Qual Life Res (2011) 20:1513–1518
DOI 10.1007/s11136-011-9886-7

BRIEF COMMUNICATION

Examining the ability to detect change using the TRIM-Diabetes and TRIM-Diabetes Device measures

Meryl Brod · Torsten Christensen · Mette Hammer · Anne K. Busk · Donald M. Bushnell

Accepted: 7 March 2011 / Published online: 19 March 2011

Abstract
Purpose Responsiveness is defined as the ability of an instrument to accurately detect change when it has occurred and is an essential psychometric property of a patient-reported outcomes (PRO) measure to understand and interpret study findings. This study examined the responsiveness of 2 Treatment Related Impact Measures (TRIMs): The TRIM-Diabetes (TRIM-D) and TRIM-Diabetes Device (TRIM-DD) as well as confirmed their measurement models in a randomized controlled trial (RCT) design.
Methods The data were collected in a multi-center, randomized, open-label (2 × 12 week), cross-over study of two prefilled pens in subjects with type 1 or type 2 diabetes, age 18 or older. Internal and external responsiveness were examined. To confirm the measurement model identified in the previous study, the Bentler comparative fit index (CFI) and internal consistency for the RCT sample scores were examined and compared.
Results Based on a priori criteria, tests of responsiveness were confirmed with patients having significant improvements over time ranging from 2.7 (Psychological Health) to 11.1 (Treatment Burden) ($P < 0.01$) (effect sizes ranging from 0.2 to 0.8). The previous measurement model factor structure was confirmed (CFI ranging from 0.8 to 1.0), and internal consistency of the TRIMs was similar to the developmental findings.
Conclusions The total score as well as all domain scores of the TRIMs was significantly responsive over time, thus acceptable internal and external responsiveness of TRIM-D and TRIM-DD are concluded. To date, all validation evidence supports the use of these two measures in future clinical trials.

Keywords Diabetes Mellitus · Outcome Assessment · PRO Questionnaires · Validation Studies · TRIM-Diabetes · TRIM-Diabetes Device

Abbreviations

BMI	Body mass index.
CFI	Comparative fit index.
CFA	Confirmatory factor analysis
ES	Effect size
ITSQ	Insulin treatment satisfaction questionnaire
PRO	Patient-reported outcomes
RCT	Randomized controlled trial
RMSEA	Root mean square error of approximation
TRIM-D	TRIM-Diabetes
TRIM-DD	TRIM-Diabetes Device
TS	Treatment satisfaction

Introduction

Responsiveness is defined as an instrument's ability to accurately detect change that has occurred [1, 2]. Internal responsiveness is defined as an instrument's ability to change during a prespecified time frame. External

M. Brod (✉)
The Brod Group, 219 Julia Avenue,
Mill Valley, CA 94941, USA
e-mail: mbrod@thebrodgroup.net

T. Christensen · M. Hammer · A. K. Busk
Novo Nordisk A/S, Vandtaarnsvej 114, 2860 Soeborg, Denmark

D. M. Bushnell
Health Research Associates, 6505 216th Street SW,
Suite 105, Mountlake Terrace, WA 98043, USA

responsiveness is the extent to which a measure's degree of change corresponds to an external reference value or measure (assesses an instrument's ability to reflect both change and no change in the external standard) [3, 4].

This study examined the responsiveness of 2 patient-reported outcomes (PRO) measures, the Treatment Related Impact Measure-Diabetes (TRIM-D) and Treatment Related Impact Measure-Diabetes Device (TRIM-DD), which were developed as disease-specific PRO measures to assess the impact of diabetes treatment for both type 1 and 2 diabetes and across the spectrum of pharmacological treatments and delivery methods [5]. The TRIM-D is a 28 item measure with 5 domains assessing Treatment Burden, Daily Life, Diabetes Management, Compliance and Psychological Health. The TRIM-DD is an 8 item measure with 2 domains assessing Device Bother and Device Function. Both measures can be scored independently for each domain or as a total score. Higher scores indicate a better health state. The item generation and preliminary validation were conducted following FDA guidelines for PRO measures development [1]. Initial validation data for the TRIM-D and TRIM-DD were collected via an online, cross-sectional survey of 507 US patients. The cross-sectional validation showed that both measures have acceptable psychometric properties [5].

The purpose of the current study was to continue the validation process by examining the measures' responsiveness and to confirm the measurement model under randomized controlled trial (RCT) conditions.

Methods

The data used to assess responsiveness came from a multicenter, randomized, open-label, 2 × 12 week period cross-over study of two prefilled pens in subjects with type 1 or 2 diabetes. All subjects were using insulin by vial/syringe previous to inclusion in the study and were pen naïve. Data for these analyses came from all patients who had completed the TRIM-D and TRIM-DD at randomization (baseline) and time of cross-over (week 12). Non-superiority for glucose control between groups was hypothesized. The study was approved by Sterling IRB (approval #2925), and all persons gave informed consent.

Analyses

Analyses were conducted according to an a priori statistical analysis plan. All statistical tests were two-tailed and conducted with an alpha level of 0.05 as minimal threshold for significance. As the TRIM-D and TRIM-DD are intended to be used as either a total score or as independent domains, change scores were examined for both the totals and domain scores.

