

A Comparative Study of Different Methods of Handling Missing Data in Patient Reported Outcomes, Burdens and Experiences (PROBE) Score Algorithm among People with Hemophilia

Quazi Ibrahim¹, Alfonso Iorio^{1,2}, Randall Curtis³, Michael B. Nichol⁴, Declan Noone⁵, Jeffrey S. Stonebraker⁶, Mark W. Skinner^{1,7}, Federico Germini¹, and the PROBE investigators.

¹Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada; ²Department of Medicine, McMaster University, Hamilton, ON, Canada; ³Factor VIII Computing, Berkeley, US; ⁴Sol Price School of Public Policy, University of Southern California, Los Angeles, US; ⁵Irish Haemophilia Society, Dublin, Ireland; ⁶Poole College of Management, North Carolina State University, Raleigh, US; ⁷Institute for Policy Advancement Ltd, Washington, DC

PROBE Score

- PROBE questionnaire measures quality of life in people with hemophilia and healthy controls.
- A score is calculated as the average of nine core question/item scores.
 - 0 being worst reported health status
 - 1 being best reported health status
- There is currently no validated method for calculating the PROBE score when some item scores are missing.

Key Concepts

- Responses to some questions could be more correlated than the others.
- Grouping highly correlated question/item scores into a meaningful domain
- Treating missing values within a domain

Objectives

- Group highly correlated item scores into a domain
- Compare four strategies of estimating PROBE score as an average of the available item scores when
 - ≥ 50% item scores within a domain were available,
 - only one item score within a domain was available,
 - ≥50% item scores were available irrespective of any domain,
 - 8 out of 9 item scores were available.

Methods

- Data:** PROBE phase 3 study
- Domain:** Item scores with intra-class correlation (ICC) ≥ 0.5 grouped

- Simulation:** 36 data set with artificially created missing PROBE and item scores
 - 3 types of missing data (hypothetical)
 - Missing Completely At Random (MCAR)
 - Missing At Random (MAR): missing among elderly (> 45 years)
 - Missing Not At Random (MNAR): missing from lower quartile of the score
 - 3 percentages of missing data: 10%, 15%, and 20%
 - 4 scenarios of missing item scores

Optimal Strategy: estimated vs. true PROBE

- Mean of absolute errors (MAE) (SD) < 0.05
- Test of calibration intercepts and slopes (95% CI)

True PROBE score_i = α₀ + α₁*estimated PROBE score_i + ε_i
 - Intercepts (α₀) not systematically different from 0
 - Slopes (α₁) not systematically different from 1

Results

- Pain or difficulty in daily activities domain
 - Chronic pain, acute pain, pain medication, and difficulty of daily activities
 - ICC ≥ 0.5, n=3217
 - Internally consistent: Cronbach's alpha 0.8
- Strategies 1 and 4 performed better (MAE ± SD: 0.02 ± 0.02 for MCAR and MAR) followed by strategies 2 and 3

Table 1: Closeness of estimated PROBE scores to true PROBE scores in terms of mean of absolute error

Missing Data Type	Strategy 1: 50% or more question scores available within pain/difficulty in daily activities domain	Strategy 2: 1 question score available within pain/difficulty in daily activities domain	Strategy 3: 50% or more of all question scores available	Strategy 4: 90% of all question scores available
MCAR				
10% missing	0.02 ± 0.02, 322	0.04 ± 0.03, 322	0.04 ± 0.04, 322	0.02 ± 0.02, 322
15% missing	0.02 ± 0.02, 483	0.04 ± 0.03, 483	0.04 ± 0.04, 483	0.02 ± 0.02, 483
20% missing	0.02 ± 0.02, 644	0.04 ± 0.03, 644	0.04 ± 0.04, 644	0.02 ± 0.02, 644
MAR				
10% missing	0.02 ± 0.02, 322	0.04 ± 0.03, 322	0.04 ± 0.04, 322	0.02 ± 0.02, 322
15% missing	0.02 ± 0.02, 483	0.04 ± 0.03, 483	0.05 ± 0.04, 483	0.02 ± 0.02, 483
20% missing	0.02 ± 0.02, 644	0.04 ± 0.03, 644	0.04 ± 0.04, 644	0.02 ± 0.02, 644
MNAR				
10% missing	0.03 ± 0.02, 322	0.06 ± 0.04, 322	0.06 ± 0.05, 322	0.04 ± 0.02, 322
15% missing	0.03 ± 0.02, 483	0.06 ± 0.04, 483	0.06 ± 0.05, 483	0.04 ± 0.02, 483
20% missing	0.03 ± 0.02, 644	0.06 ± 0.04, 644	0.06 ± 0.05, 644	0.03 ± 0.02, 644

