Table 10: Clinical evidence profile: robot-assisted arm training compared to any other intervention

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
Person/participant health related quality of life (SF-36 PCS, 0-100, higher values are better, change score) at end of intervention (follow-up: mean 5 weeks)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	109	106	-	MD **0.73 higher** (0.81 lower to 2.27 higher)	⨁◯◯◯ Very low	CRITICAL
Person/participant health related quality of life (SF-36 MCS, 0-100, higher values are better, change score) at end of intervention (follow-up: mean 5 weeks)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	109	106	-	MD **1.14 lower** (3.5 lower to 1.22 higher)	⨁◯◯◯ Very low	CRITICAL
Person/participant health related quality of life (EQ5D, -0.11-1, higher values are better, final values and change scores) at end of intervention (follow-up: mean 4 weeks)												
2	randomised trials	very serious[c]	not serious	not serious	serious[b]	none	255	461	-	MD **0.01 higher** (0.02 lower to 0.03 higher)	⨁◯◯◯ Very low	CRITICAL
Person/participant health related quality of life (EQ5D, 0-100, higher values are better, change score) at ≥6 months (follow-up: 12 months)												
1	randomised trials	very serious[d]	not serious	not serious	serious[b]	none	97	97	-	MD **4.67 lower** (10.58 lower to 1.24 higher)	⨁◯◯◯ Very low	CRITICAL
Person/participant health related quality of life (EQ5D, -0.11-1, higher values are better, final values) at ≥6 months (follow-up: 6 months)												
1	randomised trials	serious[e]	not serious	not serious	serious[b]	none	221	404	-	MD **0.04 lower** (0.09 lower to 0.01 higher)	⨁⨁◯◯ Low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
Activities of daily living (Barthel index, functional independence measure, stroke impact scale, MAL, Frenchay arm test, ABILHAND [different scale ranges], higher values are better, change scores) at end of intervention (follow-up: mean 5 weeks)												
25	randomised trials	very serious[a]	very serious[f]	not serious	serious[b]	none	678	640	-	SMD **0.41 SD higher** (0.16 higher to 0.67 higher)	⨁◯◯◯ Very low	CRITICAL
Activities of daily living (Barthel index, functional independence measure, Motor activity log [different scale ranges], higher values are better, final values) at end of intervention (follow-up: mean 5 weeks)												
11	randomised trials	not serious	not serious	not serious	not serious	none	389	599	-	SMD **0.14 SD higher** (0.01 higher to 0.27 higher)	⨁⨁⨁⨁ High	CRITICAL
Activities of daily living (Barthel index, functional independence measure, motor activity log [different scale ranges], higher values are better, change scores) at ≥6 months (follow-up: mean 6 months)												
9	randomised trials	not serious	serious[f]	not serious	serious[b]	none	247	222	-	SMD **0.28 SD higher** (0.09 higher to 0.46 higher)	⨁⨁◯◯ Low	CRITICAL
Activities of daily living (Barthel index, Functional Independence Measure [different scale ranges], higher values are better, final values) at ≥6 months (follow-up: mean 4 months)												
2	randomised trials	not serious	very serious[f]	not serious	not serious	none	244	426	-	SMD **0.02 SD higher** (0.14 lower to 0.17 higher)	⨁⨁◯◯ Low	CRITICAL
Arm function (FMA UE, Quick DASH, manual function test [different scale ranges], higher values are better, change scores) at end of intervention (follow-up: mean 5 weeks)												
48	randomised trials	serious[g]	serious[f]	not serious	not serious	none	1125	1042	-	SMD **0.34 SD higher** (0.26 higher to 0.43 higher)	⨁⨁◯◯ Low	CRITICAL
Arm function (FMA UE, Chedoke Arm and Hand Activity [different scale ranges], higher values are better, final values) at end of intervention (follow-up: mean 6 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
24	randomised trials	not serious	serious[f]	not serious	serious[b]	none	639	857	-	SMD **0.2 SD higher** (0.09 higher to 0.31 higher)	⨁⨁◯◯ Low	CRITICAL
Arm function (FMA UE, 0-66, higher values are better, change scores) at ≥6 months (follow-up: mean 6 months)												
11	randomised trials	serious[h]	not serious	not serious	not serious	none	288	229	-	MD **1.08 higher** (0.09 higher to 2.07 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Arm function (FMA UE, Korean DASH [different scale ranges], higher values are better, final values) at ≥6 months (follow-up: mean 4 months)												
9	randomised trials	serious[g]	very serious[f]	not serious	serious[b]	none	370	560	-	SMD **0.61 SD higher** (0.18 higher to 1.03 higher)	⨁◯◯◯ Very low	CRITICAL
Arm muscle strength (Motricity index, MRC, manual muscle test, MRC total motor power [different scale ranges], higher values are better, change scores) at end of intervention (follow-up: mean 5 weeks)												
21	randomised trials	very serious[a]	very serious[f]	not serious	serious[b]	none	548	471	-	SMD **0.45 SD higher** (0.17 higher to 0.72 higher)	⨁◯◯◯ Very low	CRITICAL
Arm muscle strength (Motricity index, MRC [different scale ranges], higher values are better, final values) at end of intervention (follow-up: mean 4 weeks)												
3	randomised trials	very serious[a]	serious[f]	not serious	serious[b]	none	57	50	-	SMD **0.89 SD higher** (0.19 higher to 1.6 higher)	⨁◯◯◯ Very low	CRITICAL
Arm muscle strength (grip strength [kg], higher values are better, change scores and final values) at end of intervention (follow-up: mean 5 weeks)												
