

Evaluating the impact of a standard error reduction (SER) based stopping rule for PROMIS CAT instruments

Kathleen Fear, Ph.D.¹; Chris DaSilva, B.S.¹;
Judith F. Baumhauer, M.D. M.P.H.²; Michael Bass, M.S.³

¹ Health Lab, University of Rochester Medical Center

² Department of Orthopedics, University of Rochester Medical Center

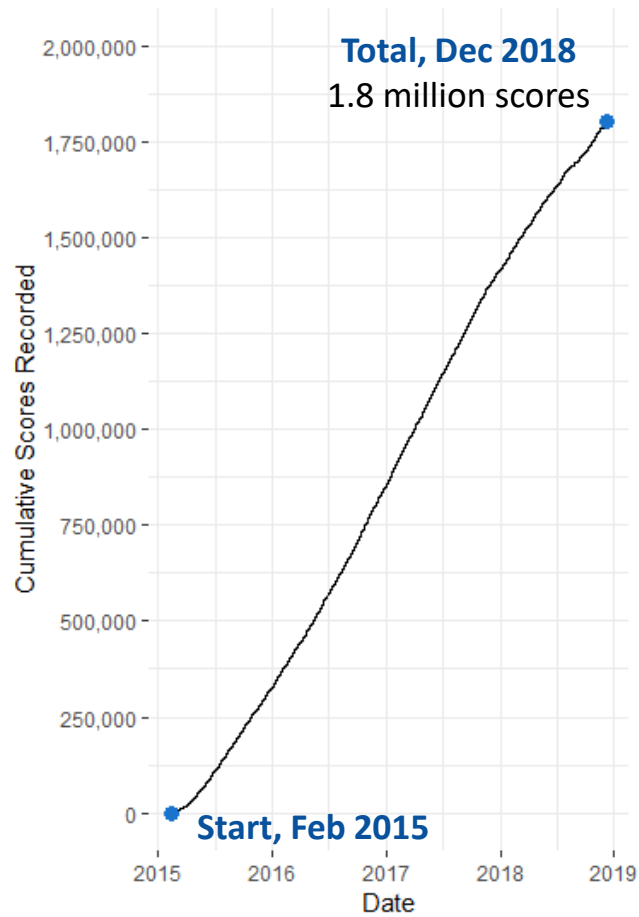
³ Northwestern University

MEDICINE *of*
THE HIGHEST ORDER



UNIVERSITY *of*
ROCHESTER
MEDICAL CENTER

Current State



Adult
Domains
1.7 million
scores
(91%)



Pediatric /
Parent Proxy
Domains
152,579
scores
(9%)