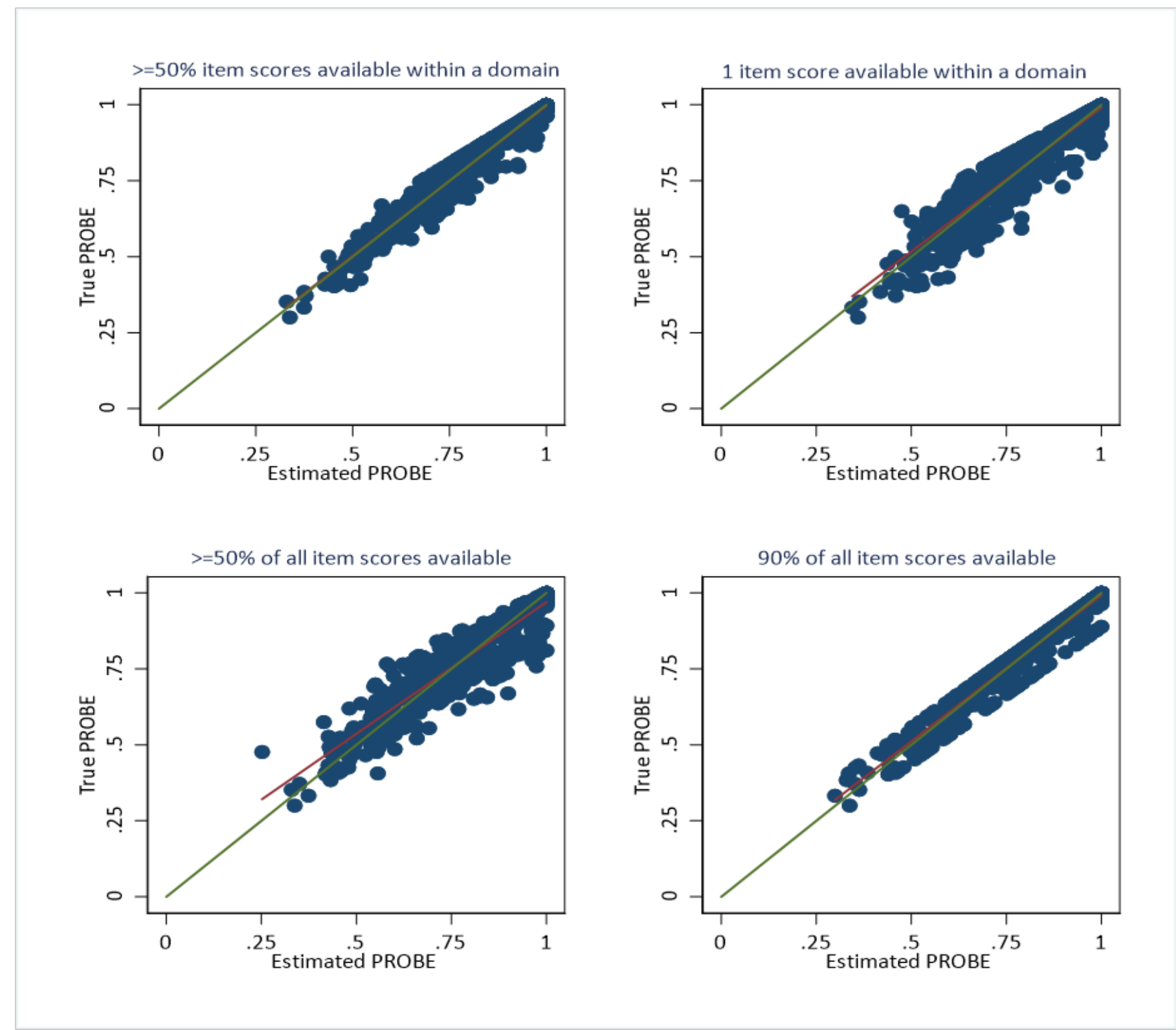
Data are presented as MAE ± SD, N

- For MCAR and MAR
 - Strategy 1 estimated PROBE score accurately
 - Strategies 2 and 4 slightly underestimated the score

