Appendix 5_GRADE tables

Review question 1

1. **Interferon compared with placebo for clinically isolated syndrome**

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Interferon	Placebo	Relative (95% CI)	Absolute		
Time to conversion to CDMS (104 weeks' follow-up) (follow-up mean 104 weeks)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	-	-	HR 0.49 (0.38 to 0.64)	-	⊕⊕⊕⊕ HIGH	CRITICAL
								0%		-		
Conversion to CDMS (follow-up 104-156 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	152/363 (41.9%)	217/360 (60.3%)	RR 0.71 (0.61 to 0.82)	175 fewer per 1000 (from 109 fewer to 235 fewer)	⊕⊕⊕O MODERATE	CRITICAL
								61.6%		179 fewer per 1000 (from 111 fewer to 240 fewer)		
New GAD lesions (number of patients free) (follow-up mean 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	65/171 (38%)	32/171 (18.7%)	RR 2.03 (1.41 to 2.93)	193 more per 1000 (from 77 more to 361 more)	⊕⊕⊕O MODERATE	CRITICAL
								18.7%		193 more per 1000 (from 77 more to 361 more)		
GAD lesions (mean number) (78 weeks' follow-up) (follow-up mean 78 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	134	114	-	MD 1 lower (1.71 to 0.29 lower)	⊕⊕OO LOW	CRITICAL
New T2 lesions (number of patients free) (follow-up mean 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	120/171 (70.2%)	50/171 (29.2%)	RR 2.4 (1.86 to 3.09)	409 more per 1000 (from 251 more to 611 more)	⊕⊕⊕O MODERATE	CRITICAL
								29.2%		409 more per 1000 (from 251 more to 610 more)		
T2 new or newly enlarging lesions (mean number) (78 weeks' follow-up) (follow-up mean 78 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	132	119	-	MD 2.9 lower (4.39 to 1.41 lower)	⊕⊕OO LOW	CRITICAL
Change in T2 lesion volume (follow-up mean 104 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	292	176	-	MD 456.9 lower (959.46 lower to 45.66 higher)	⊕⊕⊕O MODERATE	CRITICAL
Cumulative number of newly active lesions (mean number) (follow-up mean 104 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	292	176	-	MD 4.8 lower (7.06 to 2.54 lower)	⊕⊕⊕⊕ HIGH	CRITICAL
Combined unique active lesions (mean number per patient per scan) (follow-up mean 104 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	171	171	-	MD 2.1 lower (2.9 to 1.3 lower)	⊕⊕⊕O MODERATE	CRITICAL
Discontinuation due to any reason (follow-up 104-156 weeks)												

3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	72/656 (11%)	57/537 (10.6%)	RR 1.11 (0.8 to 1.54)	12 more per 1000 (from 21 fewer to 57 more)	⊕⊕⊕O MODERATE	CRITICAL
								11.7%		13 more per 1000 (from 23 fewer to 63 more)		
Discontinuation due to side effects (follow-up 104-156 weeks)												
2	randomised trials	no serious risk of bias	serious[4]	no serious indirectness	serious[2]	none	13/463 (2.8%)	6/347 (1.7%)	RR 2.17 (0.16 to 28.82)	20 more per 1000 (from 15 fewer to 481 more)	⊕⊕OO LOW	CRITICAL
								1.8%		21 more per 1000 (from 15 fewer to 501 more)		
Discontinuation of study drug due to side effects (follow-up 104-156 weeks)												
2	randomised trials	no serious risk of bias	very serious[5]	no serious indirectness	serious[2]	none	25/485 (5.2%)	8/366 (2.2%)	RR 0.98 (0.87 to 1.09)	0 fewer per 1000 (from 3 fewer to 2 more)	⊕OOO VERY LOW	CRITICAL
								2.1%		0 fewer per 1000 (from 3 fewer to 2 more)		
Discontinuation of study drug due to any reason (follow-up 104-156 weeks)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	82/485 (16.9%)	53/366 (14.5%)	RR 1.21 (0.88 to 1.67)	30 more per 1000 (from 17 fewer to 97 more)	⊕⊕⊕O MODERATE	CRITICAL
								14.3%		30 more per 1000 (from 17 fewer to 96 more)		
Mortality (risk of non-event) (follow-up 104-156 weeks)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	1/364 (0.27%)	2/361 (0.55%)	RR 1 (0.99 to 1.02)	0 fewer per 1000 (from 0 fewer to 0 more)	⊕⊕⊕O MODERATE	CRITICAL
								0.6%		0 fewer per 1000 (from 0 fewer to 0 more)		
Cognitive performance (PASAT-3") (follow-up mean 104 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[6]	none	273	166	-	MD 1.4 higher (0.29 to 2.51 higher)	⊕⊕⊕O MODERATE	IMPORTANT

[1] Unclear allocation concealment and risk of selective outcome reporting (Jacobs 2000)
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[3] Confidence intervals include a null effect and appreciable benefit
[4] Substantial heterogeneity (I2=67%)
[5] Substantial and significant heterogeneity (I2=96%; p<0.00001)
[6] Confidence intervals include a negligible effect and appreciable benefit

2. Glatiramer acetate compared with placebo for clinically isolated syndrome

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Glatiramer acetate**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Time to conversion to CDMS (follow-up median 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	-	-	OR 0.55 (0.4 to 0.76)	-	□□□□ MODERATE	CRITICAL
								0%		-		
Discontinuation due to any reason (follow-up median 156 weeks)												

1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	39/243 (16%)	23/238 (9.7%)	RR 1.66 (1.02 to 2.69)	64 more per 1000 (from 2 more to 163 more)	⨁⨁⨁◯ MODERATE	IMPORTANT
								9.7%		64 more per 1000 (from 2 more to 164 more)		
Discontinuation due to side effects (follow-up median 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	14/243 (5.8%)	4/238 (1.7%)	RR 3.43 (1.14 to 10.26)	41 more per 1000 (from 2 more to 156 more)	⨁⨁◯◯ LOW	IMPORTANT
								1.7%		41 more per 1000 (from 2 more to 157 more)		

[1] Unclear risk of detection bias and unclear allocation concealment.

[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

3. Teriflunomide compared with placebo for clinically isolated syndrome

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Teriflunomide	Placebo	Relative (95% CI)	Absolute		
Time to conversion to CDMS (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	38/214 (17.8%)	55/197 (27.9%)	HR 0.57 (0.38 to 0.87)	109 fewer per 1000 (from 31 fewer to 162 fewer)	⊕⊕OO LOW	CRITICAL
								28.3%		110 fewer per 1000 (from 32 fewer to 164 fewer)		
Conversion to CDMS (number of participants) (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	38/214 (17.8%)	55/197 (27.9%)	RR 0.64 (0.44 to 0.92)	101 fewer per 1000 (from 22 fewer to 156 fewer)	⊕⊕OO LOW	CRITICAL
								28.3%		102 fewer per 1000 (from 23 fewer to 158 fewer)		
Disability progression (number of participants)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	16/214 (7.5%)	10/99 (10.1%)	RR 0.74 (0.35 to 1.57)	26 fewer per 1000 (from 66 fewer to 58 more)	⊕⊕OO LOW	CRITICAL
								10.1%		26 fewer per 1000 (from 66 fewer to 58 more)		
Atrophy (mean change from baseline) (follow-up mean 108 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	99	68	-	MD 0 higher (0.01 lower to 0.01 higher)	⊕⊕OO LOW	CRITICAL
GAD lesions (mean number of lesions per MRI scan) (follow-up mean 108 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	74	110	-	MD 0.56 lower (1.17 lower to 0.06 higher)	⊕⊕OO LOW	CRITICAL
T2 lesion component (volume) (mean change from baseline) (follow-up mean 108 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	99	68	-	MD 0.07 lower (0.21 lower to 0.06 higher)	⊕⊕OO LOW	CRITICAL
Discontinuation of study drug due to any reason (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	51/216 (23.6%)	56/197 (28.4%)	RR 0.83 (0.6 to 1.15)	48 fewer per 1000 (from 114 fewer to 43 more)	⊕⊕OO LOW	CRITICAL
								28.3%		48 fewer per 1000 (from 113 fewer to 42 more)		
Discontinuation of study drug due to side effects (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	18/216 (8.3%)	18/197 (9.1%)	RR 0.91 (0.49 to 1.7)	8 fewer per 1000 (from 47 fewer to 64 more)	⊕⊕OO LOW	CRITICAL
								9.1%		8 fewer per 1000 (from 46 fewer to 64 more)		
Infection (number of participants) (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	97/216 (44.9%)	77/191 (39.3%)	RR 1.11 (0.89 to 1.4)	43 more per 1000 (from 43 fewer to 157 more)	⊕⊕OO	IMPORTANT

								39.4%		43 more per 1000 (from 43 fewer to 158 more)	LOW	
Serious infection (number of participants) (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	7/216 (3.2%)	1/191 (1%)	RR 3.09 (0.65 to 14.72)	22 more per 1000 (from 4 fewer to 144 more)	⊕⊕OO LOW	CRITICAL
								2%		42 more per 1000 (from 7 fewer to 274 more)		
Mortality (follow-up mean 108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/216 (0%)	1/197 (0.51%)	RR 1.01 (0.99 to 1.02)	0 more per 1000 (from 0 fewer to 0 more)	⊕⊕OO LOW	CRITICAL
								1%		0 more per 1000 (from 0 fewer to 0 more)		

[1] High risk of bias due to incomplete outcome data
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