**44
seconds**

Median time to complete assessment

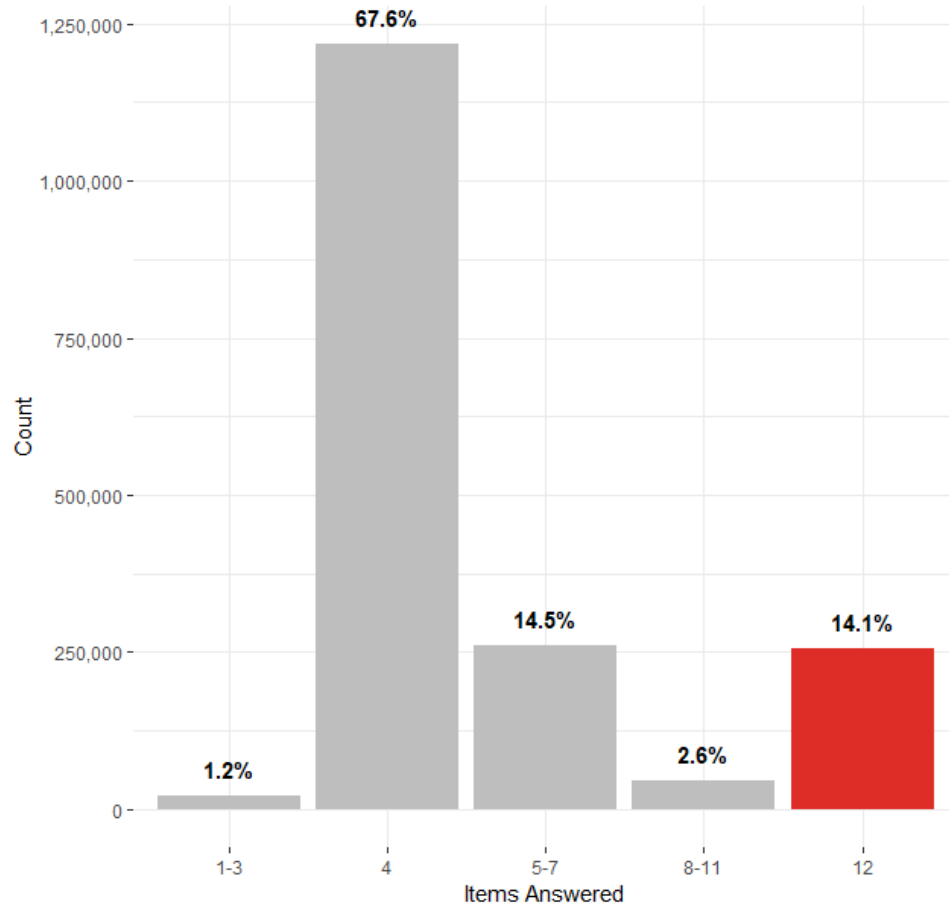
**4
items**

Median items answered per assessment

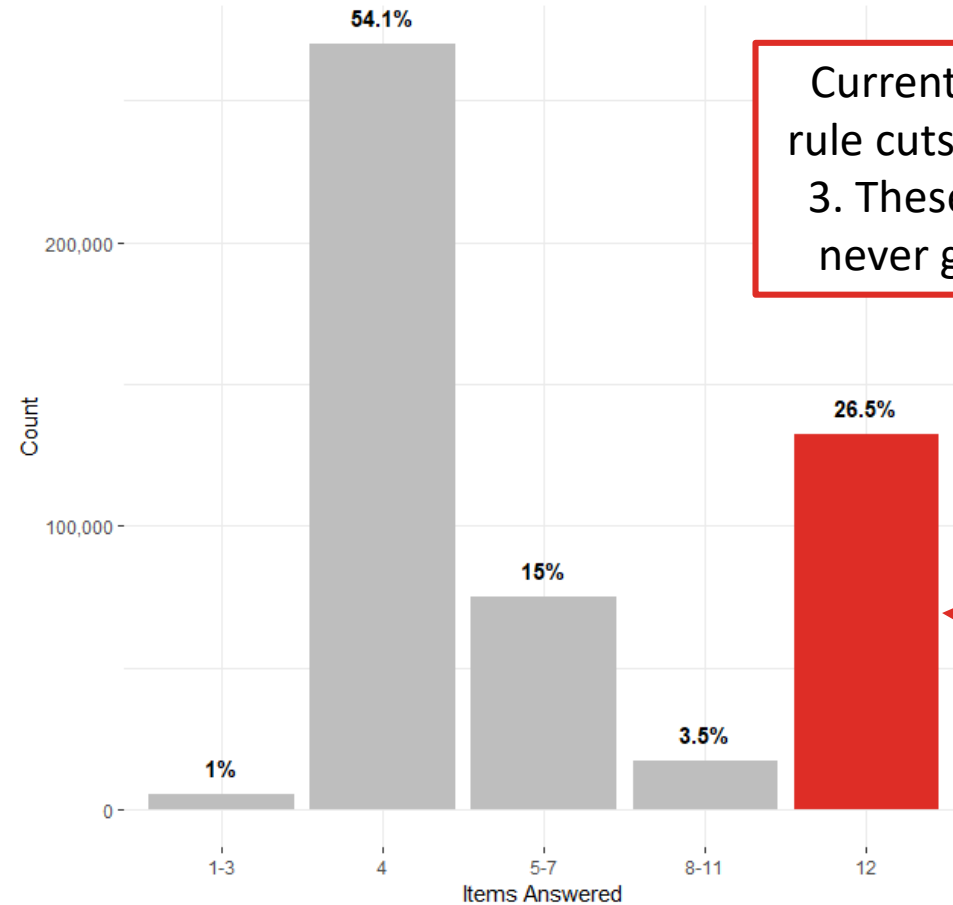
**98.2%
completed**

Some patients experience a higher burden:

All Assessments



Adult Depression

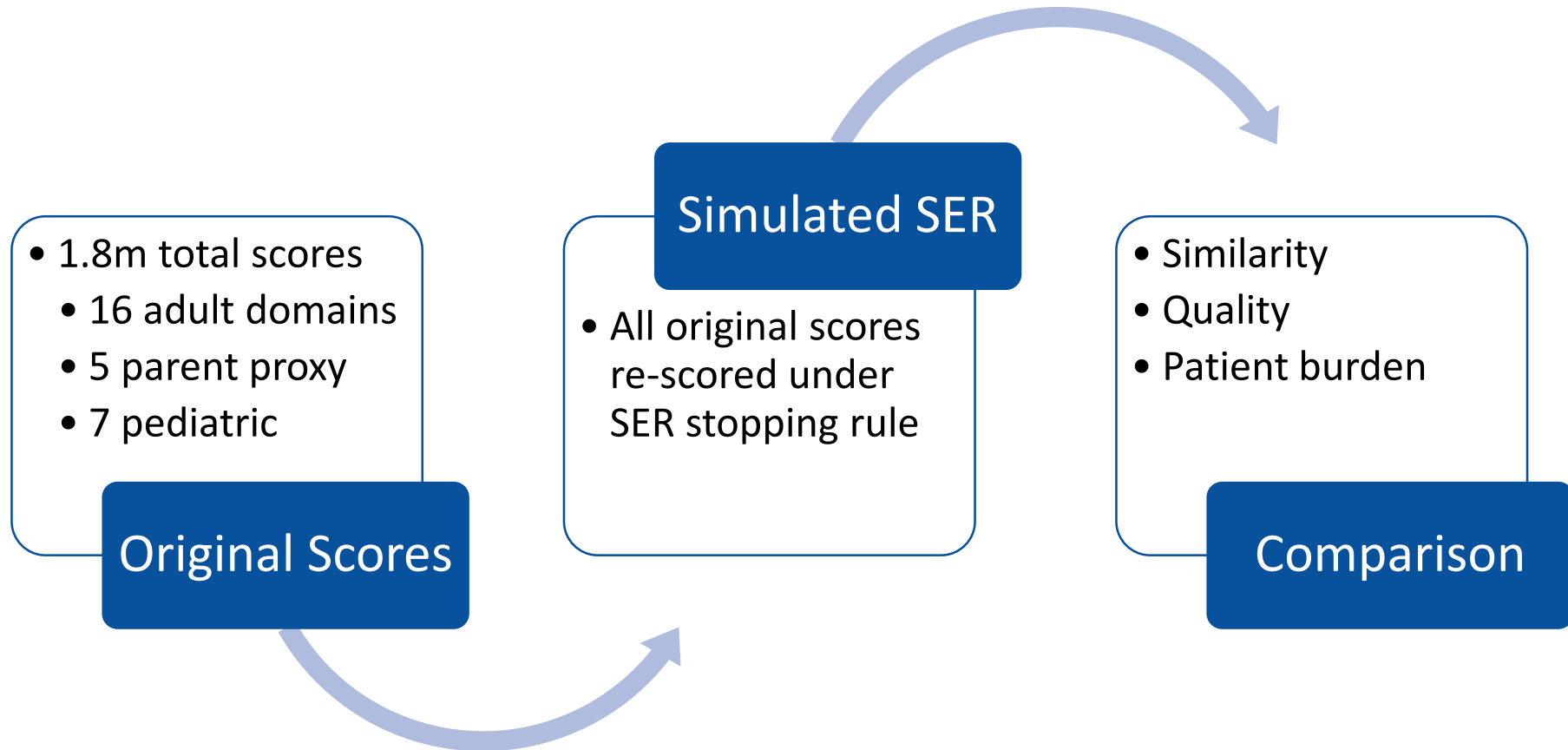


Current stopping rule cuts off at SE < 3. These patients never get there.

New stopping rule: standard error reduction

If administering an additional item doesn't reduce the standard error by more than 0.1, assessment is completed.

