INTRODUCTION

The purpose of this research was to assess the measurement equivalence of an IVR version of the EQ-5D with the original paper-based version. In addition, this study was conducted to test the measurement equivalence of this IVR version of the EQ-5D with the original paper version of the instrument. Testing the measurement equivalence of INTERACTIVE VOICE RESPONSE (IVR) AND PAPER VERSIONS OF THE EQ-5D

METHODS

This study utilized a crossover design with subjects randomly assigned to one of two assessment orders. 1) paper then IVR or 2) IVR then paper. A convenience sample of non-treatment outpatient cancer clinic patients (n=119) were asked to complete each assessment two days apart. The analyses tested for mean differences (repeated measures ANOVA) and intraclass correlation coefficients (ICC) to assess measurement stability over time. Equivalence of the means was established if the 95% confidence interval (CI) of the mean difference was within the minimally important difference (MID) interval. 0.035 to 0.035 for the index and -0.3 to 3 for the EQ VAS. Adequacy of the ICC was established by comparing the ICC 95% lower CI with a critical value of 0.70.

RESULTS

The per protocol analysis included 109 subjects for the EQ VAS and 113 subjects for the index. For the EQ-SD index, the means (SD) of the paper and IVR administrations were 0.790 (0.172) and 0.800 (0.180), respectively. The 95% CI of the mean difference was -0.024 to 0.006, which was within the equivalence interval. The ICC was 0.894 (95% lower CI 0.857), significantly different from 0.70. For the EQ VAS, the means (SD) were 72.0 (19.7) for paper and 74.1 (19.8) for IVR. The 95% CI of the mean difference was -3.784 to -0.484, partially within the equivalence interval. The ICC was 0.897 (95% lower CI 0.859) also significantly different from 0.70.

DISCUSSION

This analysis provides evidence that the EQ-SD scores on the IVR version were equivalent to those obtained on the original paper version.

OBJECTIVE

Electronic data capture technologies, such as interactive voice response (IVR) systems, are emerging as important alternatives for collecting patient-reported outcomes data. The objective of this study was to assess the measurement equivalence of an IVR version of the EQ-SD with the original paper version.

METHODS

This study utilized a crossover design with subjects randomly assigned to one of two assessment orders: 1) paper then IVR or 2) IVR then paper. A convenience sample of non-treatment outpatient cancer clinic patients (n=119) were asked to complete each assessment two days apart. The analyses tested for mean differences (repeated measures ANOVA) and intraclass correlation coefficients (ICC) to assess measurement stability over time. Equivalence of the means was established if the 95% confidence interval (CI) of the mean difference was within the minimally important difference (MID) interval: 0.035 to 0.035 for the index and -3 to 3 for the EQ VAS. Adequacy of the ICC was established by comparing the ICC 95% lower CI with a critical value of 0.70.

RESULTS

The per protocol analysis included 109 subjects for the EQ VAS and 113 subjects for the index. For the EQ-SD index, the means (SD) of the paper and IVR administrations were 0.790 (0.172) and 0.800 (0.180), respectively. The 95% CI of the mean difference was -0.024 to 0.006, which was within the equivalence interval. The ICC was 0.894 (95% lower CI 0.857), significantly different from 0.70. For the EQ VAS, the means (SD) were 72.0 (19.7) for paper and 74.1 (19.8) for IVR. The 95% CI of the mean difference was -3.784 to -0.484, partially within the equivalence interval. The ICC was 0.897 (95% lower CI 0.859) also significantly different from 0.70.

DISCUSSION

This analysis provides evidence that the EQ-SD scores on the IVR version were equivalent to those obtained on the original paper version.

ABSTRACT

OBJECTIVE: Electronic data capture technologies, such as interactive voice response (IVR) systems, are emerging as important alternatives for collecting patient-reported outcomes data. The objective of this study was to assess the measurement equivalence of an IVR version of the EQ-SD with the original paper version.

METHODS: This study utilized a crossover design with subjects randomly assigned to one of two assessment orders: 1) paper then IVR or 2) IVR then paper. A convenience sample of non-treatment outpatient cancer clinic patients (n=119) were asked to complete each assessment two days apart. The analyses tested for mean differences (repeated measures ANOVA) and intraclass correlation coefficients (ICC) to assess measurement stability over time. Equivalence of the means was established if the 95% confidence interval (CI) of the mean difference was within the minimally important difference (MID) interval: 0.035 to 0.035 for the index and -3 to 3 for the EQ VAS. Adequacy of the ICC was established by comparing the ICC 95% lower CI with a critical value of 0.70.