Table 2: Calibration intercepts, and slopes for estimated PROBE score in case of missing data

Missing Data Type	Strategy 1: 50% or more question scores available within pain/difficulty in daily activities domain		Strategy 2: 1 question score available within pain/difficulty in daily activities domain		Strategy 3: 50% or more of all question scores available		Strategy 4: 90% of all question scores available	
	Intercept (95% CI)	Slope (95% CI)	Intercept (95% CI)	Slope (95% CI)	Intercept (95% CI)	Slope (95% CI)	Intercept (95% CI)	Slope (95% CI)
MCAR								
10% missing	0.01 (-0.01, 0.02)	0.99 (0.97, 1.01)	0.03 (-0.00, 0.06)	0.96 (0.93, 1.00)	0.11 (0.08, 0.14)	0.85 (0.82, 0.89)	0.04 (0.02, 0.06)	0.95 (0.93, 0.97)
15% missing	0.01 (0.00, 0.03)	0.98 (0.96, 1.00)	0.04 (0.02, 0.07)	0.94 (0.91, 0.97)	0.09 (0.06, 0.11)	0.89 (0.86, 0.92)	0.03 (0.01, 0.04)	0.97 (0.95, 0.98)
20% missing	0.02 (0.00, 0.03)	0.98 (0.96, 0.99)	0.02 (0.00, 0.05)	0.96 (0.94, 0.99)	0.11 (0.09, 0.13)	0.87 (0.84, 0.90)	0.05 (0.03, 0.06)	0.94 (0.93, 0.96)
MAR								
10% missing	0.01 (-0.01, 0.03)	0.99 (0.96, 1.01)	0.06 (0.03, 0.09)	0.93 (0.89, 0.97)	0.13 (0.10, 0.16)	0.84 (0.80, 0.88)	0.05 (0.03, 0.07)	0.94 (0.92, 0.97)
15% missing	0.02 (0.00, 0.03)	0.98 (0.96, 1.00)	0.04 (0.02, 0.07)	0.95 (0.92, 0.98)	0.14 (0.11, 0.16)	0.83 (0.80, 0.86)	0.04 (0.03, 0.06)	0.95 (0.93, 0.97)
20% missing	0.01 (-0.00, 0.02)	0.99 (0.97, 1.00)	0.04 (0.02, 0.06)	0.95 (0.92, 0.98)	0.10 (0.08, 0.13)	0.87 (0.84, 0.90)	0.03 (0.02, 0.04)	0.96 (0.95, 0.98)
MNAR								
10% missing	0.03 (-0.00, 0.06)	0.94 (0.89, 0.99)	0.15 (0.10, 0.20)	0.72 (0.65, 0.80)	0.28 (0.24, 0.31)	0.56 (0.51, 0.61)	0.13 (0.10, 0.15)	0.80 (0.76, 0.84)
15% missing	0.04 (0.01, 0.06)	0.92 (0.89, 0.96)	0.14 (0.10, 0.18)	0.74 (0.68, 0.81)	0.28 (0.25, 0.30)	0.56 (0.51, 0.60)	0.13 (0.11, 0.15)	0.79 (0.76, 0.82)
20% missing	0.04 (0.02, 0.06)	0.93 (0.89, 0.96)	0.14 (0.11, 0.17)	0.74 (0.69, 0.79)	0.26 (0.23, 0.28)	0.58 (0.55, 0.62)	0.10 (0.08, 0.12)	0.84 (0.81, 0.87)

Figure 1: Scatterplots of true PROBE scores against the estimated PROBE scores by the proposed four strategies when 20% data were missing at random (MAR)



The red line is the regression line, the green line represents the perfect fit.

Conclusion

- We recommend estimating PROBE scores as a simple average of the available item scores for MCAR and MAR when
 - at least one item score within pain/difficulty in daily activities domain is available (strategy 1 & 2) OR
 - 8 out of 9 item scores are available irrespective of the domain (strategy 4).