Our approach



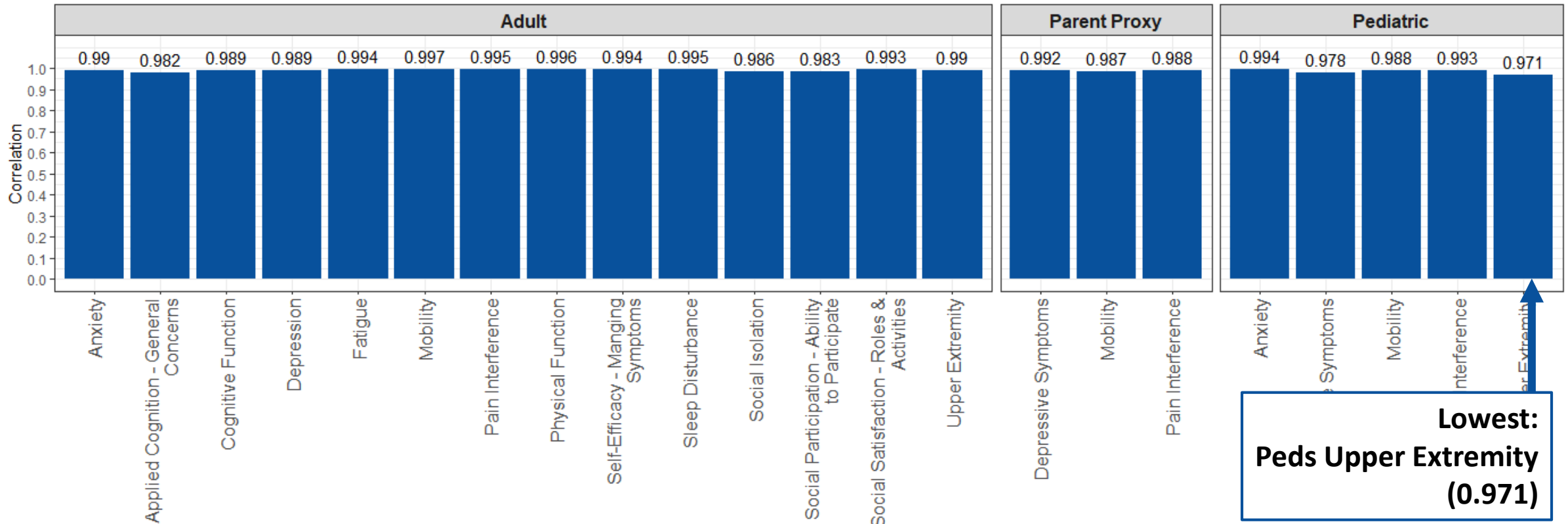
Score Similarity

How well do scores generated using the standard method agree with those generated by SER?

Measures:

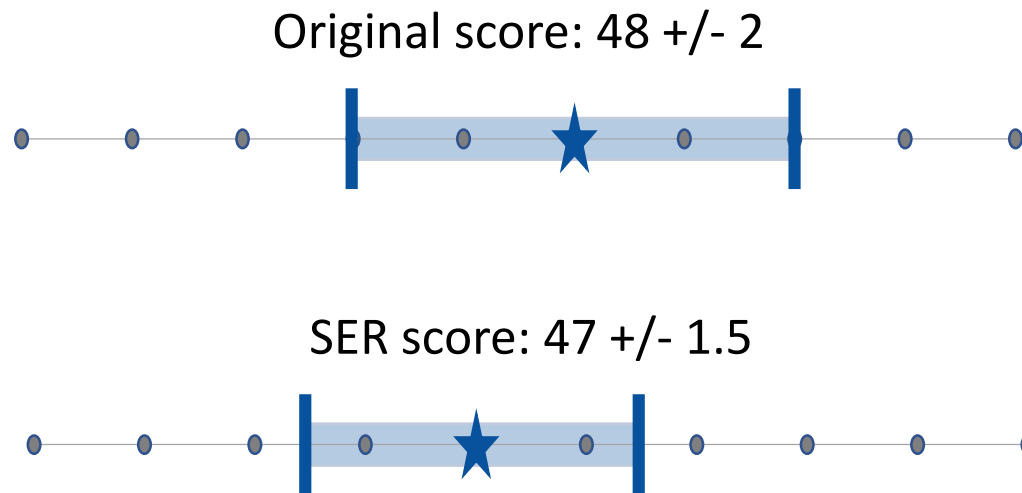
- Correlation between old scores and new
- Overlap between old scores + standard error and new