RESULTS: The per protocol analysis included 109 subjects for the EQ VAS and 113 subjects for the index. For the EQ-SD index, the means (SD) of the paper and IVR administrations were 0.790 (0.172) and 0.800 (0.180), respectively. The 95% CI of the mean difference was -0.024 to 0.006, which was within the equivalence interval. The ICC was 0.894 (95% lower CI 0.857), significantly different from 0.70. For the EQ VAS, the means (SD) were 72.0 (19.7) for paper and 74.1 (19.8) for IVR. The 95% CI of the mean difference was -3.784 to -0.484, partially within the equivalence interval. The ICC was 0.897 (95% lower CI 0.859) also significantly different from 0.70.

DISCUSSION: This analysis provides evidence that the EQ-SD scores on the IVR version were equivalent to those obtained on the original paper version.

REFERENCES


In addition to the descriptive system, the EQ-SD also has a visual analog scale (EQ VAS) to measure an individual’s overall health status. The original EQ VAS is a thermometer-like 20-cm vertical line with endpoints labeled “worst imaginable health state” and “best imaginable health state” anchored at 0 and 100, respectively.

The EQ-SD descriptive system was adapted to the IVR system using the exact wording for the items and responses. For the EQ VAS, the IVR system asked respondents to “picture in their minds” a scale with 100 at the top (i.e., “best health state you can imagine”) and 0 at the bottom (i.e., “worst health state you can imagine”) and enter a number between 0 and 100 representing their health status.

DATA ANALYSIS: Descriptive statistics of cancer type, age, and sex were calculated to characterize the sample, no subgroup analyses were performed. All statistical analyses were performed using SPSS version 16.0 (Chicago, Illinois) and evaluated using an alpha level of 0.05.

Mean Differences: Testing of the mean differences was based on analysis of variance (i.e., split-plot ANOVA) with factors for mode, period of administration (first or second) and subject; the p-values from the significance tests will be reported. The split-plot ANOVA also accounts for the interaction effect (period x mode effect, often termed as carryover). The adjusted mean differences between modes were estimated together with the associated 95% confidence interval (CI) for the difference. Equivalence on this measure was considered to have been established if the 95% CI excludes the MD used in the sample size calculations, namely 0.07 for the EQ index and 6 points for the EQ VAS.

Reliability: To analyze the reliability of the instrument, the analyses were carried out upon the ICC (Shrout and Fleiss 1979). It approaches 1.0 when the between-groups effect is very large relative to the within-groups variance, indicative of the grouping variable having no effect. The ICC is calculated based on the ANOVA model that includes factors for mode and subject. A one-sided 95% CI for the lower bound was computed using the formula provided in McGraw and Wong (1996). Measurement equivalence was considered to have been established if the lower bound of the 95% CI exceeded 0.70.

RESULTS: A total of 184 subjects agreed to participate. Of those, 139 subjects completed both administrations for a participation rate of 75.5% (Figure 2). The respondents were 67.6% female and had a mean age of 61.5 years. The ages ranged from 19 to 86.

The analyses for the index score and the EQ VAS were based on analysis of variance (i.e., split-plot ANOVA) with factors for mode, period of administration (first or second) and subject; the adjusted mean differences between modes were estimated together with the associated 95% confidence interval (CI) for the difference. Equivalence on this measure was considered to have been established if the 95% CI excludes the MD used in the sample size calculations, namely 0.07 for the EQ index and 6 points for the EQ VAS.

Reliability: To analyze the reliability of the instrument, the analyses were carried out upon the ICC (Shrout and Fleiss 1979). It approaches 1.0 when the between-groups effect is very large relative to the within-groups variance, indicative of the grouping variable having no effect. The ICC is calculated based on the ANOVA model that includes factors for mode and subject. A one-sided 95% CI for the lower bound was computed using the formula provided in McGraw and Wong (1996). Measurement equivalence was considered to have been established if the lower bound of the 95% CI exceeded 0.70.

MEAN DIFFERENCES: For the EQ-SD index, the means (SD) of the paper and IVR administrations were 0.172 and 0.880 (0.180), respectively. The tests for an order effect and for an order by mode interaction based on the split-plot ANOVA were not statistically significant. The adjusted means (i.e., least squares means (SE)) were 0.789 (0.016) for the paper version and 0.798 (0.017) for the IVR version. The adjusted mean difference was 0.009 and the 95% CI of the mean difference was -0.204 to 0.006, which was within the equivalence interval.

The EQ VAS means (SD) were 72.0 (19.7) for paper and 74.1 (19.8) for IVR. Similarly, no order effect or mode by order interaction was present in the analysis of means for the EQ VAS. The adjusted means were 72.03 (1.85) for the paper and 74.16 (1.89) for the IVR. The adjusted mean difference (SE) was 2.13 (0.83), which was associated 95% CI of the mean difference was 3.39 to -0.484, partially contained within the equivalence interval of -3 to +3.

CONCLUSION

The evidence presented here, when taken in totality, supports the measurement equivalence of the IVR version of the EQ-SD with the original paper version. In conclusion, we recommend that the data from this IVR version of the EQ-SD be treated as equivalent to that of the paper version.