Computerized administration of health-related quality of life instruments compared to interviewer administration may reduce sample size requirements in clinical research: a pilot randomized controlled trial among rheumatology patients

J. Thumboo[1,2], H.-L. Wee[1], Y.-B. Cheung[3], D. Machin[3,4], N. Luo[5], D. Feeny[6], K.-Y. Fong[1,2]

[1]Department of Rheumatology & Immunology, Singapore General Hospital; [2]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore; [3]National Cancer Centre Singapore; [4]Clinical Trials & Epidemiology Research Unit; [5]Department of Pharmacy, National University of Singapore, Republic of Singapore; [6]Health Utilities Inc., Canada.

Abstract

Objectives

Computerized health-related quality of life (HRQoL) administration may facilitate clinical trials incorporating HRQoL assessment in rheumatology patients by reducing sample size requirements. We tested this hypothesis in a pilot randomized controlled trial.

Methods

Chinese-speaking adult rheumatology outpatients were randomized to computerized (PC) or interviewer (IA) administration of the EQ-5D (utility & VAS), Health Utilities Index (HUI2 & HUI3) and Family Functioning Measure (FFM). We compared measurement variability (i.e., variance) between PC and IA for each instrument before (Levene's test) and after adjusting for the effects of age, gender and education (multivariable modeling) and computed the variance ratio (VR) for PC over IA.

Results

In 138 patients (mean age: 48), the mean (SD) time for administration was similar for PC (n = 67) and IA (n = 71) at 17.7 (7.94) versus 17.3 minutes (7.49), respectively. More subjects expressed a preference for PC (n = 21) over IA (n = 13). Mean HRQoL scores were not significantly different for PC versus IA except for higher VAS scores with IA (difference -7.7, 95% CI –14.0 to 1.3, p = 0.018). Variances and adjusted VR were smaller with PC for the EQ-5D (adjusted VR 0.34, 95% CI 0.18 to 0.65), HUI3 (0.49, 0.27 to 0.89) and FFM (0.95, 0.61 to 1.46), but larger for the HUI2 (1.30, 0.67 to 2.55) and VAS (1.05, 0.55 to 2.00).

Conclusions

The reduced variability in 3 of 5 instruments and good acceptance of computerized HRQoL assessment, if confirmed in larger studies, may lead to smaller sample size requirements, with potential reductions in cost and recruitment time for clinical trials and cohort studies.

Key words

Computers, data collection, health surveys, quality of life, sample size, statistics.

Julian Thumboo, FRCP (Edin); Hwee-Lin Wee, PhD; Yin-Bun Cheung, PhD; David Machin, PhD; Nan Luo, PhD; David Feeny, PhD; Kok-Yong Fong, FRCP (Edin).

This study was funded by programme grant 03/1/27/18/226 from the Biomedical Research Council of Singapore.

Please address correspondence to: Dr. Julian Thumboo, Dept of Rheumatology and Immunology, Singapore General Hospital, Outram Road, Singapore 169608.
E -mail: julian.thumboo@sgh.com.sg

Received on July 3, 2006; accepted in revised form on February 26, 2007.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2007.

Competing interests: Dr. David Feeny has a proprietary interest in Health Utilities Incorporated, Dundas, Ontario, Canada. HUInc. distributes copyrighted Health Utilities Index (HUI) materials and provides methodological advice on the use of HUI.

Introduction

Measurement of health-related quality of life (HRQoL) in clinical research and routine care is important for improving outcomes in rheumatology patients (1, 2). Self-administration is the simplest means of administering HRQoL instruments, but is hampered by low literacy, which is present in more than 20% of the world population (3). Biased findings are likely to result from the exclusion of subjects who cannot self-administer questionnaires. Interviewer administration (IA) is therefore often used in subjects with low literacy, but is costly and logistically difficult because well-trained bilingual interviewers are often in short supply. These issues are compounded by the increasingly multi-ethnic nature of populations worldwide. Computerized (PC) administration of HRQoL instruments (using an interactive touchscreen (4)) in subjects with low literacy (5) has been well accepted by patients, does not require the presence of a bilingual interviewer and may thus potentially address the issues of illiteracy, cost and logistic difficulties.