Score similarity: Correlation between old and new scores



Score similarity:

% overlap between old score +/- SE and new



$$\frac{2 \times |\text{Original Range} \cap \text{New Range}|}{|\text{Original Range}| + |\text{New Range}|}$$

$$\frac{2 \times 2.5}{4 + 3} = .714$$

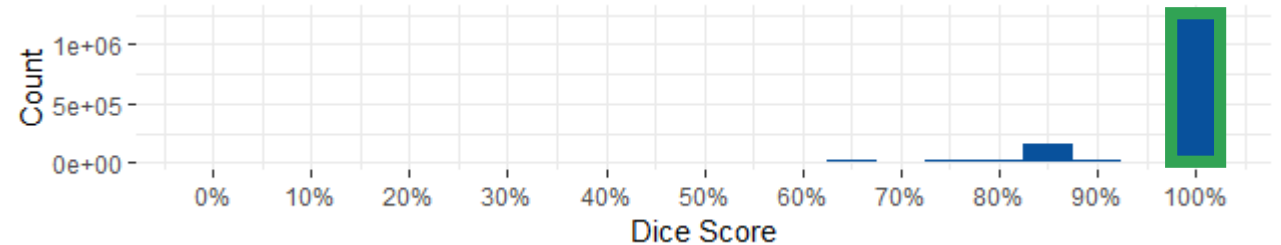
71.4% overlap
between the old score +/- SE and new

