

Original investigation

Preliminary Evaluation of a New German Translated Tobacco Quality of Life Impact Tool to Discriminate Between Healthy Current and Former Smokers and to Explore the Effect of Switching Smokers to a Reduced Toxicant Prototype Cigarette

Anuradha Kulasekaran MBBS, MD¹, Christopher Proctor PhD¹, Ermioni Papadopoulou MSc¹, Christopher J. Shepperd BSc¹, Rick Guyer PhD², Barbara Gandek PhD^{2,3}, John E. Ware Jr PhD^{2,3}

¹Group Research and Development, British American Tobacco (Investments) Ltd, Southampton, United Kingdom;

²John Ware Research Group, Worcester, MA; ³Department of Quantitative Health Sciences, UMass Medical School, Worcester, MA

Corresponding Author: Anuradha Kulasekaran, MBBS, MD, Group Research and Development, British American Tobacco (Investments) Ltd, Regents Park Road, Southampton, SO15 8TL, United Kingdom. Telephone: +44(0)2380588635; Fax: +44(0)2380793800; E-mail: anuradha_kulasekaran@bat.com

Abstract

Background: Assessment of health-related quality of life (HRQoL) is well established in clinical research, but ceiling effects in validated tools might prevent detection of changes in well respondents. Tobacco Quality of Life Impact Tool (TQOLITv1) uses conceptual and psychometric advances to enhance detection of HRQoL changes.

Methods: In a 6-month, forced-switch study, the German TQOLITv1 was assessed in healthy adult (age 23–55 years) current and matched former-smokers. At baseline, smokers were switched to reduced toxicant prototype (RTP) or conventional cigarette for 6 months. TQOLITv1 responses were collected at baseline, 3 and 6 months from current smokers whilst former smokers completed it at the latter two time points. TQOLITv1 includes SF-36v2 and new smoking-specific, physical and general-health measures.

Results: Reliability at baseline was good (Cronbach's coefficient alpha > 0.70) for all measures. The baseline percentage with the best possible score (ceiling effect) for former and current smokers was substantially better for the new physical function than SF-36 physical function measure (35% vs. 59% at ceiling, respectively). New smoking-specific measures discriminated current from former smokers better than general health measures. Smoking-specific symptoms ($r = 0.73$) were more stable from baseline to 6 months than other measures ($r = 0.38$ – 0.54) particularly more than the SF-36 mental component score ($r = 0.24$). Although both product smoking groups worsened in most HRQoL measures, changes in general and smoking-specific HRQoL impact measures favored RTP smokers.

Conclusions: The German TQOLITv1 is sufficiently reliable and valid to assess HRQoL and may be more useful than SF-36v2 in evaluation of interventions in well smoking populations including those consuming RTPs.

Introduction

Health-related quality of life (HRQoL) is an established multidimensional concept with a well-defined framework for qualifying measures used in clinical trials.¹ HRQoL assessments have gained importance in clinical research, routine health care, and regulatory review of medical interventions. The SF-36v2 general health survey has been widely used in clinical trials² and in tobacco research mostly to differentiate HRQoL between current, former, and never smokers.³⁻⁶ Considerations related to whether some of the harm caused by tobacco products might be reduced with modified-risk tobacco products (MRTPs) have led the US Institute of Medicine (IOM) and Food and Drug Administration to consider the optimum study designs for their assessment.⁷ The IOM in its report, *Scientific Studies on MRTPs*⁸ considers HRQoL among other scientific evidence in providing guidance for evaluation of MRTP impact on public health.

The concept of MRTPs (initially referred to as potential reduced-exposure products) was introduced by the IOM in its 2001 report, *Clearing the Smoke: the Scientific Basis for Tobacco Harm Reduction*.⁹ Potential reduced-exposure products were defined as products that would substantially reduce exposure to one or more tobacco toxicants and could reasonably be expected to reduce the risk of one or more smoking-related diseases or other adverse health effects. Surveillance for smoking-related diseases and construction of aggregate measures of population health impact, including HRQoL, were suggested. In 2009, the Family Smoking Prevention and Tobacco Control Act¹⁰ gave the US Food and Drug Administration the authority to regulate all tobacco products to a population-based health standard, and the IOM advocated that composite outcomes would most accurately reflect general-health outcomes, in the same way that self-reported health status and disability-adjusted life years summarize health.⁸ So far only draft guidance on MRTP research is available.¹¹

Validated HRQoL instruments, such as SF-36v2, SF-6D, and EQ-5D¹² are widely used to contribute to important health-care decisions, but they generally have broad applications. The lack of measures specific to tobacco-related HRQoL potentially limits the usefulness of these tools in tobacco research. In a clinical study of reduced toxicant prototype (RTP) versus conventional cigarettes, the SF-36v2 general health survey was used by Shepperd and colleagues to explore the impact of switching to RTP in healthy smokers, as a secondary outcome.¹³ Primary outcomes such as some vapor-phase toxicants and particulate-phase toxicants were shown to be reduced in RTP smokers by biomarker assessment. However, unpublished SF-36v2 data from this study identified ceiling effects (ie, large proportions of respondents earning the best possible score), particularly for measures of physical functioning (49.6%), role-physical (62.5%), role-emotional (63.3%), and social functioning (74.1%).¹⁴ Similarly, a study of current ($n = 77$) and former ($n = 24$) smokers also identified ceiling effects on these four SF-36 scales.⁴ Instruments with ceiling effects cannot detect improvements in scores outside the range measured. Additionally, as suggested by a review of the literature and two publically available databases,¹⁴ positive health domains such as confidence in health and psychological well-being have not been the focus of tobacco-related research and might be better in discriminating between smoking and nonsmoking populations in comparison with other widely-used general HRQoL instruments.

