MEASUREMENT EQUIVALENCE AND PATIENT PREFERENCE FOR THE SF-36v2® ON A HANDHELD DEVICE AND SMARTPHONE APP

Laura Khurana MPH, Michelle K. White PhD, Jakob B. Bjorner MD PhD, Sarah Gary PhD, Ellen Durand PhD, Antonio Otero, Susan Dallabrida PhD

1 ERT (F/K/A PHT Corporation), Boston, MA USA • 2 Optum (F/K/A QualityMetric), Lincoln, RI USA

OBJECTIVES
The Short-Form 36 Health Survey version 2.0 (SF-36v2®) is a validated patient-reported outcome instrument that measures functional health and well-being from the patient's point of view. Twelve items measure eight health domains: general health (GH), vitality (VT), social function (SF), role limitations due to emotional problems (RE), role limitations due to physical health (RP), bodily pain (BP), mental health (MH), and physical function (PF). The eight scales can be summarized in a physical component summary (PCS) and a mental health component summary (MCS).

Clinical trials are increasingly using FRID measures, and the SF-36® is the most frequently used FRID instrument in clinical trials. Although it was originally developed to support paper and pencil administration, the increased pace of electronic data collection in clinical trials necessitates that FRID instruments be adapted to electronic formats in ways that ensure measurement comparability to the original FRID-based instruments. In 2007, a single-item format (SIF) version of the SF-36v2 was developed for deployment on computer screens/tablet-sized devices. This version presents the items one at a time, and was tested in 2010 in response to increasing requests from clinical sponsors to accommodate smaller device screen sizes. A handheld SIF version was developed in 2010 which presents instructions on a single screen and each item/test alone without instructions as it is a true-on-subsequent-screen hardwired SIF version has not yet been validated. Although a similar version was tested in a rheumatoid arthritis population.

The objectives of this study were to evaluate measurement equivalence between the paper and electronic handheld SIF versions of the SF-36v2® administration using a handheld device as a smartphone app in four disease groups rheumatoid arthritis (RA), osteoarthritis (OA), diabetes, and hypertensive disease. The determinants of patient preference for mode of administration results presented here are the nine SF-36v2® general population group only, and thus only a summary analysis is possible.

METHODS
Subjects were recruited from the general public in the Boston area and screened over the phone. Eligible subjects were aged 40 or older, reported that they frequently test by a doctor that they have diabetes, high blood sugar, or sugar in their urine, and reported taking insulin or any other medication for their diabetes. 320 subjects participated in a randomized crossover study in which they completed the SF-36v2® on two modalities: paper and either the electronic handheld (PH) LogPad® UI (Figure 1A) or the smartphone App (PH/LogiD App™) (Figure 1B). Subjects completed the assessments in a single version with instruction completion between completion of the first and second modality (Figure 1C). After the study, subjects completed a survey of their preferred “mode of administration.” Subject responses to the paper SF-36v2® were entered into an electronic database for analysis. Mean score differences were analyzed as an electronic score minus paper score and 95% confidence intervals.

RESULTS
Subjects ranged in age from 30 to 79 years and were 51% female. Of the subjects had less than some college education. 30% reported owning a smartphone and 28% reported using the internet on a daily or weekly basis. About a quarter of subjects (25%) reported that they never use the internet or cannot recall the last time they did (Table 1).

Most subjects found the electronic mode either highly easy to use (80%) or easy to use (85%) and 79% were satisfied with the electronic mode faster to complete than paper. 75% found it more physically comfortable than the paper. There were 8% of subjects who preferred paper and 6% of subjects who preferred an electronic handheld mode over paper and 6% had no preference.

For subjects using electronic data capture on handheld devices, 50% of responses were captured for all items of the SF-36v2® whether subjects used the smartphone App or the LogPad App UI device. In contrast, 31% of SIF SF-36v2® administrations were missing at least one data point and as many as ten responses were missing. 23% of responses were missing on paper. Subject response rates on paper ranged from 71%-100% complete. Subjects most frequently omitted responses with multiple parts on paper and 50% of missing responses. The domains most affected were MH, VT, PF, RE and SF. 60% of the missing responses were from Question 2, which consists of nine MH and VT domain items.

Most subjects found the electronic mode either highly easy to use (80%) or easy to use (85%) and 79% were satisfied with the electronic mode faster to complete than paper. 75% found it more physically comfortable than the paper. There were 8% of subjects who preferred paper and 6% of subjects who preferred an electronic handheld mode over paper and 6% had no preference.

CONCLUSION
A study of 320 diabetes patients, we used a randomized cross-over design to compare paper and pencil administration of the SF-36v2® Health Survey with administration of a smartphone app and handheld device (LogPad). Out of 320 comparison, we found no score differences that exceeded the recommended MID. Comparison of the LogPad device and paper showed no significant score differences for any subscale or component summary. For comparisons of app and paper & pencil administration, we found score differences for the RE subscale and MCS that were below the recommended MID. Comparison of the LogiD app to paper administration, we found no scale differences found on any subscale or component summary. For comparisons of app and paper & pencil administration, we found score differences for the RE subscale and MCS that were below the recommended MID. These differences could be spurious due to multiple testing, or due to transferring the RE items from paper to smartphone app. Two subscales, RP and RE, showed no significant score differences for the remaining scales or the PF and MCS. Overall, confidence intervals are wide. ICS were low for IF and RE subscales.

REFERENCE

DISCLOSURES
The authors are employees of ERT (F/K/A PHT Corporation and Optum (F/K/A QualityMetric). For more information, contact ert.com/contact-us.

www.ert.com

Presented at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 20th Annual International Meeting
16-20 May 2015, Philadelphia, PA, USA