Score similarity: % overlap between old score +/- SE and new

Adult

77% of scores overlap by 99.9% or more

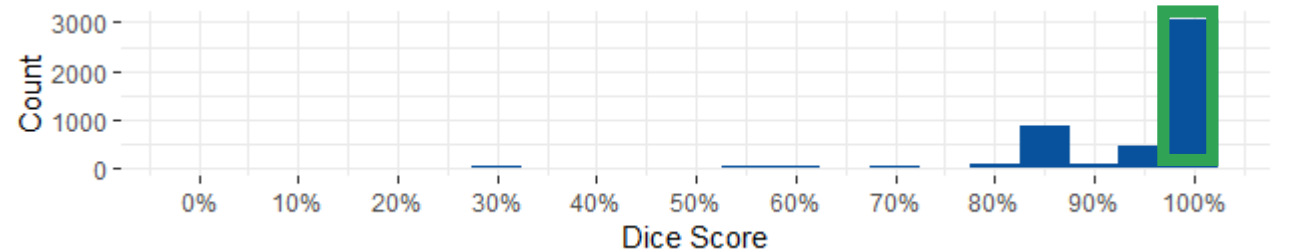
2.3% overlap by less than 50%



Parent Proxy

59.1% of scores overlap by 99.9% or more

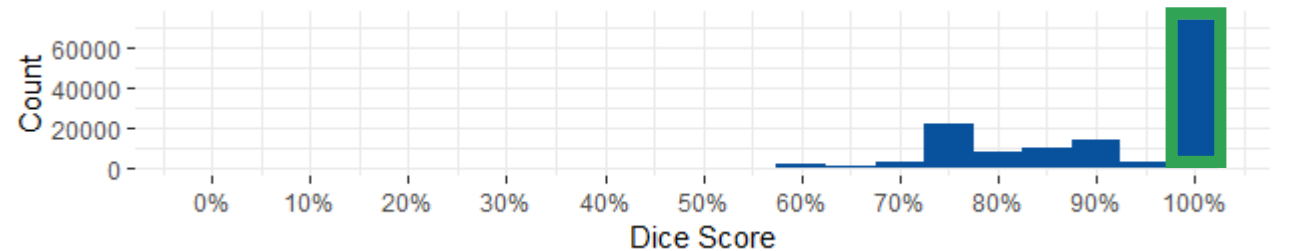
2.7% overlap by less than 50%



Pediatric

50.2% of scores overlap by 99.9% or more

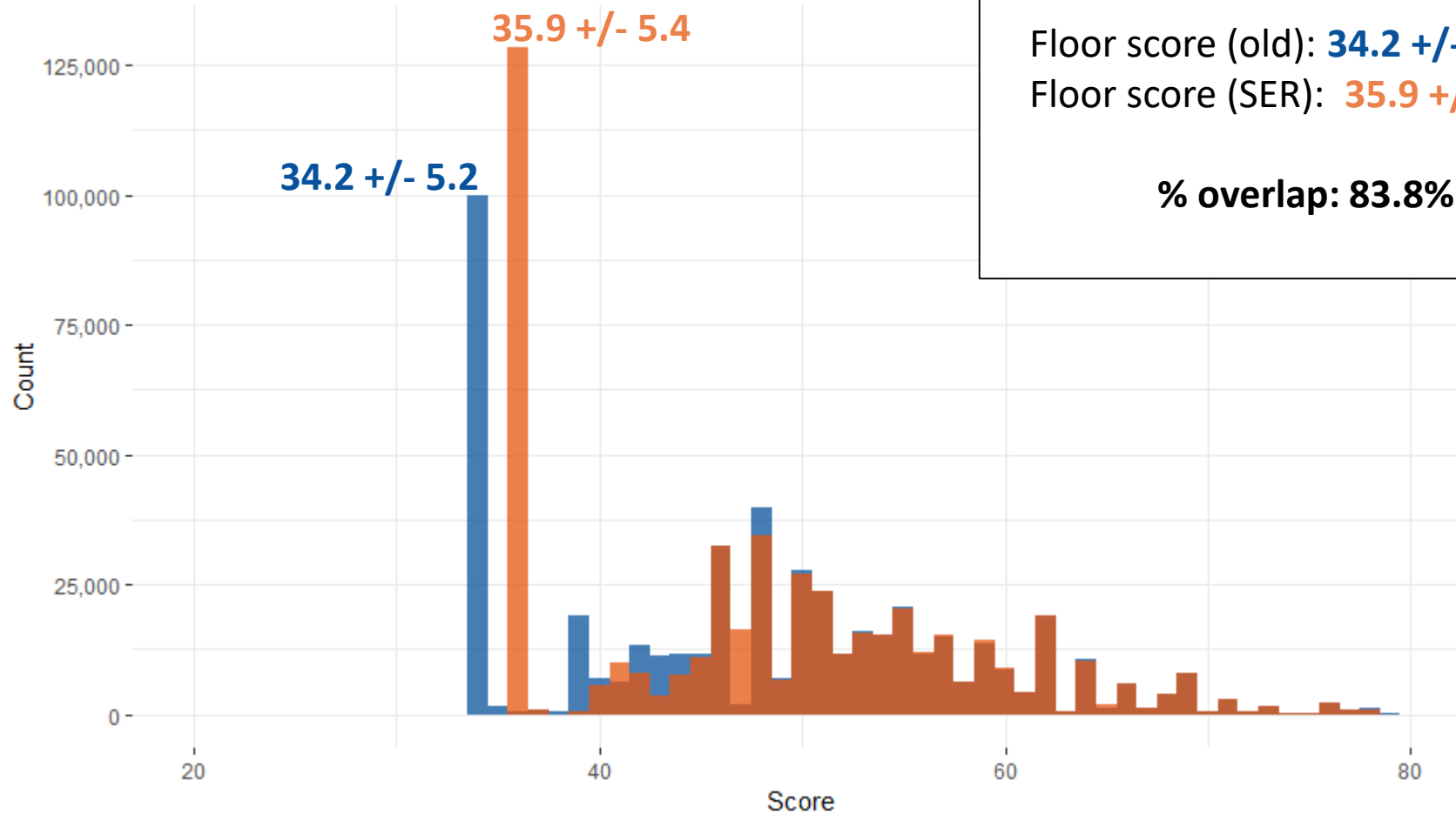
2.0% overlap by less than 50%



Score similarity: % overlap between old score +/- SE and new

		Count	Percent
Adult	Physical Function	458,695	86.2%
	Pain Interference	431,384	82.9%
	Fatigue	4,401	82.6%
	Sleep Disturbance	6,181	81.9%
	Social Satisfaction - Roles & Activities	1,769	79.0%
	Mobility	1,231	71.8%
	Anxiety	21,305	69.3%
	Applied Cognition - General Concerns	759	67.9%
	Upper Extremity	31,145	65.6%
	Depression	309,284	61.7%
	Cognitive Function	407	61.2%
	Social Isolation	388	57.6%
	Social Participation - Ability to Participate	514	49.5%
Self-Efficacy - Managing Symptoms	61	40.4%	
Parent Proxy	Pain Interference	1,120	64.0%
	Mobility	1,013	58.5%
	Depressive Symptoms	936	54.7%
Pediatric	Mobility	29,643	62.1%
	Pain Interference	29,198	59.8%
	Upper Extremity	746	39.4%
	Anxiety	253	30.0%
	Depressive Symptoms	14,065	29.3%

Score similarity: Adult Depression



Score similarity: Scores are less similar in asymptomatic ranges

Scale	Score Range	Total	Corr	99.9% + Overlap
Physical Function	> One SD Above	48,192	0.951	22,869 (47.45%)
	Within One SD	137,798	0.977	118,337 (85.88%)
	> One SD Below	345,547	0.990	316,568 (91.61%)
Pain Interference	> One SD Above	344,999	0.993	307,487 (89.13%)
	Within One SD	132,224	0.933	120,503 (91.14%)
	> One SD Below	43,052	0.357	1,045 (2.43%)
Depression	> One SD Above	124,459	0.994	114,677 (92.14%)
	Within One SD	194,675	0.990	173,121 (88.93%)
	> One SD Below	181,182	0.783	21,054 (11.62%)

Score Quality

How does the quality of PROMIS scores, individually and in aggregate, change when using SER vs the standard method?

Measures:

- Change in completion rate
- Root mean square error

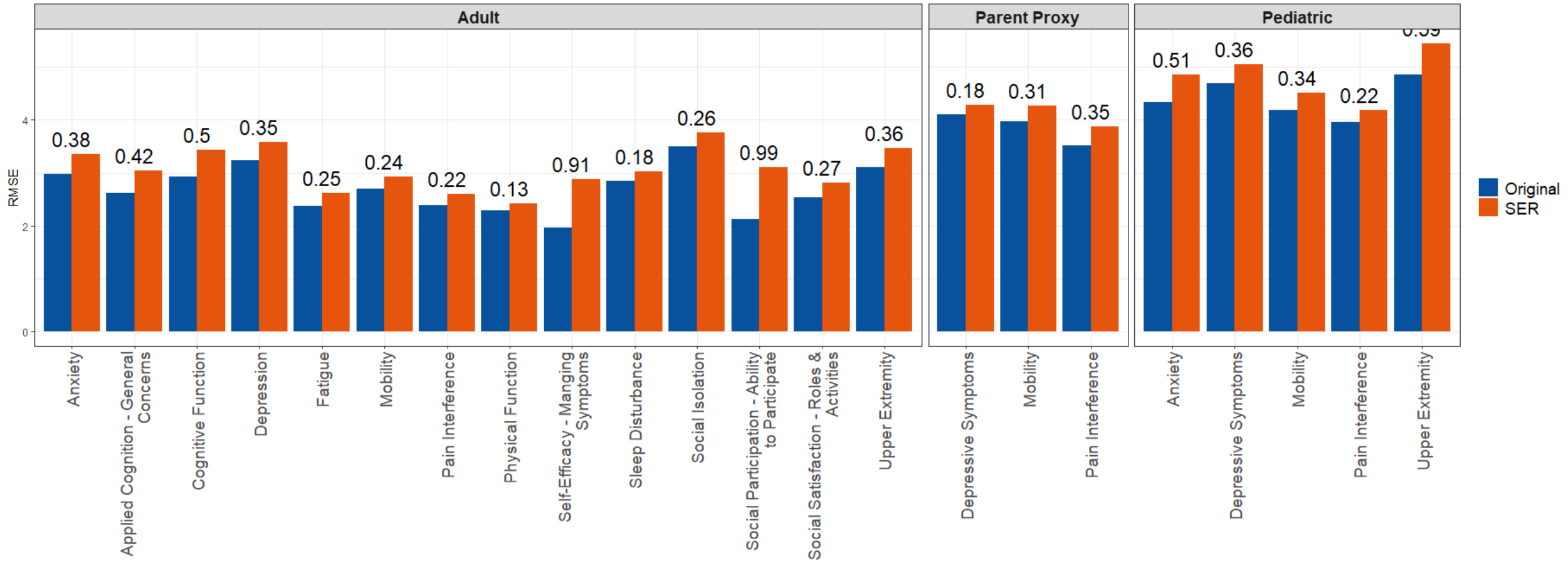
Score Quality

How does the quality of PROMIS scores, individually and in aggregate, change when using SER vs the standard method?

Measures:

- ~~• Change in completion rate~~
- Root mean square error

Score quality: RMSE



Patient Burden

How much does SER reduce the amount of patient effort required?

Measures:

- Change in number of questions administered
- Estimated change in time required to complete assessments

Patient Burden

		Overall		Asymptomatic Patients	
		<i>Difference (95% CI)</i>	<i>Estimated Time Saved (Days)</i>	<i>Difference (95% CI)</i>	<i>Estimated Time Saved (Days)</i>
Adult	Anxiety	1.31 (1.28 - 1.35)	2.9	5 (4.9 - 5)	2.8
	Applied Cognition - General Concerns	2.06 (1.85 - 2.27)	0.2	2.1 (1.8 - 2.3)	0.2
	Cognitive Function	1.41 (1.13 - 1.68)	0.1	5.1 (4.8 - 5.4)	0.1
	Depression	2.08 (2.07 - 2.09)	72.4	5.6 (5.6 - 5.6)	69.5
	Fatigue	0.29 (0.24 - 0.33)	0.2	1.3 (1.1 - 1.5)	0.1
	Mobility	1.06 (0.92 - 1.2)	0.2	6.7 (6.5 - 7)	0.1
	Pain Interference	0.68 (0.67 - 0.68)	31.2	6.7 (6.7 - 6.8)	24.0
	Physical Function	0.19 (0.19 - 0.2)	13	1.4 (1.3 - 1.4)	7.3
	Self-Efficacy - Managing Symptoms	0.42 (0.04 - 0.8)	0	2.4 (0.9 - 3.9)	0.0
	Sleep Disturbance	0.53 (0.48 - 0.59)	0.4	1.6 (1.5 - 1.8)	0.3
	Social Isolation	2.64 (2.36 - 2.93)	0.1	4.2 (3.9 - 4.6)	0.1
	Social Participation - Ability to Participate	0.81 (0.67 - 0.96)	0.1	1.9 (1.6 - 2.2)	0.1
	Social Satisfaction - Roles & Activities	0.84 (0.73 - 0.95)	0.2	2.8 (2.5 - 3.1)	0.1
	Upper Extremity	1.42 (1.39 - 1.45)	6.5	5.8 (5.7 - 5.8)	2.3
Parent Proxy	Depressive Symptoms	2.66 (2.5 - 2.82)	0.4	4.8 (4.6 - 5)	0.4
	Mobility	2.46 (2.31 - 2.61)	0.5	7.8 (7.7 - 8)	0.3
	Pain Interference	2.05 (1.91 - 2.2)	0.3	5.8 (5.5 - 6)	0.3
Pediatric	Anxiety	2.84 (2.63 - 3.04)	0.2	4.3 (4.1 - 4.4)	0.2
	Depressive Symptoms	4.12 (4.09 - 4.15)	13.3	5.8 (5.8 - 5.8)	12.8
	Mobility	2.36 (2.33 - 2.38)	10.2	7.7 (7.7 - 7.7)	7.3
	Pain Interference	1.58 (1.55 - 1.6)	6.3	3.6 (3.5 - 3.6)	5.9
	Upper Extremity	3.72 (3.56 - 3.88)	0.6	7.9 (7.8 - 8)	0.3

Conclusion

- SER scores are similar to original scores across adult domains
- Where scores diverge, it is in the healthy/asymptomatic portion of the population
- Minimal precision is lost using SER
- SER reduces patient burden, especially for healthy patients

MEDICINE *of* THE HIGHEST ORDER



Some patients experience a higher burden:

