Table 1: Quality of evidence of study outcomes (GRADE assessment)

Outcomes	Illustrative comparative risks (95% CI)	No. of participants (studies)	Quality of the evidence (GRADE)	Comments
Buzhong Yiqi Decoction (Human studies)				
Muscle mass (ASMI/RASM)	MD 0.99 higher (0.55 to 1.43 higher)	200 (3 RCTs)	⊕⊕○○ LOW[a,b]	Random effects; I^2=80%
Muscle strength (grip strength)	MD 2.14 higher (0.04 to 4.25 higher)	200 (3 RCTs)	⊕⊕○○ LOW[a,b]	Random effects; I^2=50%
Muscle function (gait speed)	MD 0.09 higher (-0.09 to 0.27 higher)	140 (2 RCTs)	⊕○○○ VERY LOW[a,b,c]	Random effects; I^2=97%
Buzhong Yiqi Decoction (Animal studies)				
Gastrocnemius muscle mass (g)	MD 0.42 higher (0.25 to 0.60 higher)	64 (3 studies)	⊕⊕○○ LOW[d,e]	Random effects; I^2=65%
Grip strength (g)	MD 240.01 higher (200.82 to 279.20 higher)	66 (4 studies)	⊕⊕⊕○ MODERATE[d]	Fixed effects; I^2=0%
Exhaustive swimming time (s)	MD 202.86 higher (128.65 to 277.06 higher)	82 (5 studies)	⊕⊕○○ LOW[d,b]	Random effects; I^2=66%
Shenqi Paste (Human studies)				
Muscle mass (SMI)	MD 0.47 higher (0.01 to 0.93 higher)	267 (2 RCTs)	⊕⊕○○ LOW[a,b]	Random effects; I^2=86%
Muscle strength (grip strength)	MD 0.51 higher (-1.12 to 2.15 higher)	267 (2 RCTs)	⊕○○○ VERY LOW[a,b,c]	Not significant
Muscle function (FTSST)	MD 0.88 lower (-1.90 to 0.14 lower)	267 (2 RCTs)	⊕⊕○○ LOW[a,b]	Not significant
Bazhen Decoction (Human studies)				
Muscle mass	MD 0.66 higher (-0.62 to 1.94 higher)	304 (2 RCTs)	⊕○○○ VERY LOW[a,b,c]	Random effects; I^2=99%
Muscle strength (grip strength)	MD 2.63 higher (-2.53 to 7.78 higher)	304 (2 RCTs)	⊕○○○ VERY LOW[a,b,c]	Random effects; I^2=98%
Muscle function (gait speed)	MD 0.15 higher (-0.11 to 0.42 higher)	304 (2 RCTs)	⊕○○○ VERY LOW[a,b,c]	Random effects; I^2=100%
Shenling Baizhu Powder (Human studies)				
Muscle mass (SMI)	MD 0.23 higher (-0.09 to 0.55 higher)	317 (2 RCTs)	⊕⊕○○ LOW[a,b]	Random effects; I^2=91%
Muscle strength (grip strength)	MD 1.65 higher (0.74 to 2.56 higher)	317 (2 RCTs)	⊕⊕⊕○ MODERATE[a]	Fixed effects; I^2=16%
Muscle function (gait speed)	MD 0.01 higher (-0.01 to 0.03 higher)	317 (2 RCTs)	⊕⊕○○ LOW[a,c]	Fixed effects; I^2=48%

GRADE Working Group grades of evidence:
High quality (⊕⊕⊕⊕): Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality (⊕⊕⊕○): Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality (⊕⊕○○): Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality (⊕○○○): We are very uncertain about the estimate.

Downgrading reasons:
[a]Risk of bias: Most studies did not report allocation concealment and blinding methods.
[b]Inconsistency: Substantial heterogeneity ($I^2 > 50\%$).
[c]Imprecision: Wide confidence intervals crossing the line of no effect.
[d]Risk of bias (animal studies): Assessed using SYRCLE tool; concerns about selection and detection bias.
[e]Indirectness: Animal model may not fully translate to human outcomes.

Abbreviations: ASMI, appendicular skeletal muscle index; RASM, relative appendicular skeletal muscle mass; SMI, skeletal muscle mass index; FTSST, five-times sit-to-stand test; MD, mean difference; CI, confidence interval; RCT, randomized controlled trial.