Review question 2

1. Interferon compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Interferon**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Relapse free (number of participants) (follow-up 48 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	422/512 (82.4%)	358/500 (71.6%)	RR 1.15 (1.08 to 1.23)	107 more per 1000 (from 57 more to 165 more)	□□□□ MODERATE	CRITICAL
								71.6%		107 more per 1000 (from 57 more to 165 more)		
Relapse free (number of participants) (follow-up 104 weeks)												
3	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	178/573 (31.1%)	71/387 (18.3%)	RR 1.73 (1.35 to 2.21)	134 more per 1000 (from 64 more to 222 more)	□□□□ LOW	CRITICAL
								16%		117 more per 1000 (from 56 more to 194 more)		
Relapse free (number of participants) - 156 weeks FU (follow-up 156 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	27/124 (21.8%)	17/123 (13.8%)	RR 1.58 (0.91 to 2.74)	80 more per 1000 (from 12 fewer to 240 more)	□□□□ LOW	CRITICAL
								13.8%		80 more per 1000 (from 12 fewer to 240 more)		
Annualised relapse rate (follow-up 48-104 weeks; Better indicated by lower values)												
2	randomised	serious[5]	no serious	no serious	no serious	None	959	950	-	MD 0.1 lower (0.16 to 0.04	□□□□	CRITICAL

	trials		inconsistency	indirectness	imprecision					lower)	MODERATE	
Disability progression confirmed at 3 months (number of participants worsened) (follow-up 48 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	31/512 (6.1%)	50/500 (10%)	RR 0.61 (0.39 to 0.93)	39 fewer per 1000 (from 7 fewer to 61 fewer)	⨁◯◯◯ LOW	CRITICAL
								10%		39 fewer per 1000 (from 7 fewer to 61 fewer)		
Disability progression confirmed at 6 months (number of participants worsened) (follow-up 104 weeks)												
2	randomised trials	serious[6]	no serious inconsistency	no serious indirectness	serious[3]	none	53/532 (10%)	75/537 (14%)	RR 0.71 (0.51 to 0.98)	41 fewer per 1000 (from 3 fewer to 68 fewer)	⨁◯◯◯ LOW	CRITICAL
								21.8%		63 fewer per 1000 (from 4 fewer to 107 fewer)		
Disability progression (number of participants worsened) (follow-up 156 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	33/124 (26.6%)	48/123 (39%)	RR 0.68 (0.47 to 0.98)	125 fewer per 1000 (from 8 fewer to 207 fewer)	⨁◯◯◯ LOW	CRITICAL
								39%		125 fewer per 1000 (from 8 fewer to 207 fewer)		
Discontinuation due to side effects - 48 weeks FU (follow-up 48 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	24/512 (4.7%)	5/500 (1%)	RR 4.69 (1.8 to 12.19)	37 more per 1000 (from 8 more to 112 more)	⨁⨁◯◯ MODERATE	CRITICAL
								1%		37 more per 1000 (from 8 more to 112 more)		
Discontinuation due to any reason (follow-up 48 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	74/512 (14.5%)	44/500 (8.8%)	RR 1.64 (1.15 to 2.34)	56 more per 1000 (from 13 more to 118 more)	⨁⨁◯◯ MODERATE	CRITICAL
								8.8%		56 more per 1000 (from 13 more to 118 more)		
Discontinuation due to side effects (follow-up 104 weeks)												
3	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	48/905 (5.3%)	23/725 (3.2%)	RR 1.72 (1.04 to 2.86)	23 more per 1000 (from 1 more to 59 more)	⨁◯◯◯ LOW	CRITICAL
								1.7%		12 more per 1000 (from 1 more to 32 more)		
Discontinuation due to any reason (follow-up 104 weeks)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	110/820 (13.4%)	109/638 (17.1%)	RR 0.84 (0.65 to 1.07)	27 fewer per 1000 (from 60 fewer to 12 more)	⨁⨁◯◯ MODERATE	CRITICAL
								9.6%		15 fewer per 1000 (from 34 fewer to 7 more)		
Discontinuation due to side effects (follow-up 156 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	10/124 (8.1%)	2/123 (1.6%)	RR 4.96 (1.11 to 22.17)	64 more per 1000 (from 2 more to 344 more)	⨁◯◯◯ LOW	CRITICAL

								1.6%		63 more per 1000 (from 2 more to 339 more)		
Discontinuation due to any reason (follow-up 156 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	23/124 (18.5%)	24/123 (19.5%)	RR 0.95 (0.57 to 1.59)	10 fewer per 1000 (from 84 fewer to 115 more)	□□□□ LOW	CRITICAL
								19.5%		10 fewer per 1000 (from 84 fewer to 115 more)		
Lesion volume (mm3) (follow-up 156 weeks; Better indicated by lower values)												
1	randomised trials	serious[4,7]	no serious inconsistency	no serious indirectness	serious[3]	none	134	123	-	MD 26.5 lower (90.6 lower to 37.6 higher)	□□□□ LOW	CRITICAL
Lesion volume (mm3) (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	serious[7]	no serious inconsistency	no serious indirectness	serious[3]	none	82	82	-	MD 48.3 lower (169.42 lower to 72.82 higher)	□□□□ LOW	CRITICAL
New or newly enlarging T2 lesions (mean number) (follow-up 48 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	457	476	-	MD 7.3 lower (8.85 to 5.75 lower)	□□□□ MODERATE	CRITICAL
T2 active lesions (number of participants with no activity) (follow-up 104 weeks)												
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	serious[3]	none	91/367 (24.8%)	16/184 (8.7%)	RR 2.8 (1.69 to 4.63)	157 more per 1000 (from 60 more to 316 more)	□□□□ LOW	CRITICAL
								8.7%		157 more per 1000 (from 60 more to 316 more)		
Combined unique active lesions (number of participants with no activity) (follow-up 104 weeks)												
2	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	serious[3]	none	48/132 (36.4%)	8/66 (12.1%)	RR 2.97 (1.49 to 5.92)	239 more per 1000 (from 59 more to 596 more)	□□□□ LOW	CRITICAL
								12.1%		238 more per 1000 (from 59 more to 595 more)		
Percent brain volume change (follow-up 48 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	512	500	-	MD 0.1 lower (0.2 lower to 0 higher)	□□□□ MODERATE	CRITICAL
Percent brain volume change (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	serious[9]	no serious inconsistency	no serious indirectness	no serious imprecision	none	447	450	-	MD 0.11 lower (0.28 lower to 0.06 higher)	□□□□ MODERATE	CRITICAL
Cumulative number of GdE lesions at months 12 and 24 (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	serious[9]	no serious inconsistency	no serious indirectness	no serious imprecision	none	447	450	-	MD 1.44 lower (1.97 to 0.91 lower)	□□□□ MODERATE	CRITICAL
Cumulative number of new/enlarged T2 lesions at months 12 and 24 (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	serious[9]	no serious inconsistency	no serious indirectness	no serious imprecision	none	447	450	-	MD 6.66 lower (9.04 to 4.28 lower)	□□□□ MODERATE	CRITICAL

[1] Unclear risk of detection bias

[2] Unclear risk of randomisation sequence generation (IFNB MS Group 1993). Unclear allocation concealment (IFNB MS Group 1993 and Jacobs 1996). Unclear risk of detection bias (IFNB MS Group 1993). Unclear risk of selective outcome reporting (all studies).

[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

[4] Method of randomisation sequence generation and allocation concealment unclear. Unclear risk of detection bias and selective outcome reporting.
[5] Unclear detection bias (Calabresi 2014). Unclear risk of performance bias - interferon was not blinded (Vollmer 2014)
[6] Unclear risk of performance bias (Vollmer 2014). Unclear risk of detection bias (Calabresi 2014)
[7] Unclear allocation concealment. Unclear risk of selective outcome reporting.
[8] Unclear risk of selective outcome reporting
[9] Unclear risk of performance bias

2. Glatiramer acetate compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Glatiramer acetate**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Relapse free (number of participants) (follow-up 52-104 weeks)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	1006/1418 (70.9%)	550/950 (57.9%)	RR 1.17 (1.1 to 1.24)	98 more per 1000 (from 58 more to 139 more)	□□□□ MODERATE	CRITICAL
								59%		100 more per 1000 (from 59 more to 142 more)		
Annualised relapse rate (follow-up 52-96 weeks; Better indicated by lower values)												
2	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	1293	824	-	MD 0.14 lower (0.21 to 0.06 lower)	□□□□ MODERATE	CRITICAL
Disability progression (number of participants worsened) (follow-up 96-104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	82/475 (17.3%)	98/489 (20%)	RR 0.86 (0.66 to 1.11)	28 fewer per 1000 (from 68 fewer to 22 more)	□□□□ LOW	CRITICAL
								22.8%		32 fewer per 1000 (from 78 fewer to 25 more)		
Discontinuation due to any reason (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	84/943 (8.9%)	31/461 (6.7%)	RR 1.32 (0.89 to 1.97)	22 more per 1000 (from 7 fewer to 65 more)	□□□□ MODERATE	CRITICAL
								6.7%		21 more per 1000 (from 7 fewer to 65 more)		
Discontinuation due to any reason (follow-up 06-104 weeks)												
2	randomised trials					none	87/485 (17.9%)	102/489 (20.9%)	RR 0.86 (0.66 to 1.11)	29 fewer per 1000 (from 71 fewer to 23 more)		CRITICAL
								18.5%		26 fewer per 1000 (from 63 fewer to 20 more)		
Discontinuation due to side effects (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	10/360 (2.8%)	11/363 (3%)	RR 0.92 (0.39 to 2.13)	2 fewer per 1000 (from 18 fewer to 34 more)	□□□□ MODERATE	CRITICAL

								3%		2 fewer per 1000 (from 18 fewer to 34 more)		
Discontinuation due to side effects (follow-up 96-104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	34/1068 (3.2%)	7/587 (1.2%)	RR 2.63 (1.17 to 5.9)	19 more per 1000 (from 2 more to 58 more)	□□□□ LOW	CRITICAL
								1.1%		18 more per 1000 (from 2 more to 54 more)		
New or newly enlarged T2 lesions (mean number) (follow-up 96 weeks; Better indicated by lower values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	153	139	-	MD 9.4 lower (14.26 to 4.54 lower)	□□□□ LOW	CRITICAL
GAD lesions (mean number) (follow-up 96 weeks; Better indicated by lower values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	161	144	-	MD 1.3 lower (2.26 to 0.34 lower)	□□□□ LOW	CRITICAL
Relapse free (number of participants) (follow-up 128 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	42/125 (33.6%)	31/126 (24.6%)	RR 1.37 (0.92 to 2.02)	91 more per 1000 (from 20 fewer to 251 more)	□□□□ LOW	CRITICAL
								24.6%		91 more per 1000 (from 20 fewer to 251 more)		
Disability progression (number of participants worsened) (follow-up 128 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	29/125 (23.2%)	37/126 (29.4%)	RR 0.79 (0.52 to 1.2)	62 fewer per 1000 (from 141 fewer to 59 more)	□□□□ LOW	CRITICAL
								29.4%		62 fewer per 1000 (from 141 fewer to 59 more)		
Discontinuation due to any reason (follow-up 128 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	23/125 (18.4%)	29/126 (23%)	RR 0.8 (0.49 to 1.3)	46 fewer per 1000 (from 117 fewer to 69 more)	□□□□ LOW	CRITICAL
								23%		46 fewer per 1000 (from 117 fewer to 69 more)		
Cumulative gad-e T1 lesions at months 6 and 12 (mean) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	884	441	-	MD 0.73 lower (1.15 to 0.31 lower)	□□□□ HIGH	CRITICAL
Cumulative new or newly enlarging T2 lesions at months 6 and 12 (mean) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	884	441	-	MD 1.94 lower (3.03 to 0.85 lower)	□□□□ HIGH	CRITICAL
Percentage change in brain volume from baseline to month 12(mean) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	840	423	-	MD 0.07 lower (0.19 lower to 0.06 higher)	□□□□ HIGH	CRITICAL

[1] High risk of performance bias and attrition bias (different reasons for drop-out across groups) (Fox 2012). Unclear risk of selection bias and reporting bias (no protocol available) (Johnson 1995).
[2] High risk of performance bias and attrition bias (different reasons for drop-out across groups) (Fox 2012).
[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[4] Unclear risk of selection bias and reporting bias (no protocol available) (Johnson 1995).

