

Electronic self-report: How do we evaluate success in the migration and use of electronic versions of symptom scales and HRQOL questionnaires?

Brian Tiplady, PRO Consulting
James Pierce, Perceptive Informatics
Barbara Marino, PHT Corporation
ISOQOL 30th October 2009

Two main ePRO Modalities

IVR

- Uses ordinary household telephone
- Spoken presentation of questions and response options
- Responses made by key press

Device

- Uses small computer (PDA or Tablet)
- Questions and response options displayed on screen
- Responses made by tapping screen

Question 3

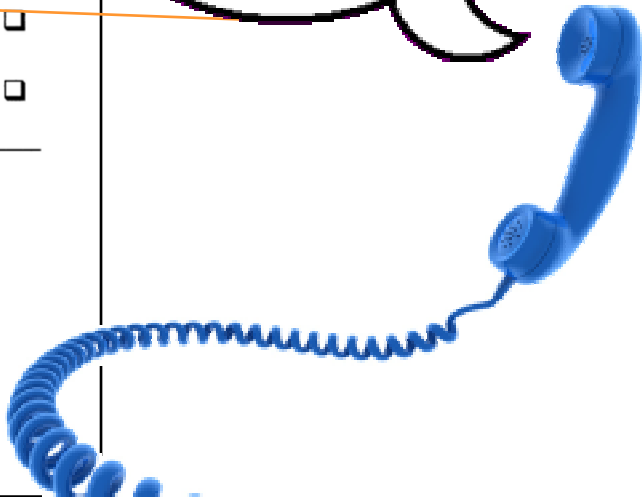
During the last week, how many hours did you spend on each of the following activities? Please tick one box only on each row.

	None	Some but less than 1 hour	1 hour but less than 3 hours	3 hours or more
a. Physical exercise such as swimming, jogging, aerobics, football, tennis, gym workout, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Cycling, including cycling to work and during leisure time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. Walking, including walking to work, shopping, for pleasure, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d. Housework / Childcare.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e. Gardening / DIY.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Walking, including walking to work, shopping, for pleasure, etc.

Key

Inactive	Sedentary job and no recreational physical activity
Moderately inactive	Sedentary job and some but less than 1 hour recreational physical activity per week OR Standing job and no recreational physical activity
Moderately active	Sedentary job and 1-2.9 hours recreational physical activity per week OR Standing job and some but less than 1 hour recreational physical activity per week OR Physical job and no recreational physical activity
Active	Sedentary job and 3 hours or more recreational physical activity per week OR Standing job and 1-2.9 hours recreational physical activity per week OR Physical job and some but less than 1 hour recreational physical activity per week OR Heavy manual job



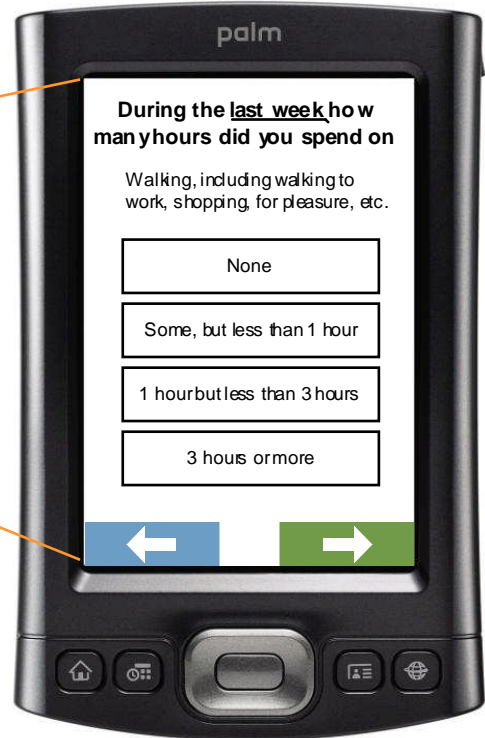
Question 3

During the last week, how many hours did you spend on each of the following activities? Please tick one box only on each row.