5	randomised trials	not serious	not serious	not serious	serious[b]	none	63	60	-	MD **0.92 higher** (0.39 lower to 2.22 higher)	⨁⨁⨁◯ Moderate	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
Arm muscle strength (grip strength [Newton meter], higher values are better, change score and final value) at end of intervention (follow-up: mean 6 weeks)												
2	randomised trials	not serious	not serious	not serious	not serious	none	56	58	-	MD **0.64 lower** (4.18 lower to 2.91 higher)	⨁⨁⨁⨁ High	CRITICAL
Arm muscle strength (MRC total, MRC total motor power [different scale ranges], higher values are better, change scores) at ≥6 months (follow-up: mean 5 months)												
4	randomised trials	serious[i]	very serious[f]	serious[i]	not serious	none	95	69	-	SMD **0.48 SD higher** (0.57 lower to 1.53 higher)	⨁◯◯◯ Very low	CRITICAL
Arm muscle strength (MRC total, MI [different scale ranges], higher values are better, final value) at ≥6 months (follow-up: mean 2 months)												
2	randomised trials	very serious[k]	not serious	serious[i]	not serious	none	42	42	-	SMD **1.05 SD higher** (0.59 higher to 1.51 higher)	⨁◯◯◯ Very low	CRITICAL
Arm muscle strength (grip strength [kg], higher values are better, change score and final value) at ≥6 months (follow-up: mean 6 months)												
2	randomised trials	not serious	not serious	not serious	serious[b]	none	35	36	-	MD **1.06 higher** (1.02 lower to 3.14 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Spasticity (MAS, MAS total [different scale ranges], lower values are better, change scores) at end of intervention (follow-up: mean 5 weeks)												
16	randomised trials	serious[i]	serious[f]	not serious	not serious	none	410	351	-	SMD **0.23 SD lower** (0.46 lower to 0.01 lower)	⨁⨁◯◯ Low	CRITICAL
Spasticity (MAS, MAS total [different scale ranges], lower values are better, final values) at end of intervention (follow-up: mean 5 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
10	randomised trials	very serious[k]	not serious	not serious	not serious	none	168	188	-	SMD **0.21 SD lower** (0.42 lower to 0)	⨁⨁◯◯ Low	CRITICAL
Spasticity (MAS, MAS total [different scale ranges], lower values are better, change scores) at ≥6 months (follow-up: mean 5 months)												
7	randomised trials	serious[l]	not serious	serious[l]	not serious	none	137	110	-	SMD **0.09 SD lower** (0.34 lower to 0.17 higher)	⨁⨁◯◯ Low	CRITICAL
Spasticity (MAS, MAS total [different scale ranges], lower values are better, final values) at ≥6 months (follow-up: mean 3 months)												
4	randomised trials	very serious[k]	not serious	serious[l]	serious[b]	none	72	81	-	SMD **0.2 SD lower** (0.52 lower to 0.12 higher)	⨁◯◯◯ Very low	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale total, 0-100, higher values are better, final values and change scores) at end of intervention (follow-up: mean 7 weeks)												
5	randomised trials	serious[l]	not serious	not serious	serious[b]	none	130	154	-	MD **5.31 higher** (2.6 higher to 8.02 higher)	⨁⨁◯◯ Low	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - hand function domain [different scale ranges], higher values are better, change scores) at end of intervention (follow-up: mean 3 weeks)												
5	randomised trials	serious[g]	very serious[f]	not serious	serious[b]	none	218	164	-	SMD **0.8 SD higher** (0.31 lower to 1.91 higher)	⨁◯◯◯ Very low	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (SS-QOL, 49-245, higher values are better, final value) at end of intervention (follow-up: 4 weeks)												
1	randomised trials	very serious[m]	not serious	not serious	very serious[b]	none	17	20	-	MD **2.21 lower** (23.36 lower to 18.94 higher)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - strength domain, 0-100, higher values are better, change score) at end of intervention (follow-up: 10 weeks)												
1	randomised trials	not serious	not serious	not serious	not serious	none	81	36	-	MD **3.45 higher** (2.58 higher to 4.32 higher)	⨁⨁⨁⨁ High	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - memory domain, 0-100, higher values are better, change score) at end of intervention (follow-up: 10 weeks)												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	81	36	-	MD **0.19 higher** (0.52 lower to 0.9 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - emotion domain, 0-100, higher values are better, change score) at end of intervention (follow-up: 10 weeks)												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	81	36	-	MD **1.24 lower** (1.7 lower to 0.78 lower)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - communication domain, 0-100, higher values are better, change score) at end of intervention (follow-up: 10 weeks)												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	81	36	-	MD **0.32 lower** (0.94 lower to 0.3 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - ADL domain, 0-100, higher values are better, change scores and final value) at end of intervention (follow-up: mean 8 weeks)												
3	randomised trials	serious[n]	very serious[f]	not serious	not serious	none	304	438	-	MD **0.12 higher** (4.56 lower to 4.8 higher)	⨁◯◯◯ Very low	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - mobility domain, 0-100, higher values are better, change score and final value) at end of intervention (follow-up: mean 11 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	not serious	not serious	not serious	not serious	none	294	431	-	MD **0.44 higher** (3.91 lower to 4.79 higher)	⨁⨁⨁⨁ High	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - social participation domain, 0-100, higher values are better, change score and final value) at end of intervention (follow-up: mean 11 weeks)												