Most clinical studies of exposure to tobacco toxicants enroll smokers with good health status.¹⁵⁻¹⁷ This approach seems appropriate to assess the performance of MRTPs before they reach the

market. However, to obtain robust data on changes in HRQoL in healthy tobacco users, improving the measurement range to raise the “ceiling” is an issue of importance, clinically, socially, and economically. The magnitude of change that is relevant for measures that raise the ceiling scores in healthy populations (ie, considered as a minimal important difference (MID) must be determined to support interpretation of impact of modified tobacco or nicotine products.

The Tobacco Quality of Life Impact Tool (TQOLITTMv1) was designed on the basis of end-point models that integrate clinical and HRQoL outcomes¹⁸ as recommended for the regulatory review of new drugs, to overcome measurement issues with legacy tools.¹ The conceptual framework of TQOLITv1 takes into account clinical parameters, new smoking-specific measures (both symptoms and impact on HRQoL attributable to smoking) and general health ([Supplementary Appendix 1](#)). The advantages of disease-specific measures have been achieved by focusing on specific symptoms such as smokers cough. Conceptually, however, this is not enough. Specific symptoms are not HRQoL because they do not capture and broadly represent life or its quality in terms of what people are able to do or how they feel. Life is affected to the extent that symptoms are severe enough to limit functional activities and how people feel in everyday life. For this reason, TQOLIT also measures the impact attributed to smoking in everyday life. The conceptual framework places emphasis on the sequence of events that occur in relation to changing smoking behaviors, that is, toxicant exposure may cause changes in clinical parameters that could result in symptoms, all of which eventually have an impact on HRQoL attributed to smoking and also general health outcomes. Advanced psychometric methods^{19,20} have been used to increase the range over which reliable scores can be measured, particularly ceiling scores, to enhance detection of HRQoL changes in healthy smokers. A certified German translation of the TQOLITv1 was prepared (see Methods below) and was administered for the first time in a 6-month clinical study of healthy current, former and never smokers.²¹ TQOLIT was a secondary outcome in that study making it possible to evaluate the psychometric performance (reliability, validity and stability) of previously used and new TQOLITv1 measures. Here we report validation tests of these measures and explore the ability of the instrument to detect changes in smokers HRQoL over time.

Methods

Study Design and Participants

The study was a single-center, single-blind, controlled clinical study with a forced-switching design that was conducted in Hamburg, Germany. The study was approved by the independent ethics committee of the Ärztekammer, Hamburg and registered (ISRCTN81286286). Participants were enrolled after they responded to advertisements in the local media and on the clinic's website. All participants were informed about the design and purpose of the study verbally and in writing, and all gave written informed consent which had information on the ethically approved stipend for participation.

Sample size calculations were based on the primary endpoints of the study (ie, biomarkers of exposure) using data from a previous RTP study.¹³ Sample size was calculated using MINITAB software version 15 and based on one-way analysis of variance. These calculations indicated that a sample size of 50 yielded at least an 80% power for all biomarkers of exposure included. A level of attrition during this longitudinal study was assumed and therefore 140 healthy adult current smokers of cigarettes with 6–8 mg International Organization

for Standardization tar yields were recruited. This smoking group was later split into approximately equivalent test and control groups, each group with a sample size of approximately 70. In addition, 60 healthy adult ex-smokers were recruited. The slight reduction in ex-smokers compared to smokers group sizes were based on an assumption that attrition would be lower for the non-smoking groups.

Participants underwent medical screening to ensure compliance with protocol inclusion and exclusion criteria.²¹ As a result 128 smokers and 58 ex-smokers were eligible for enrolment. Smokers were aged 23–55 years and ex-smokers 28–55 years. These study samples were deemed adequate for the evaluations of TQOLITv1 with methods used in previous psychometric evaluations.^{22,23} Current smokers had to have smoked regularly for at least 5 years smoking from the minimum legal smoking age in Germany (18 years). Former smokers had to have quit for at least 5 years after having smoked for at least 5 years. Participants were recruited and groups filled according to order of screening, age, gender, and availability at the start of the study, and the demographics for all groups were matched as far as possible. To ensure that enough participants were recruited to all groups, participant availability was assured, and groups were well matched for age and gender, full randomization was not possible. During the first 2 weeks, all smokers were supplied with a control conventional cigarette typical of products widely sold at the time. Numbers of cigarettes supplied were based on self-reported daily consumption at screening, plus two packs. Participants were not told at screening that their supply of cigarettes were free or would be based on their self-reported consumption. On Day 14 (baseline), half of smokers were switched to the RTP and half continued to smoke the control product for the remainder of the study. Smokers were allowed to smoke *ad libitum* and were asked to record daily consumption in electronic diaries every day for the duration of the study.

