 (0.14, 7.05)	0	0	0	-2	0	⊕⊕ **Low**	0.94 (0.11, 6.51)	**1 fewer per 1,000** (-27 to 77)	0	0	⊕⊕ **Low**
Treatment-related AEs	204/314	231/310	**0.63 (0.45, 0.90)**	0	0	0	0	0	⊕⊕⊕⊕ **High**	**0.63 (0.45, 0.89)**	**12 fewer per 1,000** (-28 to -3)	0	0	⊕⊕⊕⊕ **High**

n/N in square brackets where no direct comparison between interventions of interest is available and reflects the number of patients in the network.

Legend: Uncombined estimates represent either direct estimates, if available, or indirect NMA estimates otherwise. Combined estimates are NMA estimates for comparisons where direct estimates were available. For uncombined estimates start with high quality evidence. -1 symbolizes a choice to rate down (e.g. high quality to moderate quality evidence); 0 symbolizes choice to not rate down; -- = not applicable because the NMA estimate is the only estimate.

The final quality of evidence updates that of the uncombined evidence. The quality can be moved up if the uncombined score was penalized for precision, which was overcome in network estimates. It can be moved down if the estimates are no longer precise or if there is evidence of inconsistency in loops containing the comparison (i.e. violation of transitivity).

Precision – We rated down for precision if the confidence interval crossed 1.1 or 0.9 and if there were less than 50 total events. ***Consistency*** – We assessed the consistency for direct treatment comparisons using I^2 estimates and visual inspection of point estimates. An I^2 of 75% or higher indicates considerable heterogeneity. This was conducted along the shortest indirect pathway with the largest number of trials for indirect estimates. ***Risk of Bias*** – For direct estimates we rated down for risk of bias if the majority of studies within a comparison were considered to be at high risk of bias and similarly along the principal indirect pathway for indirect estimates. ***Indirectness*** – Estimates obtained solely from indirect evidence were rated down for indirectness.

Table 6: GRADE quality of evidence assessments for efficacy outcomes for the DTG + 2 NRTIs vs LPV/r + RAL comparison in second-line patients

Outcome	No of patients		Direct Effect	Uncombined Estimates						Combined Estimates				
	DTG + 2 NRTIs	LPV/r + RAL		Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Quality of direct evidence	Odds ratio (95% CrI)	Absolute effects	Indirect evidence precision	Network Transitivity	Overall quality of evidence
Viral supp. at 4 weeks	[208/312]	[330/645]	--	0	0	-1	-1	0	⊕⊕ Low	1.49 (0.94, 2.33)	**99 more per 1,000** (-13 to 207)	0	0	⊕⊕ Low
Viral supp. at 12 weeks	[238/312]	[487/648]	--	0	0	-1	0	0	⊕⊕⊕ Moderate	**1.96 (1.27, 2.94)**	**114 more per 1,000** (41 to 184)	0	0	⊕⊕⊕ Moderate
Viral supp. at 24 weeks	[257/312]	[526/641]	--	0	0	-1	-1	0	⊕⊕ Low	**1.61 (1, 2.56)**	**66 more per 1,000** (0 to 128)	0	0	⊕⊕ Low
Viral supp. at 48 weeks	[261/312]	[733/952]	--	0	0	-1	0	0	⊕⊕⊕ Moderate	**2.00 (1.30, 3.12)**	**99 more per 1,000** (39 to 153)	0	0	⊕⊕⊕ Moderate
VS for >100,000 at 48 weeks	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Change in CD4 at 24	[312]	[921]	--	0	0	-1	-1	0	⊕⊕ Low	--	**17.33 cells/ml lower** (-36.02, 1.13)	0	0	⊕⊕ Low
Change in CD4 at 48	[312]	[703]	--	0	0	-1	-1	0	⊕⊕ Low	--	24.98 cells/ml lower (-46.43, -2.83)	0	0	⊕⊕ Low
Change in body weight at 48 weeks	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Mortality	[1/314]	[77/1394]	--	0	0	-1	-2	0	⊕ **Very low**	0.34 (0.01, 3.23)	**8 fewer per 1,000** (-20 to 23)	0	0	⊕ **Very low**
Discontinuations	[34/312]	[72/961]	--	0	0	-1	-1	0	⊕⊕ Low	0.75 (0.43, 1.33)	**14 fewer per 1,000** (-39 to 15)	0	0	⊕⊕ Low
Discontinuations due to AEs	[7/314]	[255/1394]	--	0	0	-1	-1	0	⊕⊕ Low	0.39 (0.14, 0.98)	**24 fewer per 1,000** (-44 to -1)	0	0	⊕⊕ Low
Neuropsychiatric AEs (any grade)	[19/312]	[8/270]	--	0	0	-1	-1	0	⊕⊕ Low	1.75 (0.56, 5.88)	**22 more per 1,000** (-25 to 71)	0	0	⊕⊕ Low
Overall resistance	[2/312]	[17/390]	--	0	0	-1	-2	0	⊕ **Very low**	0.70 (0.08, 5)	**7 fewer per 1,000** (-36 to 60)	0	0	⊕ **Very low**
Treatment emergent SAEs	[17/314]	[255/1394]	--	0	0	-1	-1	0	⊕⊕ Low	1.08 (0.5, 2.27)	**6 more per 1,000** (-48 to 86)	0	0	⊕⊕ Low