2	randomised trials	not serious	not serious	not serious	serious[b]	none	291	430	-	MD **2.81 higher** (5.98 lower to 11.6 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - stroke recovery domain, 0-100, higher values are better, change score) at end of intervention (follow-up: 10 weeks)												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	81	36	-	MD **1.11 higher** (0.21 higher to 2.01 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - physical domain, 0-100, higher values are better, change score) at end of intervention (follow-up: 10 weeks)												
1	randomised trials	not serious	not serious	not serious	not serious	none	81	36	-	MD **3.52 higher** (2.99 higher to 4.05 higher)	⨁⨁⨁⨁ High	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - hand function domain, 0-100, higher values are better, final value) at end of intervention (follow-up: mean 12 weeks)												
1	randomised trials	serious[o]	not serious	not serious	not serious	none	213	395	-	MD **2.6 lower** (6.75 lower to 1.55 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale total, 0-100, higher values are better, change score and final value) at ≥6 months (follow-up: mean 5 months)												
2	randomised trials	serious[p]	not serious	serious[i]	not serious	none	56	34	-	MD **4.36 higher** (1.64 lower to 10.36 higher)	⨁⨁◯◯ Low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - hand function domain, 0-100, higher values are better, final values and change scores) at ≥6 months (follow-up: mean 7 months)												
3	randomised trials	serious[n]	not serious	not serious	not serious	none	320	499	-	MD **0.27 lower** (3.98 lower to 3.45 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - ADL domain, higher values are better, change score and final value) at ≥6 months (follow-up: mean 4 months)												
2	randomised trials	serious[n]	not serious	not serious	not serious	none	223	402	-	MD **2.21 lower** (5.71 lower to 1.28 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - mobility domain, higher values are better, final value) at ≥6 months (follow-up: 6 months)												
1	randomised trials	serious[o]	not serious	not serious	not serious	none	213	395	-	MD **1.7 lower** (5.77 lower to 2.37 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Stroke-specific Patient-Reported Outcome Measures (Stroke Impact Scale - social participation domain, higher values are better, final value) at ≥6 months (follow-up: 6 months)												
1	randomised trials	serious[o]	not serious	not serious	not serious	none	210	394	-	MD **3 lower** (7.23 lower to 1.23 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Withdrawal for any reason at end of intervention (follow-up: mean 6 weeks)												
72	randomised trials	not serious	serious[q]	not serious	very serious[r]	none	160/1890 (8.5%)	177/2064 (8.6%)	**RD 0.00** (-0.02 to 0.02)	**0 fewer per 1,000** (from 20 fewer to 20 more)[s]	⨁◯◯◯ Very low	CRITICAL
Withdrawal for any reason at ≥6 months (follow-up: mean 6 months)												
21	randomised trials	not serious	serious[q]	not serious	very serious[r]	none	56/736 (7.6%)	79/936 (8.4%)	**RD -0.02** (-0.04 to 0.01)	**20 fewer per 1,000** (from 40 fewer to 10 more)[s]	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
Adverse events (cardiovascular events) at end of intervention (follow-up: 3 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	5/257 (1.9%)	2/513 (0.4%)	**RR 4.99** (0.97 to 25.55)	**16 more per 1,000** (from 0 fewer to 96 more)	⨁⨁◯◯ Low	CRITICAL
Adverse events (cardiovascular events) at ≥6 months (follow-up: 6 months)												
1	randomised trials	not serious	not serious	not serious	very serious[b]	none	2/257 (0.8%)	2/513 (0.4%)	**RR 2.00** (0.28 to 14.09)	**4 more per 1,000** (from 3 fewer to 51 more)	⨁⨁◯◯ Low	CRITICAL
Adverse events (injuries and pain) at end of intervention (follow-up: mean 7 weeks)												
5	randomised trials	not serious	serious[q]	not serious	very serious[r]	none	69/285 (24.2%)	71/270 (26.3%)	**RD 0.03** (-0.07 to 0.13)	**30 more per 1,000** (from 70 fewer to 130 more)[s]	⨁◯◯◯ Very low	CRITICAL
Adverse events (injuries and pain) at ≥6 months (follow-up: mean 6 months)												
3	randomised trials	not serious	not serious	not serious	not serious	none	0/149 (0.0%)	0/150 (0.0%)	**RD 0.00** (-0.02 to 0.02)	**0 fewer per 1,000** (from 20 fewer to 20 more)[s]	⨁⨁⨁⨁ High	CRITICAL
Adverse events (other reported adverse events) at end of intervention (follow-up: mean 6 weeks)												
19	randomised trials	not serious	serious[q]	not serious	very serious[r]	none	56/745 (7.5%)	86/991 (8.7%)	**RD 0.01** (-0.01 to 0.04)	**10 more per 1,000** (from 10 fewer to 40 more)[s]	⨁◯◯◯ Very low	CRITICAL
Adverse events (other reported adverse events) at ≥6 months (follow-up: mean 5 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	robot-assisted arm training	all other interventions	Relative (95% CI)	Absolute (95% CI)		
10	randomised trials	not serious	serious[q]	not serious	very serious[r]	none	46/514 (8.9%)	86/760 (11.3%)	**RD 0.00** (-0.03 to 0.04)	**0 fewer per 1,000** (from 30 fewer to 40 more)[s]	⨁◯◯◯ Very low	CRITICAL