Computerized HRQoL assessment also offers several other advantages. First, it enables data to be collected, scored and reported in real time, with the potential for improving physician-patient communication (6). Second, missing data are less common with computerized compared to the pen-and-paper assessment of HRQoL (7). Third, subjects can complete surveys in privacy, which is helpful in eliciting answers to questions that are potentially sensitive (8). Furthermore, as sample size requirements are directly proportional to the amount of variability in the data (*i.e.*, the greater the variability, the larger the sample size required), we postulated that computerized HRQoL assessment might have the additional advantage of reducing sample size requirements by reducing the variability arising from the use of several interviewers (currently the only available approach for HRQoL assessment among low literacy subjects). This could provide greater efficiency in HRQoL assessment, with shorter recruitment times and reduced costs. We also sought to determine if the mean HRQoL scores for both PC and IA would be similar (because a similar stimulus to subjects would be provided by both modes of administration), which would allow the pooling of data obtained using these modes of administration. The aims of this pilot study were therefore to compare, using a randomized trial design, the variability and mean scores in rheumatology patients completing four generic HRQoL instruments based on either PC or IA administration.

Methods

Key features of the smiling touch screen

We developed a multimedia touchscreen program (the Smiling Touchscreen (9), Fig, 1) for use among subjects with varying levels of functional and computer literacy. We chose to develop the program in Chinese as this is the commonest language used worldwide and in Singapore (10). Nevertheless, computerized administration, if proven effective, could conceivably be adopted for use in other languages or clinical populations. The program was developed using an iterative process of discussion among the investigators with patient input regarding the screen design and layout, and final development and testing in a convenience sample of rheumatology outpatients. The key features of the Smiling Touchscreen that make it user friendly for subjects with low levels of literacy include: (a) the presentation of each question individually with both visual and auditory stimuli, (b) voice-text synchronization, which allows subjects with low reading literacy to follow the audiovisual playback with relative ease, (c) providing a choice of audiovisual playback speeds (fast, normal or slow) to suit the subject's literacy level, (d) demonstration and practice screens to familiarize the patient with the system, (e) replay buttons for the question stem and *individual* response options so that subjects may listen to these without repeating the entire question, and (f) large font size to cater to elderly subjects (in whom low literacy is particularly prevalent). The program runs on a Windows-based personal

computer, with responses captured in ASCII format.

Patients and study design

Recruiters enrolled Chinese-speaking patients being seen for routine care at a rheumatology outpatient clinic in a tertiary acute-care hospital for this Institutional Review Board-approved study. Inclusion criteria were age 21 and above, ability to converse in Chinese, and the absence of cognitive impairment as assessed by the recruiter. Consenting patients were randomized to PC or IA administration of four generic HRQoL instruments: the EQ-5D (utility & VAS), the Health Utilities Index Mark 2 (HUI2) & 3 (HUI3), and the Family Functioning Measure (FFM)) administered in the stated order.

In the IA branch of the study, face-to-face interviews (not self-administration) were performed by 4 different interviewers and in the PC branch, subjects completed these instruments using the Smiling Touchscreen by themselves after a facilitator provided a brief orientation to the Smiling Touchscreen (for patients who had never previously used a touchscreen computer). Subjects assigned to PC administration were interviewed to obtain socio-demographic information and feedback on the Smiling Touchscreen. Reading and computer literacy were assessed by subject self-reporting using a 0-100 visual analogue scale (VAS); this was done because at the time of the study there were no measures of literacy validated for use in Singapore.

Before this study commenced, interviewers underwent training and participated in several supervised mock interviews. Facilitators were similarly trained in interview techniques and shown how to demonstrate the use of the Smiling Touchscreen to subjects who were unfamiliar with touchscreens or computers. Subjects completed the survey using the Smiling Touchscreen by themselves, with facilitators available to answer queries (required by a minority of subjects; see Results). Assignment of the patients to one of the two study branches was concealed from recruiters, patients and interviewers, and was performed independently by a study coordinator (who was not a recruiter or interviewer) who randomized patients through a unique patient study number using a randomization list (generated by STATA Intercooled v.8, STATA Corporation, College Station, 2003). Study monitoring was also performed by the study coordinator.

