

**Supplementary Table 4.** GRADE certainty assessment summary of findings table

Number of studies	Number of patients	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Other considerations	Anticipated effect (pooled proportion [95% CI])	Certainty of evidence
BMI (follow-up: range 12 months to 24 months)									
10	2,197	Not serious	Serious	Not serious	Not serious	Detected	None	28% reduction [24-32]	⊕○○○ Very low
NAS (follow-up: range 12 months to 24 months)									
5	1906	Not serious	Serious	Not serious	Not serious	Not detected	None	40% reduction [11-60]	⊕○○○ Very low
Steatosis (follow-up: range 12 months to 24 months)									
7	1550	Not serious	Serious	Not serious	Not serious	Not detected	None	37% reduction [19-56]	⊕○○○ Very low
Lobular inflammation (follow-up: range 12 months to 24 months)									
6	270	Not serious	Serious	Not serious	Not serious	Not detected	None	36% reduction [16-56]	⊕○○○ Very low
Ballooning degeneration (follow-up: range 12 months to 24 months)									
5	259	Not serious	Serious	Not serious	Not serious	Not detected	None	38% reduction [24-52]	⊕○○○ Very low
Fibrosis (≥F2) (follow-up: range 12 months to 24 months)									
10	995	Not serious	Serious	Not serious	Not serious	Not detected	None	18% reduction [8-28]	⊕○○○ Very low
MRI-PDFF (follow-up: range 12 months to 24 months)									
6	274	Not serious	Serious	Not serious	Not serious	Not detected	None	72% reduction [67-76]	⊕⊕○○ Low