TQOLITv1 Modules, Administration, and Scoring

The survey modules in TQOLITv1 covered demographics and smoking-specific and general HRQoL measures. The new smoking-specific measures included symptoms and HRQoL impact attributed to smoking. New general measures included improved physical functioning and general health (GH) confidence measures (Table 1). For example, in contrast to SF-36v2 items, which define perfect physical health as the absence of limitations in physical activities, new items also measured how easy it was to perform them, in order to raise the ceiling. Further, evaluations of health in general (eg, excellent to poor) were expanded to include “confidence” in health at present and in the future in order to broaden that construct (Table 1). Other general measures administered included the eight health domains in the SF-36v2 as administered in previous studies. TQOLITv1 includes certified German-language translations. The SF-36 was translated using International Quality of Life Assessment Project methods.^{24,25} New smoking-specific and general items were translated into German using comparable methods, including forward/backward translation and lay person qualitative review, as documented by the Mapi Institute, Lyon, France.²⁶ TQOLITv1 was administered to smokers at baseline, 3 months and 6 months and to former smokers at 3 months (baseline) and 6 months. Completion of the survey was voluntary, although the research benefits were reiterated to participants as encouragement. Participants entered responses via an electronic data capture system (CRF Health, Helsinki, Finland) on a tablet device with a one-item-at-a-time interface. Clinic staff were given an instruction manual for the questionnaire¹⁴ to ensure that standards for administration of the survey were met, such as suitable

environments for administration, recommended time for completion (10–20 minutes) and scripts to confirm respondents’ ability to read the questions and to introduce and conclude administration. If respondents had queries about items, staff could read the item aloud verbatim but could not interpret the wording.

Responses captured in the electronic data capture were transferred to an accredited drug development system (Statistical Analysis System [SAS] Institute Inc., Cary, NC). Data quality was assessed for completeness of data, defined as the percentage of items answered across multi-item scale scores (count of the total number of items with valid responses divided by the total number of possible responses for each scale).

Norm-based scoring algorithms were used to calculate domain scores. *T*-score transformations were performed to achieve a mean of 50 and *SD* of 10 in the US adult general population surveyed in 2011 for all new measures.¹⁴ Published scoring algorithms for the SF-36v2 were used to compute the physical component summary (PCS) and mental component summary (MCS) scores.²⁷ For all general measures, higher scores indicate better health; for smoking-specific symptom and smoking impact measures, higher scores indicate worse health. For simplicity in this article, we refer to the best (healthiest) possible score in either direction as the ceiling. HRQoL scores were computed with licensed software (QualityMetric Health Outcomes Scoring Software 4.5 and John Ware Research Group Inc. QOLIX).

For all HRQoL measures, a three-point (0.30 *SD* unit) minimal important difference (MID) threshold was defined as meaningful.^{2,27} This threshold was derived from various criteria, including self-evaluations of change, correlations with improvements in disease markers, predictive studies of hospitalization and mortality^{2,28–30} and variations in rates of achieving MID thresholds across conditions in extensive clinical trial evidence.²

Statistical Analysis

Descriptive statistics for demographic variables were expressed as mean \pm *SD* for continuous variables (age, average daily cigarette consumption in past 30 days and number of years of smoking) and as percentages for categorical variables (gender, employment, education, ethnicity).

For TQOLITv1 scales and general measures, adjusted means for current and former smokers were calculated at study entry. To compensate for data imbalances, a regression model was used for each measure that included smoking status (current or former), age and categorical variables.

Pearson product-moment correlations and estimated point-biserial correlation coefficients were used to evaluate associations between pairs of measures (eg, symptoms and HRQoL impact) and between a measure and smoking status (eg, symptoms and current smoking status). It was hypothesized that favorably-scored general and unfavorably-scored smoking specific measures would be negatively correlated. Stability of measures over time was evaluated with product-moment correlations between measures at baseline and 3 months and 6 months for current smokers. Internal consistency reliability was estimated for all HRQoL measures with Cronbach’s coefficient alpha; a value of 0.70 or higher is recommended for group comparisons.³¹ Ceiling effects were examined separately for current and former smokers.

Exploratory analyses of changes from baseline to 6 months in RTP and control smokers included changes in least square means from repeated measures analysis of variance. For this assessment, a

model was determined by including categorical variables and smoking group (RTP or control, time point and change in cigarette consumption) and continuous variables for each smoking specific and general measure. Responder analysis (percentage of smokers with improved, unchanged or worsened scores, based on the MID threshold of 0.30 *SD* units for all measures) and effect sizes (magnitude of change) were additionally used to explore change over time. The effect size was calculated for all smoking-specific and general measures separately for each smoking group by subtracting the baseline mean scores from 6 months mean scores and dividing by the *SD* of baseline scores of all participants. Effect size was evaluated as small (0.20), moderate (0.50), or large (0.80) according to the guidelines proposed by Cohen.³²

Statistical analyses were performed with SAS (version 9.3) and SPSS (version 19). The significance level of the two-tailed tests was set at $P < .05$. Tukey's adjustment was used to account for multiplicity in repeated measures analysis of variance and regression modeling.

Results

Participants

Baseline demographic characteristics are described for smokers including those in the RTP and control groups, and former smokers (Supplementary Appendix 2). TQOLITv1 was completed by all enrolled current smokers ($n = 128$) at baseline and former smokers ($n = 58$) at 3 months. Data quality was good as determined by 100% completed item responses.

Reliability and Score Distributions

Reliability coefficients in current and former smokers exceeded 0.70 for all measures (Table 1), an accepted standard for group comparisons. As expected for young healthy smokers, ceiling effects (percentage with best possible score) were substantial for smoking-specific HRQoL impact and measures of physical functioning for both current and former smokers. However, no current smoker had the best score for smoking symptoms and only 5.2% of former smokers were at the ceiling on this measure (Table 1).