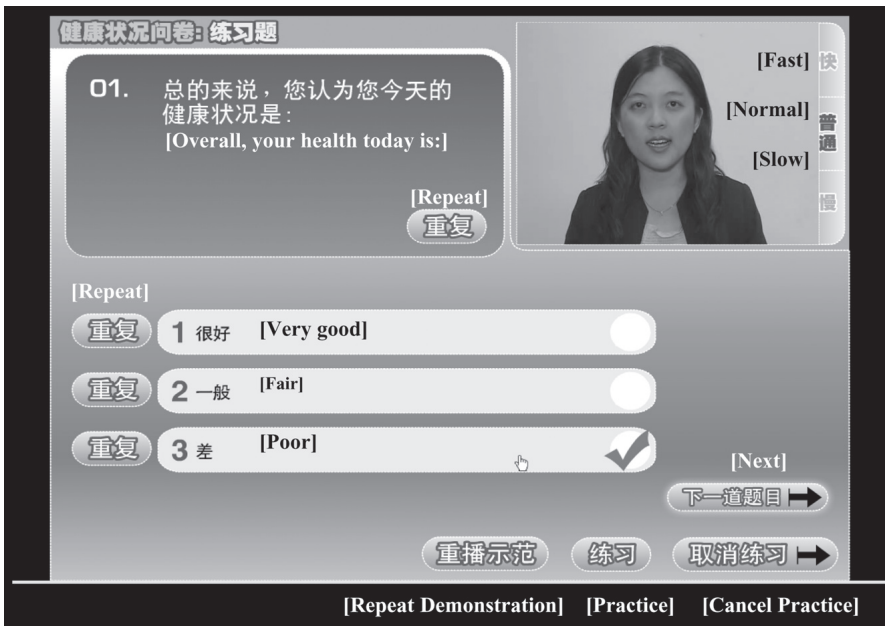

Fig. 1. Screenshot of the Smiling Touchscreen.
English translation of the contents is given in square brackets.

Instruments

The Chinese versions of the EQ-5D (11), HUI2 (12), HUI3 (11, 12), and FFM (13) used in this study have previously

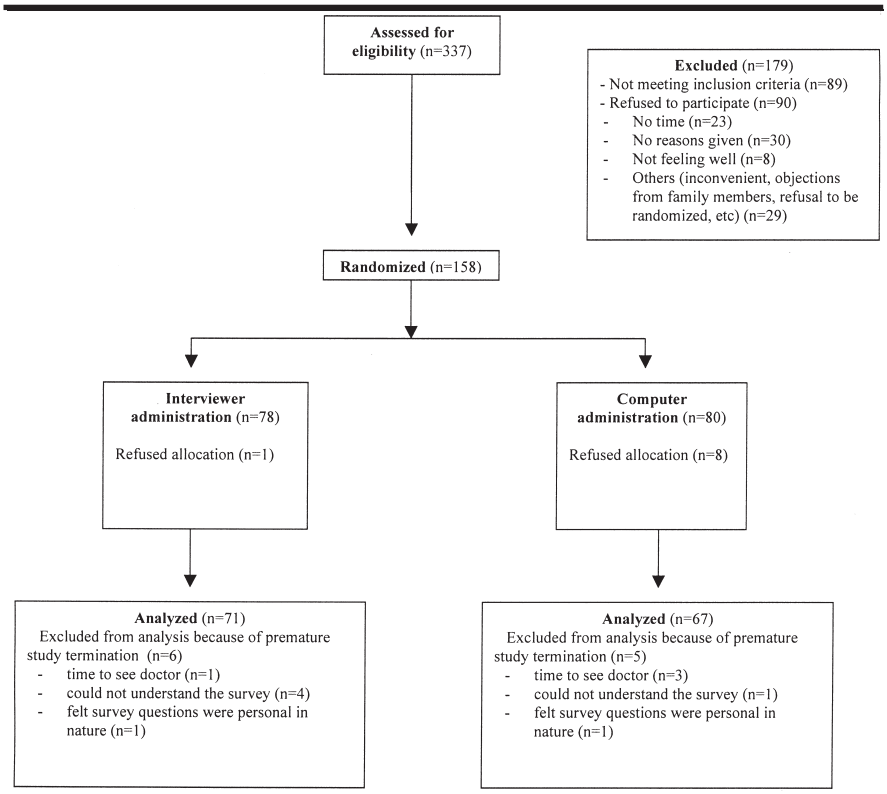

Fig. 2. Flowchart of recruitment of study participants.