Treatment emergent AEs	[204/314]	[528/528]	--	0	0	-1	-1	0	⊕⊕ **Low**	0.99 (0.91, 1.07	**1 fewer per 1,000** (-18 to 17)	0	0	⊕⊕ **Low**
Treatment-related SAEs	[2/314]	[22/433]	--	0	0	-1	-2	0	⊕ **Very low**	1.15 (0.12, 9.09)	**2 more per 1,000** (-25 to 80)	0	0	⊕ **Very low**
Treatment-related AEs	[47/314]	[20/703]	--	0	0	-1	-1	0	⊕⊕ **Low**	0.48 (0.24, 0.96)	**23 fewer per 1,000** (-54 to -1)	0	0	⊕⊕ **Low**

n/N in square brackets where no direct comparison between interventions of interest is available and reflects the number of patients in the network.

Legend: Uncombined estimates represent either direct estimates, if available, or indirect NMA estimates otherwise. Combined estimates are NMA estimates for comparisons where direct estimates were available. For uncombined estimates start with high quality evidence. -1 symbolizes a choice to rate down (e.g. high quality to moderate quality evidence); 0 symbolizes choice to not rate down; -- = not applicable because the NMA estimate is the only estimate.

The final quality of evidence updates that of the uncombined evidence. The quality can be moved up if the uncombined score was penalized for precision, which was overcome in network estimates. It can be moved down if the estimates are no longer precise or if there is evidence of inconsistency in loops containing the comparison (i.e. violation of transitivity).

Precision – We rated down for precision if the confidence interval crossed 1.1 or 0.9 and if there were less than 50 total events. ***Consistency*** – We assessed the consistency for direct treatment comparisons using I^2 estimates and visual inspection of point estimates. An I^2 of 75% or higher indicates considerable heterogeneity. This was conducted along the shortest indirect pathway with the largest number of trials for indirect estimates. ***Risk of Bias*** – For direct estimates we rated down for risk of bias if the majority of studies within a comparison were considered to be at high risk of bias and similarly along the principal indirect pathway for indirect estimates. ***Indirectness*** – Estimates obtained solely from indirect evidence were rated down for indirectness.

Table 7: GRADE quality of evidence assessments for efficacy outcomes for the DTG + 2 NRTIs vs DRV/r + 2 NRTIs comparison in second-line patients