Responsiveness analyses

To examine internal responsiveness, *t* tests were used to examine differences in TRIM scores between baseline and week 12 (time of cross-over) with the expectation that significant improvement over time would be shown. Effect size (ES), measured by Cohen's *d*, was examined by calculating the mean change in score divided by the standard deviation of the mean baseline TRIM score. ES was categorized: small, 0.2–0.3; medium, 0.4–0.7; and large, 0.8 or above [6].

External responsiveness was examined by testing the hypothesis that there will be a linear relationship between the TRIMs and treatment satisfaction (TS) as assessed by the insulin treatment satisfaction questionnaire (ITSQ) [7]. The ITSQ, a disease-specific PRO assessing insulin TS, has been shown to be reliable and valid [7, 8]. Pearson correlation coefficients between the change in ITSQ overall summary score (from baseline to week 12) and the change in each item and domain of the TRIM-D and TRIM-DD were examined.

Confirmatory analyses of measurement model

A confirmatory factor analysis (CFA) was conducted using the Bentler comparative fit index (CFI) and root mean square error of approximation (RMSEA) to determine the goodness of fit between the models previously identified [5] and the current sample data. The criterion used to indicate acceptable fit was a CFI of at least 0.90 [9] and an RMSEA of 0.06 [9] or less.

Internal consistency reliability was examined and compared with the original sample with Cronbach's alpha, a statistic calculated from the pairwise correlations between items. Alphas range between zero and one, with coefficients of greater than 0.70 indicating acceptable reliability [10].

Results

In the cross-over study, 242 subjects completed the TRIM-D and TRIM-DD at baseline and week 12 (Table 1).

Responsiveness analyses

Internal responsiveness

All TRIM-D and TRIM-DD domains and overall total scores and most individual items (TRIM-D: 23/28; TRIM-DD: 6/8) changed significantly after 12 weeks of randomized treatment. For the Treatment Burden, Diabetes Management, Daily Life, and total TRIM-D, these significant change

Table 1 Sample description

Age	Mean (SD) ($n = 242$)	58.0 (13.9)
		Range 22–87
Gender	N (%) Male	147 (60.7%)
	N (%) Female	95 (39.3%)
Body mass index (BMI) at randomization	Mean (SD) ($n = 242$)	31.4 (6.1)
		Range 18.7–44.9
Diabetes type	N (%) Type 1	70 (28.9%)
	N (%) Type 2	172 (71.1%)
HbA1c at randomization	Mean (SD) ($n = 240$)	7.3 (0.9)
		Range 5.2–10.2
Ethnicity	N (%) White	199 (82.2%)
	N (%) Black	29 (12.0%)
	N (%) Asian	7 (2.9%)
	N (%) Other	7 (2.9%)

scores were associated with large to moderate ES. For the Psychological Health and Compliance domains, the significant change scores were associated with a small ES. Score changes ranged from 18.6 (ES 0.84, TRIM-D Treatment Burden) to 3.1 (ES 0.17, TRIM-D Psychological Health). For the TRIM-DD domains and total score, large changes (9.4–10.1) along with moderate ES (0.43–0.56) were seen (Table 2).

External responsiveness

Strong associations were found between the ITSQ change, TRIM-D Total score ($r = 0.72$, $P < 0.001$) and TRIM-DD Total score ($r = 0.68$, $P < 0.001$). Moderate to strong correlations were noted between the ITSQ overall summary score and items from the domains: Treatment Burden (r ranging between 0.32 and 0.53), Daily Life (0.37–0.45), Diabetes Management (0.22–0.38), Psychological Health (0.35–0.51), Device Function (0.30–0.51), and Device Bother (0.40–0.57). Lower associations were noted between ITSQ score and the Compliance domain (0.14–0.25).

Confirmatory measurement model analyses

Fit statistics

The model fit statistics for the TRIM-D and TRIM-DD Total domains were confirmed and are presented in Table 3.

Table 2 Responsiveness of the TRIM-Diabetes and TRIM-Diabetes Device items and domains