Comparison of new and SF-36v2 physical functioning measures indicated that the new measure substantially reduced the percentage of respondents with ceiling effects for current and former smokers combined (SF-36v2 59% vs. 35% for the new measure; Figure 1). However, the ceiling effect was not improved by the new GH-confidence measure when compared with the SF-36v2 GH measure (Table 1).

Validity

PCS and MCS scores did not differ significantly between current and former smokers at baseline (Table 1). However, the smoking-specific measures (symptoms and HRQoL impact) were significantly different (both $P < .001$) in favor of former over current smokers. Former smokers also had significantly better scores than current smokers on the new physical function measure.

Significant correlations observed between all measures at baseline were in the hypothesized direction for current and former smokers combined (Table 2). Correlations in current smokers were significant between smoking-specific impact and general measures as compared to weaker correlations for the same measures in former smokers. Significant correlations were seen between smoking-specific

symptoms with the SF-36v2 MCS and the new physical functioning and GH-confidence measures in current smokers (Table 2). Additionally, smoking-specific HRQoL impact correlated significantly with most measures among current smokers. In former smokers, correlations were generally weaker (Table 2), although significant correlations were seen between smoking symptoms and the SF-36v2 GH measure and the new GH-confidence measure. No significant correlations were seen for the smoking impact scale with other measures among former smokers.

Stability Over Time

All measures showed substantial stability from baseline to 3 months and less stability at 6 months from baseline in current smokers (Table 3). Stability estimates were generally large enough to justify the use of baseline measures as covariates in analyses of cigarette consumption over time. Two noteworthy exceptions were the SF-36v2 MCS and GH scales, for which stability estimates at 6 months dropped to 0.24 and 0.38 in current smokers (Table 3).

Comparison of HRQoL Scores in RTP and Control Smokers

Sixty one control smokers and 67 RTP smokers completed the study. Twelve smokers dropped out at 3 months of follow-up and two more dropped out at 6 months, with similar dropout rates in the control and RTP smoking groups. The reasons for dropouts included missed visits or noncompliance ($n = 8$), personal reasons ($n = 2$), adverse events ($n = 3$) and serious adverse events ($n = 1$).

Mean cigarette consumption was approximately 20 cigarettes per day at baseline in all smokers. After baseline, cigarette consumption increased in both groups of smokers with mean consumption at 6 months reaching 28 cigarettes per day in the RTP group and 29 cigarettes per day in the control group. Where consumption increased, this was reviewed with the individual and when appropriate, the supply adjusted accordingly.

RTP smokers did not worsen in SF-36 PCS and MCS scores, smoking-specific HRQoL impact and new general measures when compared to control smokers (Figure 2), and these findings were substantiated by the percentages who changed by the MID amount in the two groups (Supplementary Appendix 3). In control smokers, mean score change in smoking-specific HRQoL impact (baseline to 6 months: 58–62) was significant ($P = .01$), although the computed number of responders were relatively fewer than RTP smokers who had an insignificant worsening in smoking-specific HRQoL impact (Supplementary Appendix 3). Significant predictors for smoking-specific HRQoL impact were change in cigarette consumption, years of smoking, education and ethnicity.

Cohen's effect size indicates small to moderate values for control smokers compared to small values for RTP smokers (Supplementary Appendix 4).

Discussion

HRQoL has gained importance in clinical research of various interventions and is expected to have a similar role in tobacco research. TQOLITv1, a new survey tool that integrates new smoking-specific with widely-used general HRQoL measures, and improves at least one of the latter (new physical function measure), is available in

Table 1. Summary Information for Smoking-Specific and General Health-Related Quality of Life Measures

Measure ^a	N of items	Reliability ^b	% Best score at baseline ^c		Mean score (SD) at baseline	Floor definition	Ceiling definition
			Current smokers (n = 128)	Former smokers (n = 58)			
Smoking-specific symptoms ^d (-)	8	0.71	0.0	5.2	58.1 (8.6)	48.5** (6.6)	Constant occurrence of all smoking-specific symptoms, including smoker's cough, loss of taste and smell.
Smoking impact on HRQoL (-)	7	0.71	40.2	94.8	50.2 (6.4)	45.6** (2.9)	Very severe quality of life impact attributed to smoking.
SF-36v2 physical component summary (+)	35	0.72	0.0	0.0	56.3 (4.3)	56.2 (4.1)	Limitations in self-care, physical, social, and role activities, severe bodily pain, frequent tiredness, health rated "poor."
SF-36v2 mental component summary (+)	35	0.82	0.0	0.0	53.0 (6.0)	52.0 (6.0)	Frequent positive affect, absence of psychological distress and limitations in usual social/role activities due to emotional problems.
SF-36v2 physical functioning (+)	10	0.74	57.5	63.8	55.2 (3.1)	55.7 (2.4)	No limitations due to health in all types of physical activities including the most vigorous.
SF-36v2 general health (+)	5	0.74	14.2	8.6	56.0 (6.9)	55.0 (6.0)	Evaluates personal health as poor and believes it is likely to get worse.
New physical function (+)	4	0.79	33.1	37.9	55.0 (4.9)	56.4* (3.7)	Unable to do usual physical activities ranging from walking to playing sports.
New general health confidence ^e (+)	3	0.73	25.2	25.9	58.0 (7.6)	57.4 (7.7)	High self-evaluated confidence in current health and health outlook in the future.