3. Teriflunomide compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Teriflunomide	Placebo	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 48-108 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	484/728 (66.5%)	400/751 (53.3%)	RR 1.25 (1.16 to 1.36)	133 more per 1000 (from 85 more to 192 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								53%		132 more per 1000 (from 85 more to 191 more)		
Annualised relapse rate (follow-up 48-108 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	728	752	-	MD 0.18 lower (0.24 to 0.11 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Disability progression (number of participants worsened) (follow-up 104-108 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	130/728 (17.9%)	175/751 (23.3%)	RR 0.76 (0.62 to 0.93)	56 fewer per 1000 (from 16 fewer to 89 fewer)	⨁⨁⨁◯ MODERATE	CRITICAL
								23.4%		56 fewer per 1000 (from 16 fewer to 89 fewer)		
Mortality (risk of non-event) (follow-up 48 weeks)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	2/372 (0.5%)	1/389 (0.3%)	RR 1 (0.99 to 1.01)	0 fewer per 1000 (from 0 fewer to 0 more)	⨁◯◯◯ LOW	CRITICAL
								0.3%		0 fewer per 1000 (from 0 fewer to 0 more)		
Discontinuation due to side effects (follow-up 48-108 weeks)												
2	randomised trials	serious[1]	serious[4]	no serious indirectness	serious[3]	none	96/730 (13.2%)	55/752 (7.3%)	RR 1.77 (1.02 to 3.07)	56 more per 1000 (from 1 more to 151 more)	⨁◯◯◯ VERY LOW	CRITICAL
								7.3%		56 more per 1000 (from 1 more to 151 more)		
Discontinuation due to any reason (follow-up 48-108 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	221/730 (30.3%)	229/752 (30.5%)	RR 1 (0.86 to 1.16)	0 fewer per 1000 (from 43 fewer to 49 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								30.4%		0 fewer per 1000 (from 43 fewer to 49 more)		
GAD lesions (estimated mean change) (Better indicated by lower values)												
1	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	no serious imprecision	none	358	363	-	MD 1.07 lower (1.4 to 0.74 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Total lesion volume (change from baseline) (Better indicated by lower values)												

1	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	no serious imprecision	none	358	363	-	MD 1.49 lower (2.56 to 0.42 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Patients free from enhanced lesions (follow-up 108 weeks)												
1	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	no serious imprecision	none	230/359 (64.1%)	144/363 (39.7%)	RR 1.62 (1.39 to 1.87)	246 more per 1000 (from 155 more to 345 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								39.7%		246 more per 1000 (from 155 more to 345 more)		
Risk of not having cancer (number of participants with any neoplasm) (follow-up 48-108 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	4/730 (0.5%)	5/752 (0.7%)	RR 1 (0.99 to 1.01)	0 fewer per 1000 (from 0 fewer to 0 more)	⨁◯◯◯ LOW	CRITICAL
								0.7%		0 fewer per 1000 (from 0 fewer to 0 more)		
Risk of infection (number of participants with any infection) (follow-up 48-108 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	226/730 (31%)	277/752 (36.8%)	RR 0.85 (0.75 to 0.98)	55 fewer per 1000 (from 7 fewer to 92 fewer)	⨁⨁⨁◯ MODERATE	CRITICAL
								36.3%		54 fewer per 1000 (from 7 fewer to 91 fewer)		

[1] High risk of attrition bias (30% lost to follow-up with different reasons for drop out) (Confavreux 2014). Allocation concealment unclear (O'Conner 2011)
[2] High risk of attrition bias (30% lost to follow-up with different reasons for drop out)
[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[4] Substantial heterogeneity (I2=63%)
[5] Unclear allocation concealment

4. Dimethyl fumarate compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Dimethyl fumarate	Placebo	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 104 weeks)												
2	randomised trials	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	554/769 (72%)	434/771 (56.3%)	RR 1.28 (1.14 to 1.43)	158 more per 1000 (from 79 more to 242 more)	□□□□ LOW	CRITICAL
								56.4%		158 more per 1000 (from 79 more to 243 more)		
Annualised relapse rate (follow-up 104 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	771	771	-	MD 0.19 lower (0.25 to 0.13 lower)	□□□□ MODERATE	CRITICAL
Disability progression (number of participants worsened) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	112/768 (14.6%)	172/771 (22.3%)	RR 0.66 (0.51 to 0.85)	76 fewer per 1000 (from 33 fewer to 109 fewer)	□□□□ LOW	CRITICAL
								22%		75 fewer per 1000 (from 33 fewer to 108 fewer)		
Discontinuation due to side effects (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	126/773 (16.3%)	130/773 (16.8%)	RR 0.97 (0.78 to 1.21)	5 fewer per 1000 (from 37 fewer to 35 more)	□□□□ LOW	CRITICAL
								16.7%		5 fewer per 1000 (from 37 fewer to 35 more)		
Mortality (follow-up 104 weeks)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	0/773 (0%)	1/773 (0.1%)	RR 1 (1 to 1)	-	□□□□ MODERATE	CRITICAL
								0.1%		-		
Discontinuation due to any reason (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	170/773 (22%)	176/773 (22.8%)	RR 0.97 (0.8 to 1.16)	7 fewer per 1000 (from 46 fewer to 36 more)	□□□□ LOW	CRITICAL
								22.8%		7 fewer per 1000 (from 46 fewer to 36 more)		
GAD lesions (mean number) (follow-up 104 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	299	309	-	MD 1.64 lower (2.17 to 1.1 lower)	□□□□ MODERATE	CRITICAL
New or newly enlarged T2 lesions (mean number) (follow-up 104 weeks; Better indicated by lower values)												

2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	292	304	-	MD 13.36 lower (16.63 to 10.09 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Risk of not having cancer (number of participants with any neoplasm) (follow-up 104 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	2/410 (0.5%)	2/408 (0.5%)	RR 1 (0.99 to 1.01)	0 fewer per 1000 (from 0 fewer to 0 more)	⨁⨁◯◯ LOW	CRITICAL
								0.5%		0 fewer per 1000 (from 0 fewer to 0 more)		
Risk of serious infection (number of participants with any infection) (follow-up 104 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	10/410 (2.4%)	7/408 (1.7%)	RR 1.42 (0.55 to 3.7)	7 more per 1000 (from 8 fewer to 46 more)	⨁⨁◯◯ LOW	CRITICAL
								1.7%		7 more per 1000 (from 8 fewer to 46 more)		
Risk of infection (number of participants with any infection) (follow-up 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	88/359 (24.5%)	77/363 (21.2%)	RR 1.16 (0.88 to 1.51)	34 more per 1000 (from 25 fewer to 108 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								21.2%		34 more per 1000 (from 25 fewer to 108 more)		

[1] High risk of attrition bias (different reasons for loss to follow-up between groups). Allocation concealment unclear (Fox 2012).
[2] Substantial heterogeneity (I2=55%)
[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met
[4] High risk of attrition bias (different reasons for loss to follow-up between groups).

5. Fingolimod compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Fingolimod**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Relapse free (number of participants) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	555/783 (70.9%)	378/773 (48.9%)	RR 1.44 (1.28 to 1.63)	215 more per 1000 (from 137 more to 308 more)	□□□□ MODERATE	CRITICAL
								49.2%		216 more per 1000 (from 138 more to 310 more)		
Disability progression (number of participants worsened) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	102/783 (13%)	142/773 (18.4%)	RR 0.71 (0.56 to 0.9)	53 fewer per 1000 (from 18 fewer to 81 fewer)	□□□□ LOW	CRITICAL
								18.3%		53 fewer per 1000 (from 18 fewer to 81 fewer)		
Annualised relapse rate (follow-up 104 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	783	855	-	MD 0.21 lower (0.25 to 0.16 lower)	□□□□ MODERATE	CRITICAL
GAD lesions (number of patients with no lesions) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	565/638 (88.6%)	383/588 (65.1%)	RR 1.36 (1.27 to 1.45)	234 more per 1000 (from 176 more to 293 more)	□□□□ MODERATE	CRITICAL
								65.2%		235 more per 1000 (from 176 more to 293 more)		
New or newly enlarged T2 lesions (number of patients with no lesions) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	320/634 (50.5%)	137/590 (23.2%)	RR 2.16 (1.77 to 2.63)	269 more per 1000 (from 179 more to 378 more)	□□□□ MODERATE	CRITICAL
								23.6%		274 more per 1000 (from 182 more to 385 more)		
Discontinuation due to any reason (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	142/783 (18.1%)	186/773 (24.1%)	RR 0.75 (0.57 to 0.99)	60 fewer per 1000 (from 2 fewer to 103 fewer)	□□□□ LOW	CRITICAL
								24.4%		61 fewer per 1000 (from 2 fewer to 105 fewer)		
Discontinuation due to side effects (follow-up 104 weeks)												
2	randomised trials	serious[1]	serious[3]	no serious indirectness	serious[2]	none	123/783 (15.7%)	86/773 (11.1%)	RR 1.42 (0.92 to 2.17)	47 more per 1000 (from 9 fewer to 130 more)	□□□□ VERY LOW	CRITICAL

								11.1%		47 more per 1000 (from 9 fewer to 130 more)		
GAD lesions (mean number) (follow-up 104 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	638	578	-	MD 0.87 lower (1.1 to 0.64 lower)	□□□□ MODERATE	CRITICAL
New or newly enlarged T2 lesions (mean number) (follow-up 104 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	601	591	-	MD 7.03 lower (8.22 to 5.84 lower)	□□□□ MODERATE	CRITICAL
Change in brain volume (percent change) (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	356	329	-	MD 0.3 higher (0.16 to 0.44 higher)	□□□□ MODERATE	CRITICAL
Risk of cancer (number of participants with any neoplasm) (follow-up 104 weeks)												
2	randomised trials	serious[1]	serious[4]	no serious indirectness	serious[2]	none	17/783 (2.2%)	18/773 (2.3%)	RR 0.84 (0.21 to 3.34)	4 fewer per 1000 (from 18 fewer to 54 more)	□□□□ VERY LOW	CRITICAL
								2.3%		4 fewer per 1000 (from 18 fewer to 54 more)		
Risk of infection (number of participants with any infection) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	642/783 (82%)	612/773 (79.2%)	RR 1.04 (0.99 to 1.09)	32 more per 1000 (from 8 fewer to 71 more)	□□□□ MODERATE	CRITICAL
								78.6%		31 more per 1000 (from 8 fewer to 71 more)		

[1] High risk of attrition bias (differences in loss to follow-up between groups and different reasons for drop out). Unclear allocation concealment (Calabresi 2014b)
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

6. Natalizumab compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Natalizumab**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Relapse free (number of participants) (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	501/627 (79.9%)	189/315 (60%)	RR 1.33 (1.21 to 1.47)	198 more per 1000 (from 126 more to 282 more)	□□□□ HIGH	CRITICAL
								60%		198 more per 1000 (from 126 more to 282 more)		
Relapse free (number of participants) (follow-up 104 weeks)												
1	randomised	no serious risk	no serious	no serious	no serious	none	454/615	146/315	RR 1.59 (1.4 to	273 more per 1000 (from 185	□□□□	CRITICAL

	trials	of bias	inconsistency	indirectness	imprecision		(73.8%)	(46.3%)	1.81)	more to 375 more)	HIGH	
								46.4%		274 more per 1000 (from 186 more to 376 more)		
Cumulative disability progression (number of participants worsened) (follow-up 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	106/627 (16.9%)	91/315 (28.9%)	RR 0.59 (0.46 to 0.75)	118 fewer per 1000 (from 72 fewer to 156 fewer)	□□□□ MODERATE	CRITICAL
								28.9%		118 fewer per 1000 (from 72 fewer to 156 fewer)		
Annualised relapse rate (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	627	315	-	MD 0.51 lower (0.67 to 0.35 lower)	□□□□ HIGH	CRITICAL
Annualised relapse rate (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	627	315	-	MD 0.5 lower (0.63 to 0.37 lower)	□□□□ HIGH	CRITICAL
Discontinuation due to side effects (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	15/627 (2.4%)	6/315 (1.9%)	RR 1.26 (0.49 to 3.21)	5 more per 1000 (from 10 fewer to 42 more)	□□□□ MODERATE	CRITICAL
								1.9%		5 more per 1000 (from 10 fewer to 42 more)		
Discontinuation due to any reason (follow-up 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	52/627 (8.3%)	31/315 (9.8%)	RR 0.84 (0.55 to 1.29)	16 fewer per 1000 (from 44 fewer to 29 more)	□□□□ MODERATE	CRITICAL
								9.8%		16 fewer per 1000 (from 44 fewer to 28 more)		
GAD lesions (mean number) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	627	315	-	SMD 0.56 lower (0.7 to 0.42 lower)	□□□□ HIGH	CRITICAL
GAD lesions (mean number) (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	627	315	-	SMD 0.43 lower (0.57 to 0.3 lower)	□□□□ HIGH	CRITICAL
New or newly enlarged T2 lesions (mean number) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	627	315	-	MD 4.9 lower (5.96 to 3.84 lower)	□□□□ HIGH	CRITICAL
New or newly enlarged T2 lesions (mean number) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	627	315	-	MD 9.1 lower (10.98 to 7.22 lower)	□□□□ HIGH	CRITICAL
Risk of cancer (risk of non-event; number of participants with any neoplasm) (follow-up 104 weeks)												
1	randomised	no serious risk	no serious	no serious	serious[1]	none	5/627	1/315	RR 1 (0.99 to	0 fewer per 1000 (from 0	□□□□	CRITICAL

	trials	of bias	inconsistency	indirectness			(0.8%)	(0.3%)	1)	fewer to 0 more)	MODERATE	
								0.3%		0 fewer per 1000 (from 0 fewer to 0 more)		
Risk of infection (number of participants with any infection) (follow-up 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	527/627 (84.1%)	215/315 (68.3%)	RR 1.23 (1.13 to 1.34)	157 more per 1000 (from 89 more to 232 more)	□□□□ HIGH	CRITICAL
								68.3%		157 more per 1000 (from 89 more to 232 more)		
Mortality (risk of non-event)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	2/627 (0.3%)	0/315 (0%)	RR 1 (0.99 to 1)	-	□□□□ MODERATE	CRITICAL
								0%		-		

1 Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

7. Daclizumab compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Daclizumab	Placebo	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	163/201 (81.1%)	127/196 (64.8%)	RR 1.25 (1.11 to 1.42)	162 more per 1000 (from 71 more to 272 more)	□□□□ MODERATE	CRITICAL
								64.8%		162 more per 1000 (from 71 more to 272 more)		
Disability progression (number of participants worsened) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	11/201 (5.5%)	25/196 (12.8%)	RR 0.43 (0.22 to 0.85)	73 fewer per 1000 (from 19 fewer to 99 fewer)	□□□□ LOW	CRITICAL
								12.8%		73 fewer per 1000 (from 19 fewer to 100 fewer)		
Annualised relapse rate (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	201	196	-	MD 0.25 lower (0.37 to 0.13 lower)	□□□□ HIGH	CRITICAL
Discontinuation due to any reason												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	19/201 (9.5%)	18/196 (9.2%)	RR 1.03 (0.56 to 1.9)	3 more per 1000 (from 40 fewer to 83 more)	□□□□ MODERATE	CRITICAL
								9.2%		3 more per 1000 (from 40 fewer to 83 more)		
Discontinuation due to side effects (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	6/201 (3%)	2/196 (1%)	RR 2.93 (0.6 to 14.32)	20 more per 1000 (from 4 fewer to 136 more)	□□□□ MODERATE	CRITICAL
								1%		19 more per 1000 (from 4 fewer to 133 more)		
Brain atrophy (% change in whole brain volume) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	198	196	-	MD 0.05 lower (0.22 lower to 0.12 higher)	□□□□ LOW	CRITICAL
GAD lesions (mean number) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	199	195	-	MD 1.1 lower (1.45 to 0.75 lower)	□□□□ MODERATE	CRITICAL
New or newly enlarged T2 lesions (mean number) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision[2]	none	199	195	-	MD 5.7 lower (7.38 to 4.02 lower)	□□□□ MODERATE	CRITICAL
Risk of malignancy (risk of non-event; number of participants with any neoplasm) (follow-up 52 weeks)												
1	randomised	no serious risk	no serious	no serious	serious[2]	none	1/208	1/204	RR 1 (0.99 to	0 fewer per 1000 (from 0	□□□□	CRITICAL

	trials	of bias	inconsistency	indirectness			(0.5%)	(0.5%)	1.01)	fewer to 0 more)	MODERATE	
								0.5%		0 fewer per 1000 (from 0 fewer to 0 more)		
Risk of infection (number of participants with any infection) (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	104/208 (50%)	89/204 (43.6%)	RR 1.15 (0.93 to 1.41)	65 more per 1000 (from 31 fewer to 179 more)	□□□□ MODERATE	CRITICAL
								43.6%		65 more per 1000 (from 31 fewer to 179 more)		
Mortality (risk of non-event) (follow-up 52 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	1/201 (0.5%)	0/196 (0%)	RR 1 (0.98 to 1.01)	-	□□□□ MODERATE	CRITICAL
								0%		-		

[1] High risk of reporting bias for secondary outcomes (Quality of life reported but not specified in protocol)
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[3] 95% confidence interval around the pooled estimate of effect includes no effect and appreciable

8. Cladribine compared with placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cladribine	Placebo	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	705/889 (79.3%)	266/438 (60.7%)	RR 1.31 (1.2 to 1.42)	188 more per 1000 (from 121 more to 255 more)	⊕⊕⊕⊕ HIGH	CRITICAL
								60.7%		188 more per 1000 (from 121 more to 255 more)		
Annualised relapse rate (follow-up 96 weeks; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	889	438	-	MD 0.19 lower (0.23 to 0.14 lower)	⊕⊕⊕⊕ HIGH	CRITICAL
Discontinuation due to any reason (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	85/889 (9.6%)	58/438 (13.2%)	RR 0.72 (0.53 to 0.99)	37 fewer per 1000 (from 1 fewer to 62 fewer)	⊕⊕⊕O MODERATE	CRITICAL
								13.2%		37 fewer per 1000 (from 1 fewer to 62 fewer)		
Discontinuation due to side effects (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1,2]	none	14/889 (1.6%)	6/438 (1.4%)	RR 1.13 (0.43 to 2.94)	2 more per 1000 (from 8 fewer to 27 more)	⊕⊕OO LOW	CRITICAL
								1.4%		2 more per 1000 (from 8 fewer to 27 more)		
Risk of any infection (number of participants with any infection) (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	427/884 (48.3%)	186/436 (42.7%)	RR 1.13 (1 to 1.29)	55 more per 1000 (from 0 more to 124 more)	⊕⊕⊕⊕ HIGH	CRITICAL
								42.7%		56 more per 1000 (from 0 more to 124 more)		
Risk of serious infection (number of participants with any infection) (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1,2]	none	23/884 (2.6%)	8/436 (1.8%)	RR 1.41 (0.64 to 3.13)	8 more per 1000 (from 7 fewer to 39 more)	⊕⊕⊕O MODERATE	CRITICAL
								1.8%		7 more per 1000 (from 6 fewer to 38 more)		
Risk of cancer (number of participants with any neoplasm) (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	10/884 (1.1%)	0/436 (0%)	RR 5.37 (0.69 to 41.55)	-	⊕⊕⊕O MODERATE	CRITICAL
								0%		-		
Mortality(number of participants with any infection) (follow-up 96 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1,2]	none	4/889 (0.4%)	2/438 (0.5%)	RR 0.99 (0.18 to 5.36)	0 fewer per 1000 (from 4 fewer to 20 more)	⊕⊕OO LOW	CRITICAL
								0.5%		0 fewer per 1000 (from 4 fewer		

										to 22 more)		

[1] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[2] Confidence intervals include a negligible effect and appreciable benefit