been validated for use among Singaporeans. The EQ-5D consists of a self-classifier (5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression with 3 levels of severity: none, moderate and severe, giving rise to 243 health states) and a visual analogue scale (VAS, range 0 to 100) (14). Each health state is associated with a utility score, calculated using the U.K. utility function (as a Singapore utility function is not available) (14). The HUI2 and HUI3 are complimentary instruments. The HUI2 consists of 7 attributes [sensation, mobility, emotion, cognition, self-care, pain and fertility (optional, and not assessed using current HUI questionnaires)], with 3 to 5 levels per attribute, describing 24,000 unique health states (15). The HUI3 consists of eight attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain/discomfort), with 5 to 6 levels per attribute, describing 972,000 health states (16). Both the HUI2 and HUI3 utilities were calculated using the utility function for Canadians (as a Singapore utility function is currently not available) (15, 16). The FFM, previously validated in Singapore, is a 3-item Likert-type scale assessing the quality of interactions among family members, with higher scores (range 0 to 100) reflecting better family functioning (12).

Statistical methods

Differences between 2 groups were analysed using Mann-Whitney (continuous variables) or Chi-squared (categorical variables) tests. Measurement variability was assessed using variance (*i.e.,* the squared product of standard deviation). A variance ratio (VR) was calculated by taking the square of the standard deviation in scores for PC over that for IA (*i.e.* $[SD_{PC}/ SD_{IA}]^2$). Hence, a smaller VR would indicate smaller measurement variability for PC administration. Comparisons for each outcome score between PC and IA were made using Student's t-test for means and Levene's test for equality of variances (17). As standard deviation also varies with age, gender and education, we adjusted for these factors using a linear regression model, which allowed for non-constant variances across groups (18). The VR was estimated by the maximum likelihood method and confidence intervals were estimated from robust standard errors (18). Data were analyzed according to a modified intention-to-treat principle, where patients who did not complete the study using either mode of administration were excluded from the analyses. We then illustrated how a reduction in VR would reduce sample size requirements based on published formulae for sample size calculation, where sample size is directly proportional to variance (19).

Table I. Characteristics of study participants randomized to computerized or interviewer administration.

	Computerized administration (n = 67)	Interviewer administration (n = 71)
Female, no. (%)	51 (76)	55 (78)
Mean (SD) age in years	47.5 (11.9)	48.1 (14.1)
Years of education, no. (%)*		
≤ 6	14 (21)	14 (20)
7-10	28 (42)	27 (38)
≥ 11	24 (37)	30 (42)
Median (interquartile range) self-reported literacy using Visual Analogue Scales†		
Reading	70.0 (30.0, 90.0)	65.0 (30.0, 90.0)
Listening	80.0 (70.0, 100)	90.0 (75.0, 100)
Writing	50.0 (0, 70.0)	30.0 (0, 80.0)
Median (interquartile range) self-reported computer literacy using Visual Analogue Scales†	70.0 (0, 100)	60.0 (0, 100)
Medical diagnoses		
Inflammatory arthritis	22 (33)	31 (44)
Connective tissue diseases or systemic vasculitis	22 (33)	19 (27)
Osteoarthritis or soft tissue rheumatism	16 (24)	9 (13)
Combination of the above conditions	4 (6)	5 (7)
Others‡	3 (5)	7 (10)

*One patient (assigned to the PC arm) declined to provide information on education level and was thus excluded from further analysis using listwise deletion.
† Score range: 0 to 100.
‡ Other rheumatic diseases included osteoporosis, Raynaud's phenomenon, etc.