HRQoL = health related quality of life. Baseline data is Month 0 for current smokers and Month 3 for former smokers. Floor and ceiling definitions for SF-36 measures are adapted from Ware.³⁴

^a(+) higher score indicates better health; (-) higher score indicates worse health.

^bCronbach's coefficient alpha at baseline.

^cBest possible score at baseline: % at floor for smoking symptoms and smoking impact, % at ceiling for all other measures.

^dBad breath, yellowing teeth, cold hands/feet, loss of taste/smell, nicotine-stained fingers/teeth, smoker's cough, hoarse voice, smell of smoke in clothes/hair.

^eThe new general health scale adds items that explicitly expand that evaluation to measure "confidence" in health now and the future.

* $P < .01$; **Adjusted means from the regression model analysis.

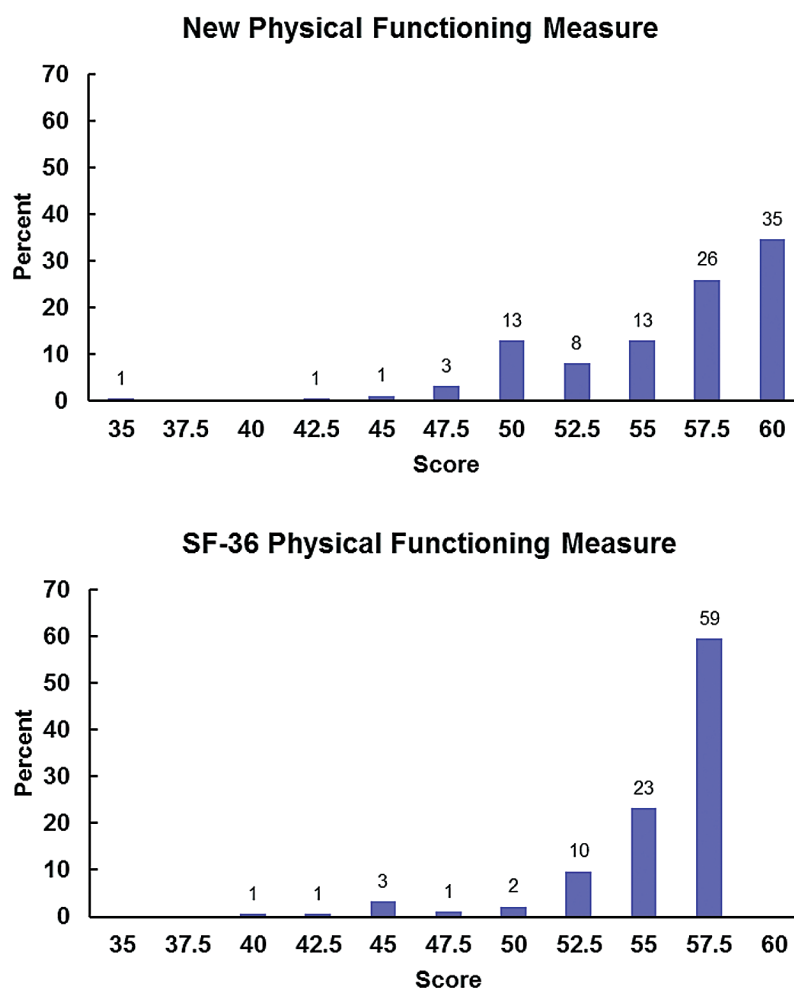


Figure 1. Improved score distribution for new and SF-36 physical functioning measures. *Note.* Percentages may not add to 100 due to rounding.

English and German. In this study the German version of TQOLITv1 was evaluated in cohorts of healthy current and former smokers in a cross-sectional analysis and over time. The improved psychometric properties of the new measures enhanced group comparisons, including confirming assumptions underlying the tool's construction and scoring of scales and leading to satisfactory score reliability and validity.

The significant correlations seen between smoking-specific measures and general measures in current smokers as compared to weaker correlations among former smokers support the validity of TQOLITv1 in measuring the effects of current smoking. The pattern of significant correlations at baseline between general and smoking-specific measures observed in this trial for the German translation is similar to the pattern observed for the English language version of TQOLITv1,³³ suggesting that the results regarding construct validity are generalizable.

The new TQOLITv1 physical functioning measure raised the ceiling enough to improve analyses involving relatively well scoring healthy smokers in comparison with the SF-36v2 physical function measure. The percentage of scores at the ceiling with the SF-36v2 was reduced from 59% to 35% with the new physical function measure. This extension in combination with the reduced ceiling effects for all the new smoking-specific measures, is likely to enhance the value of TQOLITv1 particularly for studies that

focus on well smokers. The new GH-confidence measure was not expected to increase the measurement range over the legacy GH measure as it was intended to provide conceptual improvements within the range measured by both. In addition, in light of the large ceiling effects observed in smokers for SF-36v2 social (50.7%) and role functioning (53.1%) measures in this study and in a previous study,¹⁴ improvements in items and response categories for these additional domains (analogous to those improved for physical functioning) warrant further testing to determine whether they also would better detect differences in outcomes between groups.

The stability of measures observed over the short term (3 months) in smokers and non-smokers was satisfactory to support their use in analyses of covariance. Stability of MCS dropped from baseline to 6 months in smokers due to a small number of outliers. Thus, it may not be as suitable in small, short-term studies of product effects performed with mixed model methods. Further evaluation is planned including another follow-up administration of TQOLITv1, 12 months after study completion.