9. Interferon compared with glatiramer acetate

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Interferon**	**Glatiramer acetate**	**Relative (95% CI)**	**Absolute**		
Relapse free (number of participants) (follow-up 96-104 weeks)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	773/1310 (59%)	526/865 (60.8%)	RR 0.98 (0.9 to 1.06)	12 fewer per 1000 (from 61 fewer to 36 more)	□□□□ MODERATE	CRITICAL
								61.9%		12 fewer per 1000 (from 62 fewer to 37 more)		
Annualised relapse rate (follow-up 96-104 weeks; Better indicated by lower values)												
2	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	93	56	-	MD 0.05 lower (0.21 lower to 0.11 higher)		CRITICAL
Disability progression (number of participants worsened) (follow-up 104 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	no serious imprecision	none	186/888 (20.9%)	90/448 (20.1%)	RR 1.04 (0.83 to 1.31)	8 more per 1000 (from 34 fewer to 62 more)	□□□□ MODERATE	CRITICAL
								20.1%		8 more per 1000 (from 34 fewer to 62 more)		
GAD lesions (number of patients with no lesions) (follow-up 06 weeks)												
1	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	serious[3]	none	186/230 (80.9%)	154/230 (67%)	RR 1.21 (1.08 to 1.35)	141 more per 1000 (from 54 more to 234 more)	□□□□ LOW	CRITICAL
								67%		141 more per 1000 (from 54 more to 235 more)		
New or newly enlarged T2 lesions (number of patients with no lesions)												
1	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	serious[3]	none	93/230 (40.4%)	86/230 (37.4%)	RR 1.08 (0.86 to 1.36)	30 more per 1000 (from 52 fewer to 135 more)	□□□□ LOW	
								37.4%		30 more per 1000 (from 52 fewer to 135 more)		
New T2 white matter lesion (mean number) (follow-up 104 weeks; Better indicated by lower values)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	93	56	-	MD 0.05 higher (0.29 lower to 0.39 higher)	□□□□ LOW	CRITICAL
New GAD lesions (mean number) (follow-up 104 weeks; Better indicated by lower values)												

1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	93	56	-	MD 0.15 lower (0.48 lower to 0.17 higher)	⨁◯◯◯ LOW	CRITICAL
Combined active lesions (number of participants free from) (follow-up 104 weeks)												
1	randomised trials	serious[6]	no serious inconsistency	no serious indirectness	serious[3]	none	10/36 (27.8%)	12/39 (30.8%)	RR 0.9 (0.45 to 1.83)	31 fewer per 1000 (from 169 fewer to 255 more)	⨁◯◯◯ LOW	CRITICAL
								30.8%		31 fewer per 1000 (from 169 fewer to 256 more)		
New lesions (number of participants free from) (follow-up 104 weeks)												
1	randomised trials	serious[6]	no serious inconsistency	no serious indirectness	serious[3]	none	17/36 (47.2%)	18/39 (46.2%)	RR 1.02 (0.63 to 1.66)	9 more per 1000 (from 171 fewer to 305 more)	⨁◯◯◯ LOW	CRITICAL
								46.2%		9 more per 1000 (from 171 fewer to 305 more)		
New cortical lesions (mean number) (follow-up 48 weeks; Better indicated by lower values)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	86	44	-	MD 0.36 lower (1.24 lower to 0.52 higher)	⨁◯◯◯ LOW	CRITICAL
Discontinuation due to any reason (follow-up 208 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	24/110 (21.8%)	12/56 (21.4%)	RR 1.02 (0.55 to 1.88)	4 more per 1000 (from 96 fewer to 189 more)	⨁◯◯◯ LOW	CRITICAL
								21.4%		4 more per 1000 (from 96 fewer to 188 more)		
Discontinuation due to side effects (follow-up 208 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	14/110 (12.7%)	4/56 (7.1%)	RR 1.78 (0.62 to 5.16)	56 more per 1000 (from 27 fewer to 297 more)	⨁◯◯◯ LOW	CRITICAL
								7.1%		55 more per 1000 (from 27 fewer to 295 more)		
Discontinuation due to side effects (follow-up 48-104 weeks)												
4	randomised trials	serious[7]	no serious inconsistency	no serious indirectness	serious[3]	none	51/1420 (3.6%)	33/921 (3.6%)	RR 1.15 (0.75 to 1.77)	5 more per 1000 (from 9 fewer to 28 more)	⨁◯◯◯ LOW	CRITICAL
								5.1%		8 more per 1000 (from 13 fewer to 39 more)		
Discontinuation due to any reason (follow-up 48-104 weeks)												
4	randomised trials	serious[7]	no serious inconsistency	no serious indirectness	serious[3]	none	147/1420 (10.4%)	89/921 (9.7%)	RR 1.3 (0.68 to 2.47)	29 more per 1000 (from 31 fewer to 142 more)	⨁◯◯◯ LOW	CRITICAL
								14.3%		43 more per 1000 (from 46 fewer to 210 more)		
Mortality (risk of non-event) (follow-up 104 weeks)												
1	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	0/888 (0%)	1/448 (0.2%)	RR 1 (1 to 1.01)	0 fewer per 1000 (from 0 more to 0 more)	⨁◯◯◯ LOW	CRITICAL
								0.2%		0 fewer per 1000 (from 0 more to 0 more)		

[1] Unclear allocation concealment (all studies). High risk of performance bias (Mikol 2008). Unclear risk of performance bias (O'Conner 2009). High risk of missing outcome data (O'Conner 2009).
[2] Unclear risk of perfonmance bias. Unclear allocation concealment. High risk of missing outcome data.

[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[4] Unclear allocation concealment. Unclear risk of performance bias. High risk of missing outcome data.
[5] High risk of performance bias.
[6] Unclear allocation concealment
[7] Unclear allocation concealment (all studies). High risk of performance bias (Mikol 2008). Unclear risk of performance bias (O'Conner 2009). High risk of missing outcome data (O'Conner 2009). Unclear detection bias (Calabrese 2012).

10. Teriflunomide compared with interferon

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Teriflunomide	Interferon	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 48 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	63/109 (57.8%)	88/104 (84.6%)	RR 0.68 (0.57 to 0.82)	271 fewer per 1000 (from 152 fewer to 364 fewer)	□□□□ LOW	CRITICAL
								84.6%		271 fewer per 1000 (from 152 fewer to 364 fewer)		
Annualised relapse rate (follow-up 48 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	111	104	-	MD 0.04 higher (0.17 lower to 0.25 higher)	□□□□ LOW	CRITICAL
Discontinuation due to side effects (follow-up 48 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	12/111 (10.8%)	22/104 (21.2%)	RR 0.51 (0.27 to 0.98)	104 fewer per 1000 (from 4 fewer to 154 fewer)	□□□□ LOW	CRITICAL
								21.2%		104 fewer per 1000 (from 4 fewer to 155 fewer)		
Discontinuation due to any reason (follow-up 48 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	22/111 (19.8%)	30/104 (28.8%)	RR 0.69 (0.42 to 1.11)	89 fewer per 1000 (from 167 fewer to 32 more)	□□□□ LOW	CRITICAL
								28.9%		90 fewer per 1000 (from 168 fewer to 32 more)		
Risk of infection (number of participants with any infection) (follow-up 48 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	54/111 (48.6%)	47/104 (45.2%)	RR 1.08 (0.81 to 1.43)	36 more per 1000 (from 86 fewer to 194 more)	□□□□ LOW	CRITICAL
								45.2%		36 more per 1000 (from 86 fewer to 194 more)		

[1] High risk of performance bias (interferon was open-label) and high risk of attrition bias (differential loss to follow-up between groups). Allocation concealment was not reported (unclear selection bias).
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

11. Fingolimod compared with interferon

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fingolimod	Interferon	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	354/429 (82.5%)	298/431 (69.1%)	RR 1.19 (1.11 to 1.29)	131 more per 1000 (from 76 more to 201 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								69.1%		131 more per 1000 (from 76 more to 200 more)		
Disability progression (number of participants worsened) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	25/429 (5.8%)	34/431 (7.9%)	RR 0.74 (0.45 to 1.22)	21 fewer per 1000 (from 43 fewer to 17 more)	⨁⨁◯◯ LOW	CRITICAL
								7.9%		21 fewer per 1000 (from 43 fewer to 17 more)		
Annualised relapse rate (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	429	431	-	MD 0.17 lower (0.26 to 0.08 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
GAD lesions (number of patients with no lesions) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	337/374 (90.1%)	286/354 (80.8%)	RR 1.12 (1.05 to 1.19)	97 more per 1000 (from 40 more to 154 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								80.8%		97 more per 1000 (from 40 more to 154 more)		
New or newly enlarged T2 lesions (number of patients with no lesions) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	204/372 (54.8%)	165/361 (45.7%)	RR 1.2 (1.04 to 1.39)	91 more per 1000 (from 18 more to 178 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								45.7%		91 more per 1000 (from 18 more to 178 more)		
Discontinuation due to side effects (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	45/429 (10.5%)	32/431 (7.4%)	RR 1.41 (0.92 to 2.18)	30 more per 1000 (from 6 fewer to 88 more)	⨁⨁◯◯ LOW	CRITICAL
								7.4%		30 more per 1000 (from 6 fewer to 87 more)		
Discontinuation due to any reason (follow-up 52 weeks)												
1	randomised	serious[1]	no serious	no serious	serious[2]	none	31/429	45/431	RR 0.69 (0.45	32 fewer per 1000 (from 57	⨁⨁◯◯	CRITICAL

	trials		inconsistency	indirectness			(7.2%)	(10.4%)	to 1.07)	fewer to 7 more)	LOW	
								10.4%		32 fewer per 1000 (from 57 fewer to 7 more)		
GAD lesions (mean number) (Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	374	354	-	MD 0.28 lower (0.5 to 0.06 lower)	⨁⨁⨁◯ LOW	CRITICAL
New or newly enlarged T2 lesions (mean number) (follow-up 52 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	372	361	-	MD 0.9 lower (1.62 to 0.18 lower)	⨁⨁⨁◯ LOW	CRITICAL
Risk of not having cancer (number of participants with any neoplasm) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/429 (0%)	0/431 (0%)	RR 1 (1 to 1)	-	⨁⨁⨁◯ LOW	CRITICAL
								0%		-		
Risk of infection (number of participants with any infection) (follow-up 52 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	184/429 (42.9%)	184/431 (42.7%)	RR 1 (0.86 to 1.17)	0 fewer per 1000 (from 60 fewer to 73 more)	⨁⨁⨁◯ LOW	CRITICAL
								42.7%		0 fewer per 1000 (from 60 fewer to 73 more)		

[1] Unclear risk of detection bias (unclear if rater blinded to participant treatment group). High risk of selective outcome reporting (MSCF measure not listed on protocol but reported in paper).
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[3] Confidence intervals include a negligible effect and appreciable benefit