Results

Patient characteristics

As shown in Figure 2, among 337 contacted patients, 89 were not eligible and 90 declined to participate. Of the 158 randomized patients (63.7% of eligible patients), 9 declined (8 after allocation to PC but before seeing the Smiling Touchscreen system, and 1 IA) and 11 terminated the study prematurely (5 PC, 6 IA, $p = 0.843$). Thus 67 patients completed the PC and 71 completed the IA. The mean (SD) time taken to complete the evaluation using either mode of administration (PC *vs.* IA: 17.7 (7.9) *vs.* 17.3 (7.5) minutes, $p = 0.659$) was very similar. The characteristics of these patients (mean (SD) age: 47.8 (12.7), range 21.0 to 78.0 years, female: 77%) are given in Table I and were generally similar between the two groups. Participants and decliners were also similar in terms of age (mean (SD): 53.9 (14.1), range 33.0 to 76.0 years, $p = 0.207$) and gender (75% female, $p = 0.964$).

In general, the Smiling Touchscreen was well accepted, with 90% (72/ 80) of assigned patients agreeing to use this system. Subjects who terminated the study prematurely (n = 5) did so for reasons not related to difficulties with the Smiling Touchscreen. Among subjects expressing a preference (based on

Table II. Mean scores, variances and variance ratios of EQ-5D utility, EQ-VAS, HUI2 utility, HUI3 utility and FFM scores for computerized versus interviewer administration.

	Mean (range) score (95% CI)	Standard deviation (variance)	Difference in means (95% CI) (p values)[†,‡]	Variance ratio[†,§]	Adjusted variance ratio (95% CI) (p values)[†,‖]
EQ-5D utility					
Computerized	0.788 (-0.074, 1) (0.741, 0.834)	0.191 (0.036)	0.028 (-0.046, 0.102) (0.460)	0.61	0.34 (0.18, 0.65) (0.001)
Interviewer	0.760 (-0.074, 1) (0.702, 0.818)	0.245 (0.060)			
EQ-VAS					
Computerized	67.1 (0,100) (62.0, 72.3)	21.0 (441.0)	-7.7 (-14.0, 1.3) (0.018)	1.64	1.05 (0.55, 2.00) (0.883)
Interviewer	74.8 (10, 100) (70.9, 78.7)	16.4 (269.0)			
HUI2 utility					
Computerized	0.788 (0.148, 1) (0.752, 0.823)	0.145 (0.021)	-0.038 (-0.083, 0.006) (0.723)	1.54	1.30 (0.67, 2.55) (0.439)
Interviewer	0.826 (0.348, 1) (0.798, 0.854)	0.117 (0.014)			
HUI3 utility					
Computerized	0.699 (-0.144, 1) (0.649, 0.749)	0.206 (0.042)	0.014 (-0.064, 0.091) (0.089)	0.67	0.49 (0.27, 0.89) (0.018)
Interviewer	0.685 (-0.177, 1) (0.626, 0.745)	0.251 (0.063)			
FFM					
Computerized	63.4 (16.7, 100) (58.0, 68.8)	22.1 (488.4)	1.4 (-6.3, 9.2) (0.711)	0.86	0.95 (0.61, 1.46) (0.807)
Interviewer	62.0 (0, 100) (56.3, 67.6)	23.9 (571.2)			

[†]Interviewer administration as reference group; [‡]Student's t-test; [§]Levene's test; [‖] estimated using maximum likelihood method with adjustment for age, gender and years of education.
CI: confidence interval; SD: standard deviation.

past experience with various modes of administration), 21 favored the Smiling Touchscreen over interviewer (n = 13) or self (n = 8) administration. The facilitators did not need to provide any technical help for over 65% of subjects. Only two subjects (9%) in the low literacy group (and no subjects in the high literacy group) required either continuous or a great deal of help with using the Smiling Touchscreen.

Comparison of mean scores and standard deviations for modes of administration

As seen in Table II, the difference in mean scores for subjects randomized to PC versus IA were generally small, with overlapping 95% confidence intervals, and did not reach statistical significance with the exception of the EQ-VAS. Variance ratios favoured PC administration (*i.e.*, were less than 1) for the EQ-5D, HUI3 and FFM, with and without adjusting for the influence of age, gender and education. Adjusted variance ratios reached statistical significance for the EQ-5D and FFM. Adjusted variance ratios for the HUI2 and EQ-VAS favoured IA (*i.e.*, exceeded 1) but did not reach statistical significance, and the 95% CI for these adjusted variance ratios does include values that would favor PC administration.

