Technical Appendix

Receiver operating characteristic (ROC) curve is a common technique used in medical decision-making research to determine how well a potential classifying variable can discriminate between two classes [10]. In the context of this paper, the potential classifying variables are the continuously measured scores on the QLQ-C30, and the two classes are the binary grouping of the SCNS domains/items. As a working example, we will use the physical function domain of the QLQ-C30 as our potential classifier and the "work around the home" item from the SCNS as our binary class. For this specific pairing, there were 6 individuals with a missing result on the SCNS and 1 with a missing result on the QLQ-C30, leaving 110 patients for the analysis.

The ROC curve is a graphical representation of the trade-off between the sensitivity and specificity of a potential classifier. Sensitivity is the probability that an individual will be correctly classified as "positive" given that he or she is truly positive. Specificity is the probability that an individual will be correctly classified as "negative" when he or she is truly negative. Sensitivity is also equal to the true positive rate:

$$\begin{aligned}\text{Sensitivity} &= \text{True Positive Rate}\\ &= \frac{\text{Number of positives correctly classified}}{\text{Number of true positives}}\end{aligned}$$

Specificity is also equal to 1 minus the false positive rate:

$$\begin{aligned}\text{Specificity} &= 1 - \text{False Positive Rate}\\ &= 1\\ &\quad - \frac{\text{Number of negatives incorrectly classified}}{\text{Number of true negatives}}\end{aligned}$$

Two other commonly used quantities in ROC analyses are positive predictive value (PPV) and negative predictive value (NPV). The positive predictive value is the probability of being a true positive when you are classified as positive:

$$\begin{aligned}&\text{Positive Predictive Value}\\ &= \frac{\text{Number of true positives correctly classified}}{\text{Number of classified positives}}\end{aligned}$$

The negative predictive value is the probability of being a true negative when you are classified negative:

$$\begin{aligned}&\text{Negative Predictive Value}\\ &= \frac{\text{Number of true negatives correctly classified}}{\text{Number of classified negatives}}\end{aligned}$$

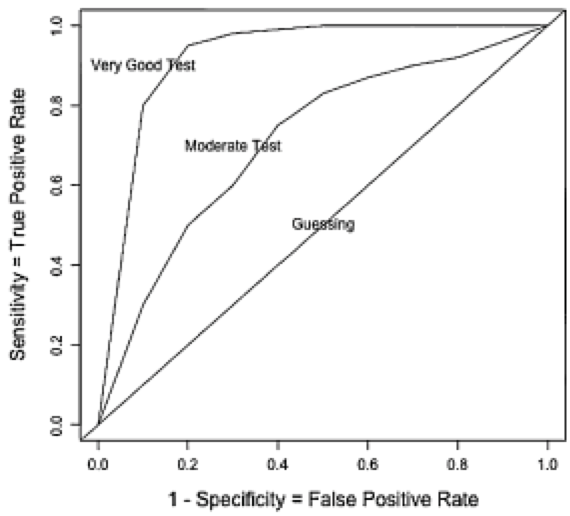

Fig. 3 Examples of three ROC curves representing predictors with chance classification ($x = y$ line), moderately good classification (*middle curve*), and very good classification (*upper curve*)

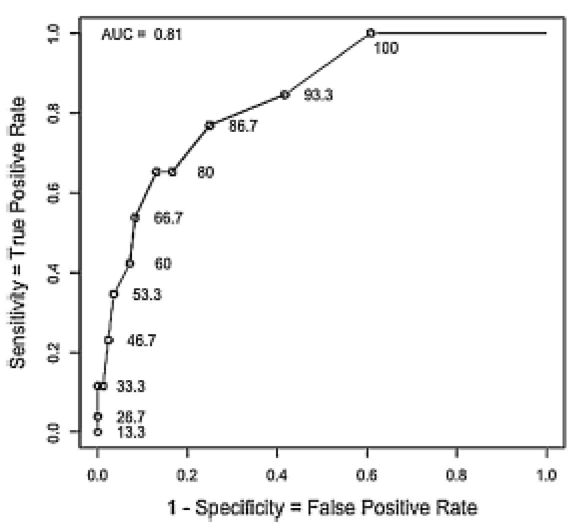

Fig. 4 ROC curve for the discriminative ability of the physical function domain of the QLQ-C30 in predicting patients who report unmet needs performing work around the home. The cut-offs of the QLQ-C30 scores that yield each possible sensitivity and specificity are shown along the curve

The ROC curve shows the trade-off between the true and false positive rates. The example in Fig. 3 illustrates three potential curves. The curve where sensitivity = 1 − specificity for all values represents mere guessing, or doing no better at classifying individuals than just flipping a fair coin. A moderately good classifier is represented by the curve in the middle, where the false positive rate

Table 3 Example of two-by-two tables for classifying patients as reporting unmet needs performing work around the home using cut-off scores of 80 and 90 on the QLQ-C30

	QLQ-C30				Total
	Cut-off = 80		Cut-off = 90		
	Classified negative	Classified positive	Classified negative	Classified positive	
SCNS					
True negative	70	14	49	35	84
True positive	9	17	4	22	26
Total	79	31	53	57	110

increases as the true positive rate increases. An even better classifier is shown in the top curve, where the false positive rate increases at a slower rate as the true positive rate increases. A perfect classifier has a single point at (0,1), where the true positive rate is 100% and the false positive rate is 0%.

The area under the ROC curve (AUC) is a commonly used measure to describe the performance of the potential classifier. The AUC is the probability that a randomly chosen true positive will be ranked higher by the classifier than a randomly chosen true negative. Thus, an AUC of .50 represents a chance classification, and an AUC of 1.0 represents perfect classification. In our example, there are 26 individuals who are true positives (report unmet needs performing work around the home) and 84 true negatives (do not report unmet needs performing work around the home). The ROC curve in Fig. 4 for the physical function domain of the QLQ-C30 has an AUC of .81, which means that of all possible pairings of the 26 positives and 84 negatives, the positive will be more highly ranked than the negative 81% of the time.

In our analysis, once a good classifier was identified, we examined potential cut-offs for the classifier by creating 2×2 tables showing the number of patients who are correctly and incorrectly classified. The physical function classifier takes on values ranging from 0 to 100 for each individual. For the purpose of illustration, in the table below we present two potential cut-offs of this score for classifying an individual as having trouble with work around the home: 80 and 90 (Table 3).

There are 31 individuals classified positive with a physical function score higher than or equal to 80 and 79 classified negative with a score less than 80. Among the 31 called positive, 17 are true positives, which results in a sensitivity of 17/26 = .65. Among the 79 individuals classified negative, 70 are true negatives, which results in a specificity of 70/84 = .83. The PPV and NPV of this cut-off are equal to 17/31 = .55 and 70/79 = .89, respectively.

Using a cut-off of 90 results in a more sensitive (sensitivity = 22/26 = .85), but less specific test (specificity = 49/84 = .58). The PPV and NPV of this cut-off are equal to 22/57 = .39 and 49/53 = .92, respectively.