12. Daclizumab compared with interferon

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Daclizumab	Interferon	Relative (95% CI)	Absolute		
Relapse free (number of participants) (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	616/919 (67%)	470/922 (51%)	RR 1.31 (1.22 to 1.42)	158 more per 1000 (from 112 more to 214 more)	□□□□ MODERATE	CRITICAL
								51%		158 more per 1000 (from 112 more to 214 more)		
Disability progression (number of participants worsened) (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	147/919 (16%)	184/922 (20%)	RR 0.8 (0.66 to 0.98)	40 fewer per 1000 (from 4 fewer to 68 fewer)	□□□□ MODERATE	CRITICAL
								20%		40 fewer per 1000 (from 4 fewer to 68 fewer)		
Annualised relapse rate (follow-up 144 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	919	922	-	MD 0.17 lower (0.22 to 0.12 lower)	□□□□ MODERATE	CRITICAL
Discontinuation due to side effects (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	56/919 (6.1%)	47/922 (5.1%)	RR 1.2 (0.82 to 1.74)	10 more per 1000 (from 9 fewer to 38 more)	□□□□ LOW	CRITICAL
								5.1%		10 more per 1000 (from 9 fewer to 38 more)		
Discontinuation due to any reason (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	195/919 (21.2%)	228/922 (24.7%)	RR 0.86 (0.73 to 1.01)	35 fewer per 1000 (from 67 fewer to 2 more)	□□□□ MODERATE	CRITICAL
								24.7%		35 fewer per 1000 (from 67 fewer to 2 more)		
New or newly enlarged T2 lesions (mean number) (Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	864	841	-	MD 5.20 lower (6.3 to 4.1 lower)	□□□□ MODERATE	CRITICAL
Risk of cancer (risk of non-event; number of participants with any neoplasm) (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	7/919 (0.8%)	8/922 (0.9%)	RR 0.88 (0.32 to 2.41)	1 fewer per 1000 (from 6 fewer to 12 more)	□□□□ LOW	CRITICAL
								0.9%		1 fewer per 1000 (from 6 fewer to 13 more)		
Risk of infection (number of participants with any infection) (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	595/919 (64.7%)	523/922 (56.7%)	RR 1.14 (1.06 to 1.23)	79 more per 1000 (from 34 more to 130 more)	□□□□ MODERATE	CRITICAL

								56.7%		79 more per 1000 (from 34 more to 130 more)		
Mortality (risk of non-event) (follow-up 144 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	1/919 (0.1%)	4/922 (0.4%)	RR 1 (1 to 1.01)	0 fewer per 1000 (from 0 more to 0 more)	□□□□ LOW	CRITICAL
								0.4%		0 fewer per 1000 (from 0 more to 0 more)		

[1] High risk of attrition bias (30% loss to follow-up). Unclear detection bias
[2] 95% confidence interval around the pooled estimate of effect includes no effect and appreciable benefit
[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

13. Alemtuzumab compared with interferon

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Alemtuzumab	Interferon	Relative (95% CI)	Absolute		
Relapse free (number of participants relapse free) (follow-up 104-156 weeks)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	657/914 (71.9%)	261/500 (52.2%)	RR 1.38 (1.26 to 1.51)	198 more per 1000 (from 136 more to 266 more)	□□□□ MODERATE	CRITICAL
								51.4%		195 more per 1000 (from 134 more to 262 more)		
Relapse free (number of participants relapse free) (follow-up 260 weeks)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	76/112 (67.9%)	45/111 (40.5%)	RR 1.67 (1.29 to 2.17)	272 more per 1000 (from 118 more to 474 more)	□□□□ LOW	CRITICAL
								40.5%		271 more per 1000 (from 117 more to 474 more)		
Annualised relapse rate (follow-up 104-156 weeks; Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	538	313	-	MD 0.25 lower (0.33 to 0.18 lower)	□□□□ MODERATE	CRITICAL
Annualised relapse rate (follow-up 260 weeks; Better indicated by lower values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	112	111	-	MD 0.23 lower (0.3 to 0.16 lower)	□□□□ LOW	CRITICAL
Disability progression (number of participants worsened) (follow-up 104-156 weeks)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	92/914 (10.1%)	84/500 (16.8%)	RR 0.59 (0.4 to 0.86)	69 fewer per 1000 (from 24 fewer to 101 fewer)	□□□□ LOW	CRITICAL
								19.8%		81 fewer per 1000 (from 28 fewer to 119 fewer)		
Disability progression (number of participants worsened) (follow-up 260 weeks)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	13/112 (11.6%)	30/111 (27%)	RR 0.43 (0.24 to 0.78)	154 fewer per 1000 (from 59 fewer to 205 fewer)	□□□□ LOW	CRITICAL
								27%		154 fewer per 1000 (from 59 fewer to 205 fewer)		
T2 Lesions (number of participants) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	362/779 (46.5%)	226/374 (60.4%)	RR 0.77 (0.6 to 1)	139 fewer per 1000 (from 242 fewer to 0 more)	□□□□ MODERATE	CRITICAL
								60.4%		139 fewer per 1000 (from 242 fewer to 0 more)		
Discontinuation due to side effects (follow-up 104-156 weeks)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	21/919 (2.3%)	39/496 (7.9%)	RR 0.31 (0.17 to 0.55)	54 fewer per 1000 (from 35 fewer to 65 fewer)	□□□□ LOW	CRITICAL
								7.4%		51 fewer per 1000 (from 33 fewer to 61 fewer)		

Discontinuation due to side effects (follow-up 260 weeks)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	5/108 (4.6%)	14/107 (13.1%)	RR 0.35 (0.13 to 0.95)	85 fewer per 1000 (from 7 fewer to 114 fewer)	⨁◯◯◯ LOW	CRITICAL
								13.1%		85 fewer per 1000 (from 7 fewer to 114 fewer)		
Discontinuation due to any reason (follow-up 104-156 weeks)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	83/935 (8.9%)	149/537 (27.7%)	RR 0.36 (0.25 to 0.52)	178 fewer per 1000 (from 133 fewer to 208 fewer)	⨁◯◯◯ LOW	CRITICAL
								31.6%		202 fewer per 1000 (from 152 fewer to 237 fewer)		
Infection (number of participants with any infection) (follow-up 104-156 weeks)												
3	randomised trials	serious[1]	serious[4]	no serious indirectness	no serious imprecision	none	658/919 (71.6%)	269/496 (54.2%)	RR 1.32 (1.1 to 1.58)	174 more per 1000 (from 54 more to 315 more)	⨁◯◯◯ LOW	CRITICAL
								46.7%		149 more per 1000 (from 47 more to 271 more)		
Infection (number of participants with any infection) (follow-up 260 weeks)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	77/108 (71.3%)	54/107 (50.5%)	RR 1.41 (1.13 to 1.76)	207 more per 1000 (from 66 more to 384 more)	⨁◯◯◯ LOW	CRITICAL
								50.5%		207 more per 1000 (from 66 more to 384 more)		
Mortality (risk of non-event) (follow-up 104-156 weeks)												
3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	4/919 (0.4%)	0/496 (0%)	RR 1 (0.99 to 1)	-	⨁⨁⨁◯ MODERATE	CRITICAL
								0%		-		
Mortality (follow-up 260 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	1/108 (0.9%)	1/107 (0.9%)	RR 1 (0.97 to 1.03)	0 fewer per 1000 (from 0 fewer to 0 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								0.9%		0 fewer per 1000 (from 0 fewer to 0 more)		
Autoimmune disorders (number of participants with any disorder) (104-156 weeks' follow-up) (follow-up 104-156 weeks)												
3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	8/919 (0.9%)	1/496 (0.2%)	RR 2.68 (0.56 to 12.9)	3 more per 1000 (from 1 fewer to 24 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								0%		-		
Autoimmune disorders (number of participants with any disorder) (follow-up 260 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	2/108 (1.9%)	1/107 (0.9%)	RR 1.98 (0.18 to 21.53)	9 more per 1000 (from 8 fewer to 192 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								0.9%		9 more per 1000 (from 7 fewer to 185 more)		
Malignancy (number of participants with any) (follow-up 260 weeks)												
3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	4/919 (0.4%)	3/496 (0.6%)	See comment	0 more per 1000 (from 10 fewer to 10 more)	⨁⨁⨁◯ MODERATE	CRITICAL
								0.9%		0 more per 1000 (from 15		

										fewer to 15 more)		

[1] High risk or performance bias (all studies were open label). High risk of detection bias in Coles 2012 and Cohen 2012 - "In the absence of a masked rater, unmasked raters could submit EDSS assessments"
[2] High risk of performance bias (open-label)
[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[4] Substantial and significant heterogeneity (I2=71%; p=0.03)

14. Ocrelizumab compared with interferon

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Ocrelizumab**	**Interferon**	**Relative (95% CI)**	**Absolute**		
Disability improvement (confirmed at 12 weeks) (follow-up 96 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	130/628 (20.7%)	96/614 (15.6%)	RR 1.32 (1.04 to 1.68)	50 more per 1000 (from 6 more to 106 more)	□□□□ LOW	CRITICAL
								15.6%		50 more per 1000 (from 6 more to 106 more)		
Disability improvement (confirmed at 24 weeks) (follow-up 96 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	98/628 (15.6%)	71/614 (11.6%)	RR 1.35 (1.02 to 1.79)	40 more per 1000 (from 2 more to 91 more)	□□□□ LOW	CRITICAL
								11.6%		41 more per 1000 (from 2 more to 92 more)		
Disability progression (follow-up 96 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	73/724 (10.1%)	109/655 (16.6%)	RR 0.6 (0.46 to 0.8)	67 fewer per 1000 (from 33 fewer to 90 fewer)	□□□□ LOW	CRITICAL
								16.7%		67 fewer per 1000 (from 33 fewer to 90 fewer)		
Infections and infestations (number of participants) (follow-up 096 weeks)												
1	randomised trials	serious[1]	no serious inconsistency[3]	no serious indirectness	no serious imprecision	none	482/825 (58.4%)	433/826 (52.4%)	RR 1.11 (1.02 to 1.22)	58 more per 1000 (from 10 more to 115 more)	□□□□ MODERATE	CRITICAL
								52.4%		58 more per 1000 (from 10 more to 115 more)		
One or more serious adverse event (number of participants) (follow-up 96 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	57/825 (6.9%)	72/826 (8.7%)	RR 0.79 (0.57 to 1.11)	18 fewer per 1000 (from 37 fewer to 10 more)	□□□□ LOW	CRITICAL
								8.7%		18 fewer per 1000 (from 37 fewer to 10 more)		
Influenza-like illness (follow-up 96 weeks)												
1	randomised trials	serious[1]	no serious inconsistency[3]	no serious indirectness	no serious imprecision	none	38/825 (4.6%)	177/826 (21.4%)	RR 0.21 (0.15 to 0.3)	169 fewer per 1000 (from 150 fewer to 182 fewer)	□□□□ MODERATE	CRITICAL