All exploratory analyses with TQOLIT indicated a low magnitude of HRQOL changes in RTP smokers. TQOLITv1 detected changes with intervention, a significant change in smoking specific HRQoL impact in those smoking conventional cigarettes. The

Table 2. Correlations Between Measures at Baseline, Current, and Former Smokers

Group/measure ^a	Smoking status ^b	Smoking symptoms	Smoking impact	SF-36v2				
				PCS	MCS	PF	GH	New PF
Current and former smokers (n = 186)								
Smoking symptoms (-)	0.48**							
Smoking impact (-)	0.37**	0.36**						
SF-36v2 PCS (+)	0.01	-0.09	-0.09					
SF-36v2 MCS (+)	0.08	-0.14	-0.32**	-0.07				
SF-36v2 PF (+)	-0.09	-0.11	-0.28**	0.52**	0.20*			
SF-36v2 GH (+)	0.07	-0.15*	-0.20**	0.60**	0.39**	0.35**		
New PF (+)	-0.15*	-0.30**	-0.25**	0.40**	0.26**	0.61**	0.48**	
New GH confidence (+)	0.04	-0.25**	-0.27**	0.36**	0.36**	0.32**	0.62**	0.41**
Current smokers (n = 128)								
Smoking symptoms (-)								
Smoking impact (-)		0.25**						
SF-36v2 PCS (+)		-0.10	-0.17					
SF-36v2 MCS (+)		-0.23**	-0.43**	-0.07				
SF-36v2 PF (+)		-0.08	-0.33**	0.55**	0.22*			
SF-36v2 GH (+)		-0.17*	-0.26*	0.63**	0.33**	0.36**		
New PF (+)		-0.28**	-0.25**	0.44**	0.29**	0.63**	0.51**	
New GH confidence (+)		-0.29**	-0.37**	0.39**	0.32**	0.33**	0.60**	0.44**
Former smokers (n = 58)								
Smoking symptoms (-)								
Smoking impact (-)		0.10						
SF-36v2 PCS (+)		-0.16	0.18					
SF-36v2 MCS (+)		-0.12	-0.21	-0.05				
SF-36v2 PF (+)		-0.06	0.10	0.47**	0.19			
SF-36v2 GH (+)		-0.35**	-0.15	0.53**	0.55**	0.39**		
New PF (+)		-0.16	-0.00	0.32*	0.26*	0.50**	0.45**	
New GH confidence (+)		-0.34**	-0.03	0.30*	0.43**	0.30*	0.68**	0.37**

GH = general health; MCS = mental component summary; PCS = physical component summary; PF = physical functioning.

^a(+) higher score indicates better health; (-) higher score indicates worse health.

^bPoint-biserial correlation is statistically equivalent to a test of the difference between the two group means. Smoking status: current = 1, former = 0.

* $P < .05$; ** $P < .01$.

Table 3. Correlations Between Baseline and Follow-Up Measures, Current Smokers

Measure ^a	Baseline and 3 months (n = 116)	Baseline and 6 months (n = 114)
Smoking symptoms (-)	0.74	0.73
Smoking impact on HRQoL(-)	0.55	0.40
SF-36v2 PCS (+)	0.46	0.41
SF-36v2 MCS (+)	0.55	0.24*
SF-36v2 PF (+)	0.50	0.51
SF-36v2 GH (+)	0.66	0.38*
New PF (+)	0.72	0.54
New GH confidence (+)	0.50	0.41

GH = general health; HRQoL = health related quality of life; MCS = mental component summary; PCS = physical component summary; PF = physical functioning. All correlations are significantly different from zero ($P < .01$) with the exception of MCS at 6-month follow-up which is $P < .05$.

^a(+) higher score indicates better health; (-) higher score indicates worse health.

*6-month follow-up correlation significantly ($P < .05$) lower than 3-month follow-up correlation.

steeper declines seen in conventional than RTP smokers for other measures excluding smoking specific symptoms could potentially be due to an increase in cigarette consumption with the supply of free

cigarettes. This limitation of the study design needs consideration in future studies.

The estimated effect sizes for smoking-specific HRQoL impact and SF-36v2 MCS were smaller in the RTP group than in control group, which suggests greater HRQoL impact for smokers of conventional cigarettes. These exploratory findings need to be confirmed further in larger, more powerful studies that involve combustible RTPs, MRTPs (including alternative nicotine products such as e-cigarettes), or cessation with or without nicotine replacement therapies to confirm whether change of this degree or a greater degree could be detected. Likewise, the optimum times to observe changes in different domains of HRQoL as a result of individual behaviors such as the use of tobacco and/or nicotine products need to be determined.

The survey results reported are preliminary evaluations in a sample of very well-scoring healthy current and former smokers. Therefore, the outcomes cannot be generalized to the general population. In measures that continue to show ceiling effects, the range of reliable scores should be extended to enable more accurate assessment of the impact on HRQoL in specific populations (eg, healthy young non-smokers and consumers of cigarettes, MRTPs or alternative nicotine products) at individual and population levels. Further analysis of the HRQoL changes in relation to findings for other clinical parameters in this study measured at all-time point and in all groups such as biomarkers of exposure or effect is recommended to further confirm the end-point model underlying TQOLITv1.

