

Appendix 2

Statistical analysis

All data from surveys were entered into Excel spreadsheets (Microsoft Corp., Redmond, WA) and exported to Stata v14.2 (College Station, Texas) for analysis. Responses to each question were summarized by individual question within each domain as well as by overall score for each domain. For individual questions, the percent of favorable respondents to each question was calculated as number of favorable individual responses divided by total number of responses. For overall scores for each domain, we computed scores at the individual question level for each staff member. A “positive” climate score was assigned when $\geq 50\%$ of staff responses in the domain were favorable. We then calculated the overall score as the number of individuals with positive responses divided by the total number of responses. For psychometric testing of the internal reliability, we assessed responses to each question using Cronbach's α , with new variables created to standardize scoring for reverse-worded items.

To facilitate data dissemination, we summarized all results using radar charts. To summarize staff responses to individual questions, we constructed a separate radar chart for each domain, with each spoke representing the percent of favorable responses to individual questions (see Appendix 2). For summary of overall domain scores, we separated health climate domains and staff engagement domains into two radar charts. For each of these charts, each spoke represented the overall score for each domain. Color coding was used to aid visual interpretation of radar charts, with yellow and green color used to demonstrate $\geq 60\%$ favorable responses and red used to demonstrate values below 60% favorable responses. The 60% threshold was used to represent desirable goal of favorable responses based on previous experiences using the SAQ, in which units with $< 60\%$ of favorable responses in each domain were determined to have the most to gain from QI efforts and were substandard in clinical and operational performance.^{18,19,32}

All 13 domains had a Cronbach's α greater than 0.65, suggestive of moderate to high level of internal consistency and scale reliability (Table 1). Local leadership (0.96), job certainty (0.95), and personal burnout (0.91) showed the highest shared covariance amongst all items in the domain.

Table 1. Cronbach's α Coefficient of Reliability for SCORE Survey Domain	
Domain	Cronbach's α
Learning environment	0.74
Local leadership	0.96
Burnout climate	0.84
Personal burnout	0.91
Teamwork	0.75
Safety climate	0.70
Work-life balance	0.80
Growth opportunities	0.88
Workload strain	0.80
Job certainty	0.95
Intentions to leave	0.69
Advancement	0.84
Decision making	0.84

Table 1: Psychometric testing of the SCORE survey at the Pediatric Nephrology unit at the Roosevelt Hospital, Guatemala. All domains had a Cronbach's α greater than 0.65, suggestive of a moderate to high level of internal consistency. Local leadership (0.96), job certainty (0.95), and personal burnout (0.91) showed the highest shared covariance amongst all items in the domain.