Discussion

In this pilot study, we have shown that computerized HRQoL assessment in rheumatology outpatients resulted in reduced variability (*i.e.*, variance) in the mean scores for some commonly used generic HRQoL instruments. This suggests that sample size requirements in such patients are likely to be smaller in clinical trials utilizing computerized rather than interviewer administration of HRQoL instruments. Given that sample size requirements are directly proportional to variance, sample size requirements could potentially be reduced by 66% and 51% respectively in clinical trials in rheumatology patients using the EQ-5D or HUI3 as the primary outcome measure. For example, in a study using the HUI3, interviewer administration would require a sample size of 1,099 whereas PC administration would require a sample size of 541 subjects (assuming that the mean (SD) HUI3 scores by PC and IA were 0.699 (0.176) and 0.669 (0.251), respectively, based on an adjusted VR of 0.49 as found in our study, with an 80% power to detect a difference of 0.03 points). This observation, if confirmed in larger studies, could have important implications in reducing the costs and recruitment time for such research. The use of computerised administration in HRQoL assessment in rheumatology patients is further supported by the fact that (a) more subjects expressed a preference for PC over IA or self administration, and (b) the time taken to complete PC and IA were similar. In addition to clinical trials, the computerized administration of HRQoL could also facilitate the collection of patient-reported health outcomes in long-term observational studies of rheumatology

patients, which are important to improve the routine clinical care of these patients (1).

Mean scores for all the outcomes variables (except VAS) were generally similar in this study. This is encouraging, and suggests that it may be possible to pool the data obtained using PC and IA. If confirmed, this would allow greater flexibility in administering some HRQoL instruments. However, the observation that differences in the mean scores approximated the proposed level of clinical significance for the HUI2 of 0.02 to 0.04 points (20) and approached that for the VAS (10 points) (21) is a concern. Nevertheless, the 95% confidence intervals for the mean scores found for both modes of administration overlap substantially. This pilot study thus provides justification for a larger study to definitively address this issue.

It is interesting that the study hypothesis appeared to hold true for all the outcomes variables except for HUI2 and EQ-VAS. The fact that patients completing the survey by computerized administration reported lower mean HUI2 and EQ-VAS scores suggests that there may be other factors at play. One factor could be the relatively small sample size, which would have increased the influence of outliers on the data analysis. For example, if outliers in the HUI2 and VAS scores were removed, the adjusted VRs would be reduced from 1.30 to 0.85 and from 1.05 to 0.83, respectively. Thus, we are optimistic that the study hypothesis may be supported for these instruments when a larger sample size is used. Other factors may be specific to the computerized administration of visual analogue scales. One specific factor could be the ways in which the VAS was presented. In the current system, patients either touched-and-dragged a horizontal bar superimposed on the VAS or used buttons on the screen to move the horizontal bar. Other ways of presenting the VAS could be tested to identify an optimal design. Another specific factor could be that patients may have problems with the self-completion of a VAS, and score variability with interviewer administration is reduced because patients have opportunities to clarify the use of VAS.

We considered using a cross-over design (rather than the current parallel groups design), in which each subject would have undergone both the computerized and the interviewer administration of the instruments. This would have allowed an intra-subject comparison of variability rather than the inter-subject comparison reported here, and would have increased the statistical power of the results. However, we decided against a cross-over design, because the associated carry-over effect (22) would have made the interpretation of the results difficult. A carry-over effect would have occurred because subjects undergoing interviewer administration would be likely to recall their responses, and would thus respond with identical or very similar answers when undergoing computerized administration, leading to an artificially low degree of variability (and vice versa).

We recognize that there are some limitations to this study. First of all, the small sample size may have limited our ability to demonstrate the equivalence of the mean HRQoL scores obtained using computerized or interviewer administration. Nevertheless, the results of this study – given its randomized design (which would reduce any potential biases) – are encouraging and do support the usefulness of further studies to confirm these observations. Second, the generalizability of our findings to other HRQoL instruments and to alphabet-based and other picto-gram-based languages needs further investigation.