								21.4%		169 fewer per 1000 (from 150 fewer to 182 fewer)		
Mortality (risk of non-event) (follow-up 96 weeks)												
1	randomised trials	serious[1]	no serious inconsistency[3]	no serious indirectness	no serious imprecision	none	1/825 (0.1%)	2/826 (0.2%)	RR 1 (1 to 1.01)	0 fewer per 1000 (from 0 more to 0 more)	□□□□ MODERATE	CRITICAL
								0.2%		0 fewer per 1000 (from 0 more to 0 more)		
Malignancies (risk of non-event) (follow-up 96 weeks)												
1	randomised trials	serious[1]	no serious inconsistency[3]	no serious indirectness	no serious imprecision	none	4/825 (0.5%)	2/826 (0.2%)	RR 1 (0.99 to 1)	0 fewer per 1000 (from 0 fewer to 0 more)	□□□□ MODERATE	CRITICAL
								0.2%		0 fewer per 1000 (from 0 fewer to 0 more)		
Discontinuation due to adverse events (follow-up 96 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	29/827 (3.5%)	64/829 (7.7%)	RR 0.46 (0.3 to 0.7)	42 fewer per 1000 (from 23 fewer to 54 fewer)	□□□□ LOW	CRITICAL
								7.7%		42 fewer per 1000 (from 23 fewer to 54 fewer)		
Discontinuation due to any reason (follow-up 96 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	99/827 (12%)	166/829 (20%)	RR 0.6 (0.48 to 0.75)	80 fewer per 1000 (from 50 fewer to 104 fewer)	□□□□ LOW	CRITICAL
								20%		80 fewer per 1000 (from 50 fewer to 104 fewer)		

[1] Unclear risk of selection bias, attrition bias, detection bias and selective outcome reporting (full report not available).
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[3] Unclear risk - studies combined

15. Interferon compared with placebo for secondary progressive multiple sclerosis

Quality assessment						No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Interferon**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Disability progression sustained at 3 months (follow-up 156 weeks)											
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	140/360 (38.9%)	178/358 (49.7%)	RR 0.78 (0.66 to 0.92)	109 fewer per 1000 (from 40 fewer to 169 fewer)	□□□□ MODERATE	CRITICAL
							49.7%		109 fewer per 1000 (from 40 fewer to 169 fewer)		

Disability progression sustained at 6 months (follow-up 156 weeks)											
3	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	325/863 (37.7%)	347/844 (41.1%)	RR 0.92 (0.8 to 1.06)	33 fewer per 1000 (from 82 fewer to 25 more)	□□□□ MODERATE	CRITICAL
							38.2%		31 fewer per 1000 (from 76 fewer to 23 more)		
Number of participants wheelchair bound (follow-up 156 weeks)											
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	60/360 (16.7%)	88/358 (24.6%)	RR 0.68 (0.51 to 0.91)	79 fewer per 1000 (from 22 fewer to 120 fewer)	□□□□ MODERATE	CRITICAL
							24.6%		79 fewer per 1000 (from 22 fewer to 121 fewer)		
Relapse (number of participants free from) (follow-up 156 weeks)											
2	randomised trials	serious[2]	serious	no serious indirectness	no serious imprecision	340/503 (67.6%)	302/486 (62.1%)	RR 1.08 (0.94 to 1.24)	50 more per 1000 (from 37 fewer to 149 more)	□□□□ LOW	CRITICAL
							62.1%		50 more per 1000 (from 37 fewer to 149 more)		
Discontinuation due to any reason (follow-up 156 weeks)											
4	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	136/1276 (10.7%)	108/1050 (10.3%)	RR 1.05 (0.77 to 1.42)	5 more per 1000 (from 24 fewer to 43 more)	□□□□ LOW	CRITICAL
							9.7%		5 more per 1000 (from 22 fewer to 41 more)		
Discontinuation due to side effects (follow-up 156 weeks)											
3	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	serious[3]	49/599 (8.2%)	12/384 (3.1%)	RR 2.65 (1.42 to 4.95)	52 more per 1000 (from 13 more to 123 more)	□□□□ LOW	CRITICAL
							2.9%		48 more per 1000 (from 12 more to 115 more)		
Discontinuation of study drug due to any reason (follow-up 156 weeks)											
3	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	181/863 (21%)	166/844 (19.7%)	RR 1.07 (0.87 to 1.3)	14 more per 1000 (from 26 fewer to 59 more)	□□□□ MODERATE	CRITICAL
							18.4%		13 more per 1000 (from 24 fewer to 55 more)		
Discontinuation of study drug due to side effects (follow-up 156 weeks)											
2	randomised trials	serious[6]	no serious inconsistency	no serious indirectness	serious[3]	75/677 (11.1%)	27/666 (4.1%)	RR 2.73 (1.78 to 4.19)	70 more per 1000 (from 32 more to 129 more)	□□□□ LOW	CRITICAL
							4%		69 more per 1000 (from 31 more to 128 more)		
Mortality (follow-up 156 weeks)											
4	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	12/1276 (0.9%)	6/1050 (0.6%)	RR 1.5 (0.55 to 4.13)	3 more per 1000 (from 3 fewer to 18 more)	□□□□ LOW	CRITICAL
							1%		5 more per 1000 (from 4 fewer to 31 more)		

Number of participants free from new or newly enlarging T2 lesion (follow-up 156 weeks)											
2	randomised trials	serious[7]	no serious inconsistency	no serious indirectness	serious[3]	156/404 (38.6%)	48/200 (24%)	RR 1.61 (1.22 to 2.12)	146 more per 1000 (from 53 more to 269 more)	□□□□ LOW	CRITICAL
							24%		146 more per 1000 (from 53 more to 269 more)		
Combined unique activity (number of participants free) (follow-up 156 weeks)											
3	randomised trials	serious[7]	no serious inconsistency	no serious indirectness	serious[3]	188/531 (35.4%)	80/439 (18.2%)	RR 1.71 (1.17 to 2.49)	129 more per 1000 (from 31 more to 272 more)	□□□□ LOW	CRITICAL
							25.5%		181 more per 1000 (from 43 more to 380 more)		
Percent change in cerebral volume from baseline (follow-up 52 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	39	38	-	MD 0.2 lower (1.15 lower to 0.75 higher)	□□□□ MODERATE	CRITICAL
Percent change in cerebral volume from baseline (follow-up 104 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	39	33	-	MD 0.59 higher (0.86 lower to 2.04 higher)	□□□□ MODERATE	CRITICAL
Percent change in cerebral volume from baseline (follow-up 156 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	47	43	-	MD 0.5 higher (0.8 lower to 1.8 higher)	□□□□ MODERATE	CRITICAL
Absolute change in brain total lesion volume from baseline (cm3) (follow-up 52 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	329	321	-	MD 2.53 lower (3.22 to 1.84 lower)	□□□□ MODERATE	CRITICAL
Absolute change in brain total lesion volume from baseline (cm3) (follow-up 104 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	308	302	-	MD 3.83 lower (4.92 to 2.74 lower)	□□□□ MODERATE	CRITICAL
Absolute change in brain total lesion volume from baseline (cm3) (follow-up 156 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	334	330	-	MD 4.89 lower (6.11 to 3.67 lower)	□□□□ MODERATE	CRITICAL
Cumulative number of new or enlarging lesions calculated from baseline (follow-up 52 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	350	345	-	MD 2.28 lower (2.93 to 1.63 lower)	□□□□ MODERATE	CRITICAL
Cumulative number of new or enlarging lesions calculated from baseline (follow-up 104 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	350	345	-	MD 4.02 lower (5.09 to 2.95 lower)	□□□□ MODERATE	CRITICAL
Cumulative number of new or enlarging lesions calculated from baseline (follow-up 152 weeks; Better indicated by lower values)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	350	345	-	MD 5.05 lower (6.48 to 3.62 lower)	□□□□ MODERATE	CRITICAL
Number of participants who displayed =>1 active lesion during follow-up (follow-up 156 weeks)											
1	randomised trials	serious[8]	no serious inconsistency	no serious indirectness	no serious imprecision	225/350 (64.3%)	289/345 (83.8%)	RR 0.77 (0.7 to 0.84)	193 fewer per 1000 (from 134 fewer to 251 fewer)	□□□□ MODERATE	CRITICAL
							83.8%		193 fewer per 1000 (from 134 fewer to 251 fewer)		

Quality of life (follow-up 156 weeks; measured with: Multiple Sclerosis Quality of Life Inventory (MSQLI); Better indicated by higher values)											
1	randomised trials	serious[9]	no serious inconsistency	no serious indirectness	no serious imprecision	310	304	-	MD 0.25 higher (0.16 to 0.34 higher)	□□□□ MODERATE	CRITICAL

[1] Confidence intervals include a negligible effect and appreciable benefit
[2] High risk of performance bias for the North American Study Group 2004 ("Patients and treating physicians were more likely to guess treatment allocation correctly due to side effects"). High risk of bias due to incomplete outcome data (The North American Study Group 2004 - 28% of data missing from analysis - and Andersen 2004 - unequal drop-out between groups). Unclear risk of selective outcome reporting - no protocols located. Unclear risk of selection bias as method of sequence generation and allocation concealment not reported (Andersen 2004).
[3] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[4] High risk of performance bias for the North American Study Group 2004 ("Patients and treating physicians were more likely to guess treatment allocation correctly due to side effects"). High risk of bias due to incomplete outcome data (The North American Study Group 2004 - 28% of data missing from analysis - and Andersen 2004 - unequal drop-out between groups). Unclear risk of selective outcome reporting - no protocols located. Unclear risk of selection bias as method of sequence generation and allocation concealment not reported (Andersen 2004). Allocation concealment not reported (SPECTRIMS 2001).
[5] Unclear risk of selective outcome reporting - no protocols located. Unclear risk of selection bias as method of sequence generation and allocation concealment not reported (Andersen 2004). Allocation concealment not reported (SPECTRIMS 2001).
[6] High risk of performance bias for the North American Study Group 2004 ("Patients and treating physicians were more likely to guess treatment allocation correctly due to side effects"). High risk of bias due to incomplete outcome data (The North American Study Group 2004 - 28% of data missing from analysis).
[7] Unclear risk of selective outcome reporting (no protocol located). Allocation concealment unclear.
[8] Unclear risk of selective outcome reporting (unable to locate study protocol)
[9] High risk of performance bias ("Patients and treating physicians were more likely to guess treatment allocation correctly due to side effects"). High risk of bias due to incomplete outcome data. Unclear risk of selective outcome reporting - no protocol located.

16. Mitoxantrone compared with placebo for secondary progressive multiple sclerosis

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mitoxantrone	Placebo	Relative (95% CI)	Absolute		
Disability progression sustained at 3 months (follow-up 104 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	5/60 (8.3%)	14/64 (21.9%)	RR 0.38 (0.15 to 0.99)	136 fewer per 1000 (from 2 fewer to 186 fewer)	⊕⊕OO LOW	CRITICAL
								21.9%		136 fewer per 1000 (from 2 fewer to 186 fewer)		
Participants wheelchair bound (follow-up 104 weeks; assessed with: EDSS)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	3/60 (5%)	7/64 (10.9%)	RR 0.46 (0.12 to 1.69)	59 fewer per 1000 (from 96 fewer to 75 more)	⊕⊕OO LOW	CRITICAL
								10.9%		59 fewer per 1000 (from 96 fewer to 75 more)		
Discontinuation due to any reason (follow-up 104 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	15/63 (23.8%)	18/65 (27.7%)	RR 0.86 (0.48 to 1.55)	39 fewer per 1000 (from 144 fewer to 152 more)	⊕⊕OO LOW	CRITICAL
								27.7%		39 fewer per 1000 (from 144 fewer to 152 more)		
Discontinuation due to side effects (follow-up 104 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	5/63 (7.9%)	2/65 (3.1%)	RR 2.58 (0.52 to 12.81)	49 more per 1000 (from 15 fewer to 363 more)	⊕⊕OO LOW	CRITICAL
								3.1%		49 more per 1000 (from 15 fewer to 366 more)		

[1] Allocation concealment was unclear from the published report. High risk of incomplete outcome data - 27% of the sample were withdrawn from the study prior to trial completion. Unclear risk of selective outcome reporting as no study protocol was available.
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met

Review question 3

1. Interferon vs placebo for primary progressive multiple sclerosis

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Interferon	Placebo	Relative (95% CI)	Absolute		
Disability progression confirmed at 3 months (number of participants) (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	20/51 (39.2%)	24/57 (42.1%)	RR 0.97 (0.62 to 1.52)	13 fewer per 1000 (from 160 fewer to 219 more)	⊕⊕OO LOW	CRITICAL
								42.8%		13 fewer per 1000 (from 163 fewer to 223 more)		
Disability progression confirmed at 6 months (number of participants) (follow-up 104 weeks)												
1	randomised trials	serious[3]	no serious inconsistency	no serious indirectness	serious[2]	none	8/36 (22.2%)	12/37 (32.4%)	RR 0.69 (0.32 to 1.48)	101 fewer per 1000 (from 221 fewer to 156 more)	⊕⊕OO LOW	CRITICAL
								32.4%		100 fewer per 1000 (from 220 fewer to 156 more)		
Discontinuation of study drug due to any reason (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	3/51 (5.9%)	5/57 (8.8%)	RR 1.03 (0.93 to 1.14)	3 more per 1000 (from 6 fewer to 12 more)	⊕⊕OO LOW	IMPORTANT
								9.1%		3 more per 1000 (from 6 fewer to 13 more)		
Discontinuation of study drug due to side effects (follow-up 104 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	1/15 (6.7%)	0/20 (0%)	RR 0.93 (0.78 to 1.1)	-	⊕⊕OO LOW	IMPORTANT
								0%		-		
Discontinuation due to any reason (follow-up 104 weeks)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	1/51 (2%)	2/57 (3.5%)	RR 1.02 (0.95 to 1.09)	1 more per 1000 (from 2 fewer to 3 more)	⊕⊕OO LOW	IMPORTANT
								2.7%		1 more per 1000 (from 1 fewer to 2 more)		
Mortality (follow-up 104 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	0/36 (0%)	1/37 (2.7%)	RR 1.03 (0.95 to 1.11)	1 more per 1000 (from 1 fewer to 3 more)	⊕⊕⊕O MODERATE	CRITICAL
								2.7%		1 more per 1000 (from 1 fewer to 3 more)		

[1] High risk of detection bias (Leary 2003). Unclear allocation concealment and risk of selective outcome reporting (Leary 2003).
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[3] Unclear allocation concealment

2. Glatiramer acetate vs placebo for primary progressive multiple sclerosis

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Glatiramer acetate**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Disability progression (number of participants) (follow-up median 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	248/627 (39.6%)	143/316 (45.3%)	RR 0.87 (0.75 to 1.02)	59 fewer per 1000 (from 113 fewer to 9 more)	□□□□ MODERATE	CRITICAL
								45.3%		59 fewer per 1000 (from 113 fewer to 9 more)		
Time to disability progression												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	248/627 (39.6%)	143/316 (45.3%)	HR 0.87 (0.71 to 1.07)	45 fewer per 1000 (from 105 fewer to 23 more)	□□□□ MODERATE	IMPORTANT
								45.3%		45 fewer per 1000 (from 105 fewer to 23 more)		
Discontinuation of drug due to any reason (156 weeks' follow-up)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	223/627 (35.6%)	116/316 (36.7%)	RR 0.97 (0.81 to 1.16)	11 fewer per 1000 (from 70 fewer to 59 more)	□□□□ MODERATE	IMPORTANT
								36.7%		11 fewer per 1000 (from 70 fewer to 59 more)		
Discontinuation of drug due to side effects (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	48/627 (7.7%)	10/316 (3.2%)	RR 2.42 (1.24 to 4.72)	45 more per 1000 (from 8 more to 118 more)	□□□□ LOW	IMPORTANT
								3.2%		45 more per 1000 (from 8 more to 119 more)		
Mortality (risk of non-event) (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	4/627 (0.6%)	7/316 (2.2%)	RR 1.02 (1 to 1.03)	0 more per 1000 (from 0 more to 1 more)	□□□□ LOW	IMPORTANT
								2.2%		0 more per 1000 (from 0 more to 1 more)		

[1] Unclear risk of selection bias (authors did not describe method for generating the randomisation sequence or allocation of participants to intervention groups). Unclear risk of selective outcome reporting as study protocol was not located.
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

3. Fingolimod vs placebo for primary progressive multiple sclerosis

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Fingolimod**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Disability progression (number of participants) (3 criteria) (follow-up 156 weeks)												

1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	232/336 (69%)	338/487 (69.4%)	RR 0.99 (0.91 to 1.09)	7 fewer per 1000 (from 62 fewer to 62 more)	□□□□ MODERATE	CRITICAL
								69.4%		7 fewer per 1000 (from 62 fewer to 62 more)		
Disability progression (number of participants) (1 criterion) (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	154/336 (45.8%)	240/487 (49.3%)	RR 0.93 (0.8 to 1.08)	34 fewer per 1000 (from 99 fewer to 39 more)	□□□□ MODERATE	CRITICAL
								49.3%		35 fewer per 1000 (from 99 fewer to 39 more)		
Discontinuation of study drug due to side effects (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	52/336 (15.5%)	36/487 (7.4%)	RR 2.09 (1.4 to 3.13)	81 more per 1000 (from 30 more to 157 more)	□□□□ LOW	IMPORTANT
								7.4%		81 more per 1000 (from 30 more to 158 more)		
Mortality (risk of non-event) (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	1/336 (0.3%)	2/487 (0.4%)	RR 1 (0.99 to 1.01)	0 fewer per 1000 (from 0 fewer to 0 more)	□□□□ LOW	IMPORTANT
								0.4%		0 fewer per 1000 (from 0 fewer to 0 more)		
Cancer (number of participants with any neoplasm) (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	26/336 (7.7%)	12/487 (2.5%)	RR 3.14 (1.61 to 6.14)	53 more per 1000 (from 15 more to 127 more)	□□□□ LOW	IMPORTANT
								2.5%		54 more per 1000 (from 15 more to 128 more)		
Infection (number of participants with any infection) (follow-up 156 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	137/336 (40.8%)	215/487 (44.1%)	RR 0.92 (0.78 to 1.09)	35 fewer per 1000 (from 97 fewer to 40 more)	□□□□ MODERATE	IMPORTANT
								44.2%		35 fewer per 1000 (from 97 fewer to 40 more)		
Discontinuation due to any reason (follow-up 156 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	116/336 (34.5%)	170/487 (34.9%)	RR 0.99 (0.82 to 1.2)	3 fewer per 1000 (from 63 fewer to 70 more)	□□□□ MODERATE	IMPORTANT
								34.9%		3 fewer per 1000 (from 63 fewer to 70 more)		

[1] High risk of attrition bias (39% of participants were lost to follow-up)
[2] Optimal information size

4. Ocrelizumab compared with placebo for primary progressive multiple sclerosis

Quality assessment							No of patients		Effect		Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ocrelizumab	Placebo	Relative (95% CI)	Absolute		
Time to disability progression (confirmed at 12 weeks) (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	-	-	HR 0.76 (0.59 to 0.98)	-	⊕⊕⊕O MODERATE	CRITICAL
								0%		-		
Time to disability progression (confirmed at 24 weeks) (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	-	-	HR 0.75 (0.58 to 0.97)	-	⊕⊕⊕O MODERATE	CRITICAL
								0%		-		
Discontinuation of drug due to any reason (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	96/488 (19.7%)	80/244 (32.8%)	RR 0.6 (0.47 to 0.77)	131 fewer per 1000 (from 75 fewer to 174 fewer)	⊕⊕⊕O MODERATE	CRITICAL
								32.8%		131 fewer per 1000 (from 75 fewer to 174 fewer)		
Mortality (risk of non-event) (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	4/486 (0.8%)	1/239 (0.4%)	RR 1 (0.98 to 1.01)	0 fewer per 1000 (from 0 fewer to 0 more)	⊕⊕⊕O MODERATE	CRITICAL
								0.4%		0 fewer per 1000 (from 0 fewer to 0 more)		
Malignancies - number of participants (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	11/486 (2.3%)	2/239 (0.8%)	RR 2.7 (0.6 to 12.11)	14 more per 1000 (from 3 fewer to 93 more)	⊕⊕⊕O MODERATE	CRITICAL
								0.8%		14 more per 1000 (from 3 fewer to 89 more)		
Neoplasms (any) - number of participants (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	8/486 (1.6%)	7/239 (2.9%)	RR 0.56 (0.21 to 1.53)	13 fewer per 1000 (from 23 fewer to 16 more)	⊕⊕⊕O MODERATE	CRITICAL
								2.9%		13 fewer per 1000 (from 23 fewer to 15 more)		
Serious adverse events (at least 1) - number of participants (follow-up 120 weeks)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	99/486 (20.4%)	53/239 (22.2%)	RR 0.92 (0.68 to 1.23)	18 fewer per 1000 (from 71 fewer to 51 more)	⊕⊕⊕O MODERATE	CRITICAL
								22.2%		18 fewer per 1000 (from 71 fewer to 51 more)		

[1] Confidence intervals include a negligible effect and appreciable benefit
[2] Optimal information size (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.