Outcome	No of patients		Direct Effect	Uncombined Estimates						Combined Estimates				
	DTG + 2 NRTIs	DRV/r + 2 NRTIs		Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Quality of direct evidence	Odds ratio (95% CrI)	Absolute effects	Indirect evidence precision	Network Transitivity	Overall quality of evidence
Viral supp. at 4 weeks	[208/312]	[3/154]	--	0	0	-1	0	0	⊕⊕⊕ Moderate	**12.5 (3.45, 50)**	**449 more per 1,000** (268 to 573)	0	0	⊕⊕⊕ Moderate
Viral supp. at 12 weeks	[238/312]	[54/154]	--	0	0	-1	0	0	⊕⊕⊕ Moderate	**3.23 (1.89, 5.56)**	**224 more per 1,000** (119 to 334)	0	0	⊕⊕⊕ Moderate
Viral supp. at 24 weeks	[257/312]	[97/154]	--	0	0	-1	-1	0	⊕⊕ Low	1.69 (0.98, 2.94)	**73 more per 1,000** (-3 to 155)	0	0	⊕⊕ Low
Viral supp. at 48 weeks	[261/312]	[97/154]	--	0	0	-1	0	0	⊕⊕⊕ Moderate	**2.56 (1.45, 4.35)**	**141 more per 1,000** (54 to 235)	0	0	⊕⊕⊕ Moderate
VS for >100,000 at 48 weeks	[45/70]	[10/44]	--	0	0	-1	-1	0	⊕⊕ Low	2.86 (0.98, 9.09)	**249 more per 1,000** (-6 to 479)	0	0	⊕⊕ Low
Change in CD4 at 24	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Change in CD4 at 48	[312]	[149]	--	0	0	-1	-1	0	⊕⊕ Low	--	**21.14 cells/ml higher** (-2.83, 45.36)	0	0	⊕⊕ Low
Change in body weight at 48 weeks	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Mortality	[1/314]	[3/154]	--	0	0	-1	-2	0	⊕ Very low	0.13 (0, 2.33)	**26 fewer per 1,000** (-146 to 13)	0	0	⊕ Very low
Discontinuations	[34/312]	[2/154]	--	0	0	-1	-1	0	⊕⊕ Low	1.89 (0.41, 14.29)	**19 more per 1,000** (-47 to 53)	0	0	⊕⊕ Low
Discontinuations due to AEs	[7/314]	[1/154]	--	0	0	-1	-2	0	⊕ Very low	1.05 (0.1, 33.33)	**1 more per 1,000** (-88 to 28)	0	0	⊕ Very low
Neuropsychiatric AEs (any grade)	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Overall resistance	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Treatment emergent SAEs	[17/314]	[19/154]	--	0	0	-1	-1	0	⊕⊕ Low	0.88 (0.35, 2.22)	**11 fewer per 1,000** (-96 to 81)	0	0	⊕⊕ Low
Treatment emergent AEs	[204/314]	[118/154]	--	0	0	-1	-1	0	⊕⊕ Low	1.01 (0.93, 1.12)	**1 more per 1,000** (-16 to 17)	0	0	⊕⊕ Low

Treatment-related SAEs	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Treatment-related AEs	--	--	--	--	--	--	--	--	--	--	--	--	--	--

n/N in square brackets where no direct comparison between interventions of interest is available and reflects the number of patients in the network.
Legend: Uncombined estimates represent either direct estimates, if available, or indirect NMA estimates otherwise. Combined estimates are NMA estimates for comparisons where direct estimates were available. For uncombined estimates start with high quality evidence. -1 symbolizes a choice to rate down (e.g. high quality to moderate quality evidence); 0 symbolizes choice to not rate down; -- = not applicable because the NMA estimate is the only estimate.
The final quality of evidence updates that of the uncombined evidence. The quality can be moved up if the uncombined score was penalized for precision, which was overcome in network estimates. It can be moved down if the estimates are no longer precise or if there is evidence of inconsistency in loops containing the comparison (i.e. violation of transitivity).

Precision – We rated down for precision if the confidence interval crossed 1.1 or 0.9 and if there were less than 50 total events. ***Consistency*** – We assessed the consistency for direct treatment comparisons using I^2 estimates and visual inspection of point estimates. An I^2 of 75% or higher indicates considerable heterogeneity. This was conducted along the shortest indirect pathway with the largest number of trials for indirect estimates. ***Risk of Bias*** – For direct estimates we rated down for risk of bias if the majority of studies within a comparison were considered to be at high risk of bias and similarly along the principal indirect pathway for indirect estimates. ***Indirectness*** – Estimates obtained solely from indirect evidence were rated down for indirectness.