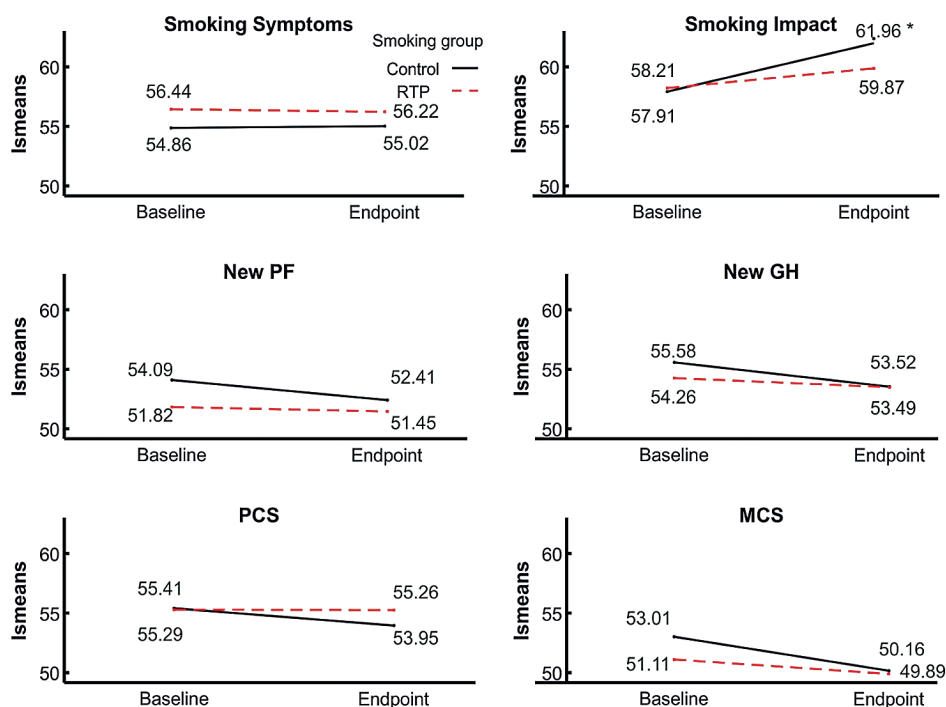


Figure 2. Descriptive statistics for changes in the least squares mean scores from baseline to 6 months. *Note.* Smoking impact = smoking specific health related quality of life impact; PF = physical function; GH = general health; PCS = physical component summary; MCS = mental component summary; RTP = reduced toxicant prototype. For smoking symptoms and impact: higher score indicates worse health; for new PF, GH, MCS and PCS: higher score indicates better health; *significant change in baseline scores at 6-month follow-up ($P < .01$).

Conclusions

The German translation of TQOLITv1 has psychometric properties that are sufficient for its use in studies of smoking populations. New TQOLIT smoking-specific HRQoL impact measures appear to be the most valid of those studied in discriminating between groups comprising individuals with different smoking behaviors. Particularly among well smokers, improvements in the range of measurement and reduced ceiling effects are likely to increase their sensitivity to changes over time. The improvements over SF-36v2 are likely to enhance evaluations of RTPs in well smoking populations.

Supplementary Material

Supplementary Appendixes 1–4 can be found online at <http://www.ntr.oxfordjournals.org>

Funding

British American Tobacco (Investments) Ltd provided an unrestricted research grant to John Ware Research Group Inc. for development and validation of TQOLITv1.

Declaration of Interests

All authors of British American Tobacco (Investments) Ltd are currently employed by the same company. Authors of John Ware Research Group Inc. performed the analysis and reported information related to development and validation of TQOLITv1.

Acknowledgments

We wish to thank Rosemarie Boulanger and Josh Ryan for administrative and technology support, respectively, at John Ware Research Group Inc.; Ingo Meyer,

the Principal Investigator at Momentum Pharma Services, MAPI for translation of TQOLIT; Audrey Richter, Nik Newland, Madeleine Ashley, and Justine Williamson at BAT for project related support. SF-36v2 is a registered trademark of Medical Outcomes Trust. TQOLIT is a trademark of John Ware Research Group.

References

1. FDA. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf. Accessed January 20, 2014.
2. Frenzl DM, Ware JE. Patient-reported functional health and well-being following drug therapy: a systematic review of randomized trials using the SF-36 Health Survey. *Medical Care*. 2014;52(5):439–445. doi:10.1097/MLR.00000000000010311.
3. Laaksonen M, Rahkonen O, Martikainen P, Karvonen S, Lahelma E. Smoking and SF-36 health functioning. *Prev Med*. 2006;42(3):206–209. doi.org/10.1016/j.ypmed.2005.12.003.
4. Olufade AO, Shaw JW, Foster SA, Leischow SJ, Hays RD, Coons SJ. Development of the Smoking Cessation Quality of Life questionnaire. *Clin Ther*. 1999;21(12):2113–2130. doi:10.1016/S0149-2918(00)87242-2.
5. Tillmann M, Silcock J. A comparison of smokers' and ex-smokers' health-related quality of life. *J Public Health Med*. 1997;19(3):268–273. <http://jpubhealth.oxfordjournals.org>. Accessed January 20, 2014.
6. Wilson D, Parsons J, Wakefield M. The health-related quality-of-life of never smokers, ex-smokers, and light, moderate, and heavy smokers. *Prev Med*. 1999;29(3):139–144. doi.org/10.1006/pmed.1999.0523.
7. Hatsukami DK, Hanson K, Briggs A, et al. Clinical trials methods for evaluation of potential reduced exposure products. *Cancer Epidemiol Biomarkers Prev*. 2009;18(12):3143–3195.
8. Henney JE, Baker TB, Bascom R, Biswal S; Institute of Medicine. *The Committee for Scientific Standards for Studies on Modified Risk Tobacco Products*. Washington, DC: National Academies Press; 2012.