	None	Some but less than 1 hour	1 hour but less than 3 hours	3 hours or more
a. Physical exercise such as swimming, jogging, aerobics, football, tennis, gym workout, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Cycling, including cycling to work and during leisure time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. Walking, including walking to work, shopping, for pleasure, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d. Housework / Childcare.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e. Gardening / DIY.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Key

Inactive	Sedentary job and no recreational physical activity
Moderately inactive	Sedentary job and some but less than 1 hour recreational physical activity per week OR Standing job and no recreational physical activity
Moderately active	Sedentary job and 1-2.9 hours recreational physical activity per week OR Standing job and some but less than 1 hour recreational physical activity per week OR Physical job and no recreational physical activity
Active	Sedentary job and 3 hours or more recreational physical activity per week OR Standing job and 1-2.9 hours recreational physical activity per week OR Physical job and some but less than 1 hour recreational physical activity per week OR Heavy manual job



What could make a difference?

- Memory
 - If material is presented on more than one screen, memory load may be increased
 - Delay between presentation of spoken questions and response options may increase memory load
- Presentation
 - Single versus grouped questions
 - Changed orientation of responses from horizontal to vertical
 - Visual versus verbal presentation

Previous work (see e.g. Couper et al., 2001) suggests that these effects are small, if they occur at all.

Planning for successful ePRO migration

- Think about what you are doing before you do it
 - Extent of validation work depends on changes made
- Understand psychometric properties of original paper based instrument
 - Proper expectation of “equivalence”
- Make sure your work is consistent with good science

Guidelines for validation work in Coons et al. (2009)

Establishing equivalence quantitatively

- Comparing scores from both modes of administration
- Cross over – all subjects use both modes
- Subjects assigned at random to assessment sequence (E-P or P-E)
- Correspondence between measures assessed:
 - Important to assess both correlation (ICC) and differences between scores
- Ensure adequate power (see e.g. McEntegart, 2010)

Presentation and Interpretation of Results

- Agreed criteria for equivalence
- Reporting standards
 - Details of ePRO application
 - Population characteristics, e.g. severity
 - Order effects, setting
- The patient's view
 - Ease of use
 - Understandability

Respiratory Case Studies

Asthma Quality of Life Questionnaire developed by Juniper et al. (1993). Current version is AQLQ(S).

- Adapted for use with IVR, PDA and tablet PC.
- Studies to test systems for ease of use, accuracy of understanding and agreement with paper originals.
 - IVR: 54 adults with moderate to severe asthma. Paper version completed in the clinic followed by the IVR version later that evening, usually at home.
 - PDA: 84 asthmatic patients in randomised crossover in single clinic session.



IVR and Paper

- Two questionnaires compared RQLQ(S) and AQLQ(S)
- ICCs 0.92 and 0.94
- Differences in mean scores,
 - 3.3 and 1.8% of scale length
 - Significant for RQLQ
 - Bias in opposite directions for two scales
- Results considered to show unsatisfactory equivalence

Source: Juniper et al. (2009)

Reported results do not accurately reflect study design limitations

- According to authors, study did not provide clear evidence of equivalence and suggested bias
- Subjects were not randomized: possible order effects
 - First assessment rated more severe than second
 - Interval between assessments sufficient to allow change of health status
- IVR and paper were in different settings
 - IVR at home paper in clinic
 - Inconsistent effects
- Results underscore importance of ensuring validation studies are appropriately designed

PDA and Paper

- Three questionnaires compared, AQLQ(S), ACQ and RQLQ(S)
- ICCs in the range 0.84 – 0.92
- Differences in mean scores,
 - ~ 1 - 3% of scale length
 - Significant for two scales, AQLQ and RQLQ
 - Electronic tended to show higher severity
- Results considered to show unsatisfactory equivalence