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Downgraded by 2 increments as the majority of the evidence was of very high risk of bias (due to bias arising from the randomisation process, bias due to deviation from the intended intervention, bias due to missing outcome data and bias in measurement of the outcome)

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 2 increments as the majority of the evidence was of very high risk of bias (due to bias in the randomisation process, bias due to missing outcome data and bias in measurement of the reported result)

d. Downgraded by 2 increments as the majority of the evidence was of very high risk of bias (due to bias in measurement of the outcome and bias in selection of the reported result)

e. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to bias in measurement of the outcome)

f. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

g. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to a mixture of bias arising from the randomisation process, bias due to deviations from the intended intervention, bias due to missing outcome data, bias in measurement of the outcome and bias in selection of the reported result)

h. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to a mixture of bias arising from the randomisation process, bias due to deviations from the intended intervention and bias due to missing outcome data)

i. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to a mixture of bias arising from the randomisation process, bias due to deviations from the intended intervention, bias due to missing outcome data and bias in measurement of the outcome)

j. Downgraded by 1 increments due to outcome indirectness (as the majority of evidence was reported at a follow up of less than 6 months)

k. Downgraded by 2 increments as the majority of the evidence was of very high risk of bias (due to bias arising from the randomisation process, bias due to deviation from the intended intervention, bias due to missing outcome data, bias in measurement of the outcome and bias in selection of the reported result)

l. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to a mixture of bias arising from the randomisation process, bias due to deviations from the intended intervention, bias due to missing outcome data and bias in measurement of the outcome)

m. Downgraded by 2 increments as the majority of the evidence was of very high risk of bias (due to bias due to deviation from the intended intervention and bias due to missing outcome data)

n. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to a mixture of bias arising from the randomisation process and bias in measurement of the outcome)

o. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to bias in measurement of the outcome)

p. Downgraded by 1 increment as the majority of the evidence was of high risk of bias (due to a mixture of bias arising from the randomisation process, bias due to deviations from the intended intervention and bias due to missing outcome data)

q. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

r. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

s. Absolute effect calculated by risk difference due to zero events in at least one arm of one study