

The following is a detailed description of the methods and results of the Rasch analysis.

Methods

We examined data-model fit using information weighted (INFIT) and unweighted (OUTFIT) mean square values (MNSQ). INFIT and OUTFIT MNSQ are chi-square statistics divided by their degrees of freedom and reported as ratios with an expected value of 1 and a range of 0 to infinity. These statistics provide an indication of the amount of useful information provided by an item. Values above 1 indicate more noise than expected by the model and values less than 1 indicate responses that are more deterministic than expected. Although there are no concrete rules about the acceptable thresholds for INFIT MNSQ and OUTFIT MNSQ, values between 0.5 and 1.5 are generally considered acceptable for use.^{13,29} In addition, Rasch reliability, an index of internal consistency similar to a Cronbach's alpha or KR-20,²⁷ Once item selection was complete, a raw HOOS_{global} score was calculated by summing the responses to each of the included items. Raw scores were then converted to a scale score using a logit transformation, with values ranging from 0 (worse possible) to 100 (best possible).^{6,15}

Results

Three additional items were identified from the full version of the HOOS for potential inclusion in the HOOS_{global}. Less than 33% of patients had preoperative responses of "none" to question S1: "Do you feel grinding, hear clicking or any other type of noise from your hip?", question P1: "How often is your hip painful?", and question Q1: "How often are you aware of your hip?" S1 had an INFIT of 2.61 and OUTFIT of 2.91; an extreme violation of the established thresholds of .5 and 1.5. After removing the item, reliability increased from 0.90 to 0.92, which

further suggests that the item was introducing only noise and not information. The raw HOOS_{global} score was then determined by summing the responses to the original six HOOS, JR items and questions P1 and Q1. The scaled HOOS_{global} score was determined using the logit conversion table (Table A1). The formula for converting raw to HOOS_{global} scaled scores was: $(- \text{Logit} + 8.5004) * 5.8621$. The HOOS_{global} questionnaire and scoring instructions can be found in the Supplemental Files available on the journal's website.

The Individual fit statistics for each item on both the HOOS JR and HOOS_{global} are shown in Table A2. There was one item (A5) whose OUTFIT exceed the previously established thresholds of .5 and 1.5; however, its INFIT did not, meaning that response patterns to this item are less predictable for people with calibrations further away from the item calibration. Overall, the data indicate sufficient fit to the model (Table A2). Person summary statistics for each instrument are shown in Table A3. Rasch reliability was 0.88 for the HOOS, JR and 0.92 for the HOOS_{global}.

Appendix Table A1. Raw score to scaled score conversion chart for the HOOS_{global}

Raw	Logit	SE	HOOS _{global}	Raw	Logit	SE	HOOS _{global}
0	-8.5583	1.9431	100.000	17	0.1401	0.5588	49.009
1	-7.0133	1.2371	90.943	18	0.4554	0.5644	47.161
2	-5.7641	1.0218	83.620	19	0.778	0.5719	45.269
3	-4.8853	0.8547	78.468	20	1.1104	0.5816	43.321
4	-4.2593	0.7358	74.799	21	1.4557	0.5942	41.297
5	-3.7727	0.6649	71.946	22	1.8183	0.6109	39.171
6	-3.3605	0.6225	69.530	23	2.2046	0.6333	36.907
7	-2.9904	0.5961	67.360	24	2.6238	0.6625	34.449
8	-2.6459	0.5789	65.341	25	3.0851	0.6961	31.745
9	-2.318	0.5672	63.419	26	3.5923	0.7271	28.772
10	-2.0011	0.5593	61.561	27	4.1387	0.7501	25.569
11	-1.6913	0.5542	59.745	28	4.7181	0.7737	22.172
12	-1.3859	0.5514	57.954	29	5.3466	0.8167	18.488
13	-1.0826	0.5503	56.177	30	6.0799	0.9067	14.189
14	-0.7797	0.5506	54.401	31	7.0879	1.1429	8.280
15	-0.4758	0.5521	52.619	32	8.5004	1.9031	0.000
16	-0.1697	0.5547	50.825				

Appendix Table A2. Item Measures, Standard Errors, and Fit Statistics

Item	HOOS, JR				HOOS _{global}			
	Measure	S.E.	INFIT MNSQ	OUTFIT MNSQ	Measure	SE	INFIT MNSQ	OUTFIT MNSQ
P5	-.3478	.0791	.92	.89	.7025	.0769	.93	.96
P10	-1.2067	.0775	.94	.93	-.0882	.0745	.86	.87
A3	.1986	.0975	.85	.90	1.2179	.0771	.90	1.21
A5	.1528	.0807	1.15	1.19	1.1689	.0782	1.21	1.76
A12	.5698	.0797	1.09	1.11	1.5856	.0774	1.12	1.30
A14	.6332	.0806	1.02	.97	1.6367	.0783	.98	.96
P1	---	---	---	---	-2.7101	.0723	.84	.90
Q1	---	---	---	---	-3.5134	.0717	1.00	1.12

Measure is the mean item parameter calibration

S.E. is the mean standard error

INFIT MNSQ is an information weighted chi square statistic

OUTFIT MNSQ is an unweighted chi square statistic

Appendix Table A3. Person Summary Statistics of the HOOS, JR and HOOS_{global} (n=608)

Instrument	Measure	S.E.	INFIT	OUTFIT	Rel.	Ext. (%)
			MNSQ	MNSQ		
HOOS, JR	-2.1624	.7427	.99	1.00	.88	128 (21.5%)
HOOS _{global}	-1.7384	.6823	.97	1.09	.92	52 (8.9%)

Measure is the mean item parameter calibration

S.E. is the mean standard error

INFIT MNSQ is an information weighted chi square statistic

OUTFIT MNSQ is an unweighted chi square statistic

Rel. is Rasch reliability

Ext. is the number and percentage of extreme scores