Source: Juniper et al. (2009)

Equivalence of Device-based ePRO to Paper-based PRO

- Meta-analysis of equivalence studies
- Included comparison of PC, Tablet and PDA implementations (E) with paper originals (P)

Volume 11 • Number 2 • 2008
VALUE IN HEALTH

Equivalence of Electronic and Paper-and-Pencil Administration of Patient-Reported Outcome Measures: A Meta-Analytic Review

Chad J. Gwaltney, PhD,^{1,5} Alan L. Shields, PhD,^{2,5} Saul Shiffman, PhD^{3,4,5}

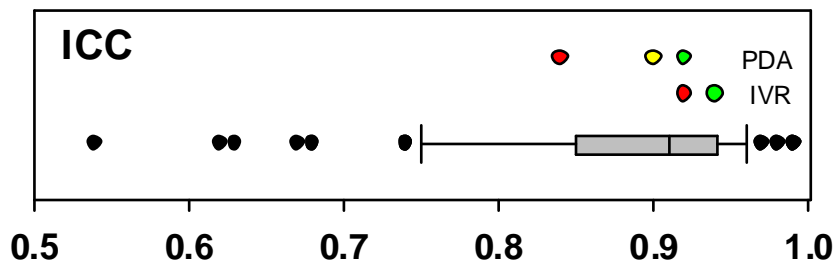
¹Brown University, Providence, RI, USA; ²East Tennessee State University, Johnson City, TN, USA; ³University of Pittsburgh, Pittsburgh, PA, USA; ⁴Invivodata, Inc., Pittsburgh, PA, USA; ⁵PRO Consulting, Pittsburgh, PA, USA

Meta-analysis

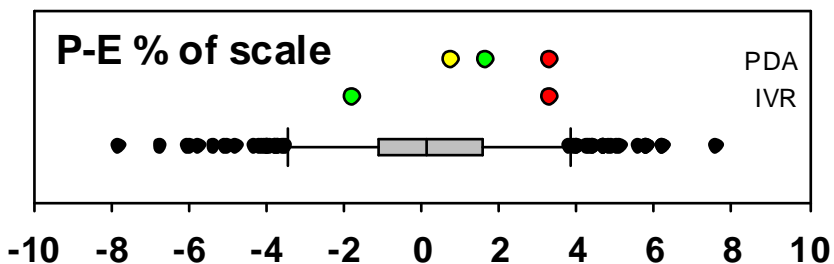
- Review of 46 studies evaluating 278 scales
- Correlations (mostly intraclass correlation, ICC) ranged from 0.67 – 0.98
- 94% of studies had correlations ≥ 0.75
- Mean E-P discrepancy 0.2% of scale length
- Mean absolute E-P difference 2%

- Data for IVR suggest similar pattern

Comparison of Equivalence with Meta-analysis



- AQLQ
- ACQ
- RQLQ



Meta-analysis
(re-drawn from
Gwaltney et al. (2008))

Intraclass Correlation

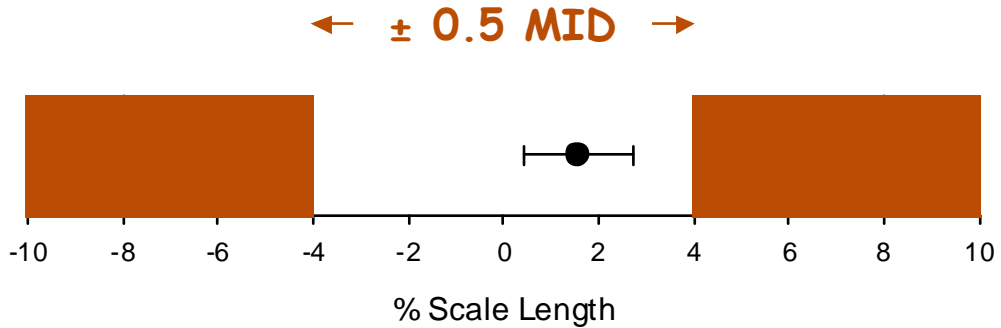
- What is an appropriate target for ICC?
 - Gwaltney et al (2008) suggest 0.75 as lower limit
 - Juniper et al specified 0.95 for AQLQ and RQLQ
- It is reasonable to take paper retest reliability into account when interpreting ICC values
 - P-E equivalence cannot be expected to be better than the reliability of the instrument assessed with P-P retest