9. Stratton K, Shetty P, Wallace R, Bondurant S. *Clearing the Smoke: Assessing the Science Base for Tobacco Harm Reduction*. Washington, DC: National Academies Press; 2001.
10. Family Smoking Prevention and Tobacco Control Act, 21 USC. 2009. www.fda.gov/TobaccoProducts/GuidanceComplianceRegulatoryInformation/ucm237092.htm. Accessed January 20, 2014.
11. FDA. Draft Guidance. Modified Risk Tobacco Product Applications. Rockville, MD: US Food and Drug Administration; 2012. www.fda.gov/downloads/TobaccoProducts/GuidanceComplianceRegulatoryInformation/UCM297751.pdf. Accessed January 20, 2014.
12. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16(3):199–208. www.health-policyjournal.com. Accessed January 20, 2014.
13. Shepperd CJ, Eldridge A, Camacho OM, McAdam K, Proctor CJ, Meyer I. Changes in levels of biomarkers of exposure observed in a controlled study of smokers switched from conventional to reduced toxicant prototype cigarettes. *Regul Toxicol Pharmacol*. 2013;66(1):147–162. doi:10.1016/j.yrtph.2013.02.007.
14. Ware JE, Guyer R. *How to Use the Tobacco Quality of Life Impact Test (TQOLIT@v1): User Manual*. Worcester, MA: JWRG Incorporated; 2013.
15. Benowitz NL, Dains KM, Hall SM, et al. Smoking behavior and exposure to tobacco toxicants during 6 months of smoking progressively reduced nicotine content cigarettes. *Cancer Epidemiol Biomarkers Prev*. 2012;21(5):761–769. doi:10.1158/1055-9965.EPI-11-0644.
16. Hatsukami DK, Lemmonds C, Zhang Y, et al. Evaluation of carcinogen exposure in people who used “reduced exposure” tobacco products. *J Natl Cancer Inst*. 2004;96(11):844–852. doi:10.1093/jnci/djh163.
17. Liu J, Liang Q, Frost-Pineda K, et al. Relationship between biomarkers of cigarette smoke exposure and biomarkers of inflammation, oxidative stress, and platelet activation in adult cigarette smokers. *Cancer Epidemiol Biomarkers Prev*. 2011;20(8):1–10. doi:10.1158/1055-9965.EPI-10-0987.
18. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA*. 1995;273(1):59–65.
19. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol*. 2008;61(1):17–33.
20. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol*. 2014;67(5):516–526. doi:10.1016/j.jclinepi.2013.10.024.
21. Shepperd CJ, Newland N, Eldridge A, Graff D, Meyer I. A single-blinded, single-centre, controlled study in healthy adult smokers to identify the effects of a reduced toxicant prototype cigarette on biomarkers of exposure and of biological effect versus commercial cigarettes. *BMC Public Health*. 2013 July 29;13:690. doi:10.1186/1471-2458-13-690.
22. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993;31(3):247–263.
23. Ware JE, Kosinski M, Bjorner JB, et al. Applications of computerized adaptive testing (CAT) to the assessment of headache impact. *Qual Life Res*. 2003;12(8):935–952.
24. Bullinger M. German translation and psychometric testing of the SF-36 Health Survey: preliminary results from the IQOLA Project. *International Quality of Life Assessment*. *Soc Sci Med*. 1995;41(10):1359–1366.
25. Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51(11):913–923.
26. Acquadro C, Conway K, Hareendran A, Aaronson N; European Regulatory Issues and Quality of Life Assessment (ERIQA) Group. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health*. 2008;11(3):509–521. doi:10.1111/j.1524-4733.2007.00292.x.
27. Ware JE Jr, Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. *User's manual for the SF-36® Health Survey (Version 2)*. 2nd ed. Lincoln, RI: QualityMetric Incorporated; 2007.
28. Guyatt GH, Naylor CD, Juniper E, Heyland DK, Jaeschke R, Cook DJ. Users' guides to the medical literature. XII. How to use articles about health-related quality of life. Evidence-Based Medicine Working Group. *JAMA*. 1997;277(15):1232–1237. doi:10.1001/jama.1997.03540390062037.
29. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77(4):371–383. doi:10.1016/S0025-6196(11)61793-X.
30. Norman G, Sloan J, Wyrwich K. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical Care*. 2003;41(5):582–592. <http://journals.lww.com/lww-medicalcare/pages/default.aspx>. Accessed January 20, 2014.
31. Nunnally JC. *Psychometric Theory*. 2nd ed. New York, NY: McGraw-Hill; 1978.
32. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
33. Ware JE, Kulasekaran A, Guyer R, Gandek B. Evaluation of the Tobacco Quality of Life Impact Tool (TQOLIT@v1) in current and former smokers in Germany and the United States. *Health Qual Life Outcomes*. In review.
34. Ware JE, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Medical Care*. 1995;33(4 Suppl):AS264–AS279.