In conclusion, these encouraging results, if confirmed in larger studies, would suggest that the reduced variability associated with computerized HRQoL assessment in clinical research on rheumatology patients could result in smaller sample size requirements, with potential reductions in cost and recruitment time when employed in clinical trials and cohort studies.

Acknowledgements

The authors wish to thank Ms Connie Tse and Mr Yeo Han Seng for their assistance with the logistics of this study.

References

1. PINCUS T, SOKKA T: Quantitative measures and indices to assess rheumatoid arthritis in clinical trials and clinical care. *Rheum Dis Clin North Am* 2004; 30: 725-51.
2. POLLARD L, CHOY EH, SCOTT DL: The consequences of rheumatoid arthritis: Quality of life measures in the individual patient. *Clin Exp Rheumatol* 2005; 23: S43-52.
3. UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION: *Education for All (EFA) Global Monitoring Report, 2002: Is the World on Track?* Paris, UNESCO, 2002.
4. GREENWOOD MC, HAKIM AJ, CARSON E, DOYLE DV: Touch-screen computer systems in the rheumatology clinic offer a reliable and user-friendly means of collecting quality-of-life and outcome data from patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2006; 45: 66-71 (E-pub 2005 Nov. 1).
5. HAHN EA, CELLA D, DOBREZ D *et al.*: The talking touch-screen: A new approach to outcomes assessment in low literacy. *Psychooncology* 2004; 13: 86-95.
6. VELIKOVA G, BOOTH L, SMITH AB *et al.*: Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol* 2004; 22: 714-24.
7. WILSON AS, KITAS GD, CARRUTHERS DM *et al.*: Computerized information-gathering in specialist rheumatology clinics: an initial evaluation of an electronic version of the Short Form 36. *Rheumatology (Oxford)* 2002; 41: 268-73.
8. TURNER CF, KU L, ROGERS SM *et al.*: Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 1998; 280: 867-73.
9. THUMBOO J, WEE HL, CHEUNG YB *et al.*: Development of a smiling touchscreen multimedia program for HRQoL assessment in subjects with varying levels of literacy. *Value Health* 2006; 9: 312-19.
10. GORDON RG, JR. (Ed): *Ethnologue: Languages of the World*, 15th ed., Dallas, SIL International, 2005.
11. LUO N, CHEW LH, FONG KY *et al.*: A comparison of the EuroQol-5D and the Health Utilities Index mark 3 in patients with rheumatic disease. *J Rheumatol* 2003; 30: 2268-74.
12. WANG Q, CHEN G: The health status of the Singaporean population as measured by a multi-attribute health status system. *Singapore Med J* 1999; 40: 389-96.
13. THUMBOO J, FONG KY, CHAN SP *et al.*: Validation of the medical outcomes study family and marital functioning measures in SLE patients in Singapore. *Lupus* 1999; 8: 514-20.
14. DOLAN P: Modeling valuations for EuroQol health states. *Med Care* 1997; 35: 1095-108.
15. TORRANCE GW, FEENY DH, FURLONG WJ *et al.*: Multi-attribute utility function for a comprehensive health status classification system: Health utilities index Mark II. *Med Care* 1996; 34: 702-22.
16. FEENY D, FURLONG W, TORRANCE GW *et*

al.: Multi-attribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002; 40: 113-28.

17. LEVENE H: Robust tests for equality of variances. *In* OLKIN I *et al*. (Eds): *Contributions to Probability and Statistics*. Stanford, Stanford University Press, 1960: 278-92.
18. CHEUNG YB, THUMBOO J, MACHIN D *et al*.: Modelling variability of quality of life scores: A study of questionnaire version and bilingualism. *Qual Life Res* 2004; 13 :897-906.
19. Inference. *In* LINDSEY JK: *Introductory Statistics: A Modelling Approach*. New York, Oxford University Press, 1995: 88-92.
20. SAMSA G, EDELMAN D, ROTHMAN ML *et al*.: Determining clinically important differences in health status measures: A general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics* 1999; 15: 141-55.
21. TODD KH: Clinical versus statistical significance in the assessment of pain relief. *Ann Emerg Med* 1996; 27: 439-41.
22. POCOCK SJ: *Clinical Trials: A Practical Approach*. Chicester, John Wiley and Sons, 1983: 111.