Electronic-Paper Difference

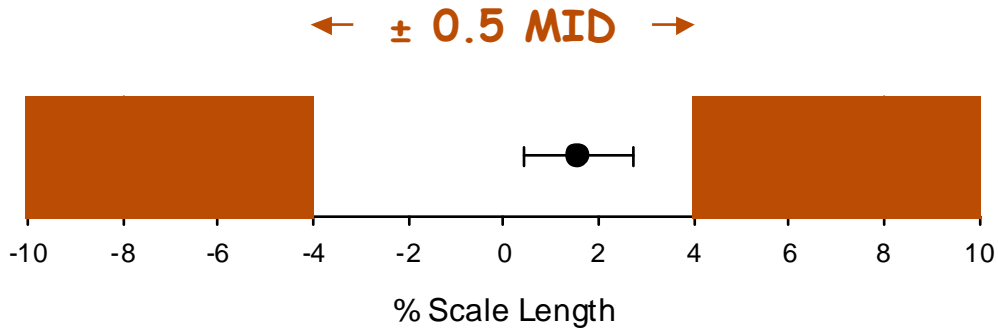
- What does a difference of $<2\%$ of scale length signify?
- Can be compared to Minimum Important Difference (MID)
 - MID of AQLQ is about 0.5 scale units or 8% of scale length.
 - P-E difference for AQLQ (1.7%) is much smaller than MID
 - How likely is this to affect a clinical trial outcome?

Analogy with bioequivalence approach

AQLQ: 95% Confidence Limits



AQLQ: 95% Confidence Limits



- Bias is small, and CIs are contained within 0.5 MID (bioequivalence criterion)
- Bias is consistent for the 3 respiratory scales

Summary

1. Importance of Design issues
 - Randomisation
 - Patient population
2. Criteria for equivalence
 - Acceptable values for ICC
 - Relate E-P to Minimal Important Difference
 - Relate to paper retest reliability

References

- Coons, S. J., Gwaltney, C. J., Hays, R. D., Lundy, J. J., Sloan, J. A., Revicki, D. A., Lenderking, W. R., Cella, D., & Basch, E. (2009) Recommendations on Evidence Needed to Support Measurement Equivalence between Electronic and Paper-based Patient-Reported Outcome (PRO) Measures: ISPOR ePRO Good Research Practices Task Force Report, *Value in Health*, 12: 419-429
- Couper, M. P., Traugott, M. W., & Lamias, M. J. (2001) Web survey design and administration, *Public Opinion Quarterly*, 65: 230-254
- Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008) Equivalence of Electronic and Paper-and-Pencil Administration of Patient-Reported Outcome Measures: A Meta-Analytic Review, *Value in Health*, 11: 322-333
- Juniper, E. F., Guyatt, G. H., Ferrie, P. J., & Griffith, L. E. (1993) Measuring quality of life in asthma, *Am.Rev.Respir Dis.*, 147: 832-838
- Juniper EF, Schatz M., Juniper BA., Zeiger RS (2008) Evaluation of 2 interactive voice response telephone versions of health related quality of life questionnaires. *Journal of Clinical Immunology* 5(3): 1-2
- Juniper, E. F., Langlands, J. M., & Juniper, B. A. (2009) Patients may respond differently to paper and electronic versions of the same questionnaires, *Respiratory Medicine*, 103: 932-934