Abbreviated item content	Baseline Mean (SD)	Week 12 Mean (SD)	Change score Mean (SD)	*t*-stat	Effect size[a]	ITSQ overall summary (Pearson *r*)
TRIM-Diabetes TOTAL SCORE ($n = 226$)	65.9 (15.0)	74.2 (13.0)	8.3 (13.5)	9.2***	0.55	0.72**
Treatment Burden ($n = 225$)	54.7 (22.1)	73.3 (19.2)	18.6 (25.0)	11.1***	0.84	0.58**
The ease and convenience of your medication	3.4 (0.8)	4.0 (0.9)	0.6 (1.1)	8.7***	0.75	0.53**
Carry your medication and supplies around with you	2.8 (1.2)	3.7 (1.0)	1.0 (1.4)	10.2***	0.75	0.48**
Store your medication	3.2 (1.2)	4.0 (1.0)	0.8 (1.4)	8.1***	0.67	0.44**
Take your medication at the right time	3.3 (1.1)	3.9 (0.9)	0.65 (1.3)	7.5***	0.55	0.44**
Prepare your medication for use	3.4 (1.1)	4.2 (0.9)	0.8 (1.4)	9.0***	0.73	0.51**
Monitor your blood sugar as often as necessary	3.1 (1.1)	3.7 (1.0)	0.6 (1.2)	7.9***	0.55	0.32**
Daily Life ($n = 226$)	68.4 (18.5)	75.6 (16.8)	7.2 (17.9)	6.0***	0.39	0.58**
Meal time planning	3.6 (0.9)	4.0 (0.9)	0.4 (1.0)	6.1***	0.44	0.45**
Social activities	3.5 (1.1)	4.0 (0.9)	0.5 (1.1)	6.3***	0.45	0.45**
Do you have to limit your daily activities?	4.0 (0.8)	4.2 (0.8)	0.2 (1.0)	2.5*	0.25	0.39**
Do you accomplish less than you would like to?	3.5 (1.1)	3.8 (1.0)	0.2 (1.1)	3.3**	0.27	0.37**
Do you feel tension in your relationships with friends or family?	4.1 (0.9)	4.3 (0.9)	0.15 (1.0)	2.4*	0.22	0.39**
Diabetes Management ($n = 226$)	52.5 (19.2)	61.7 (17.9)	9.3 (19.2)	7.2***	0.48	0.43**
Help you control your diabetes	3.3 (0.9)	3.7 (0.8)	0.35 (1.0)	5.1***	0.44	0.38**
Help you avoid high blood sugar (hyperglycemia)	3.2 (1.0)	3.6 (0.9)	0.4 (1.2)	5.0***	0.40	0.34**
Help you avoid low blood sugar (hypoglycemia)	3.4 (0.9)	3.8 (0.8)	0.35 (0.9)	5.7***	0.44	0.31**
Help you manage your weight	2.7 (1.2)	3.0 (1.1)	0.35 (1.1)	4.8***	0.30	0.22**

Table 2 continued

Abbreviated item content	Baseline Mean (SD)	Week 12 Mean (SD)	Change score Mean (SD)	*t*-stat	Effect size[a]	ITSQ overall summary (Pearson *r*)
Help you prevent feeling tired or a lack of energy	2.8 (1.0)	3.3 (1.1)	0.4 (1.0)	6.3***	0.50	0.31**
Compliance ($n = 226$)	75.7 (17.0)	79.3 (15.0)	3.7 (15.1)	3.7***	0.22	0.30**
Miss a dose	4.3 (0.7)	4.4 (0.7)	0.1 (0.7)	2.5**	0.14	0.21**
Delay or postpone taking your medication	3.9 (0.9)	4.1 (0.8)	0.2 (0.8)	3.3**	0.22	0.14*
Take your medication at a different time than prescribed	3.9 (0.9)	4.1 (0.8)	0.1 (1.0)	2.2*	0.22	0.23**
Worry that you forgot to take/or missed your last dose of medication	4.0 (0.9)	4.2 (0.8)	0.2 (1.0)	2.5*	0.22	0.25**
Psychological Health ($n = 221$)	76.2 (18.6)	79.2 (17.2)	3.1 (16.9)	2.7**	0.17	0.59**
Depressed	4.3 (0.9)	4.4 (0.9)	0.1 (0.8)	1.5 ($P = 0.137$)	0.11	0.36**
Worried that the medication is not helping to slow down or prevent complications from my diabetes	3.9 (1.0)	4.0 (1.0)	0.1 (1.1)	0.7 ($P = 0.494$)	0.10	0.44**
Nervous or anxious	4.2 (0.8)	4.4 (0.8)	0.1 (0.9)	2.2*	0.25	0.41**
Worried about my blood sugar control	3.4 (1.1)	3.5 (1.0)	0.1 (1.1)	1.7 ($P = 0.084$)	0.09	0.47**
Unhealthy	4.1 (1.0)	4.1 (1.0)	0.05 (1.0)	0.8 ($P = 0.434$)	0.00	0.40**
Angry	4.3 (1.0)	4.4 (0.9)	0.05 (0.9)	0.8 ($P = 0.417$)	0.10	0.35**
Worried about side effects from my medication	3.9 (1.1)	4.2 (1.0)	0.25 (1.1)	3.2**	0.27	0.38**
Feel embarrassed or awkward when taking your medication	4.3 (0.9)	4.5 (0.7)	0.2 (0.9)	3.8***	0.22	0.51**
TRIM-Diabetes Device TOTAL SCORE ($n = 214$)	72.2 (17.1)	81.8 (15.2)	9.6 (20.6)	6.9***	0.56	0.68**
Device Function ($n = 214$)	71.6 (18.5)	81.0 (17.0)	9.4 (22.7)	6.0***	0.51	0.56**
Learn how to use your device	3.9 (0.9)	4.4 (0.8)	0.6 (1.1)	7.3***	0.56	0.47**
Keep your device functioning properly	4.1 (0.8)	4.4 (0.8)	0.3 (1.1)	4.2***	0.38	0.42**
Adjust your medication for small dose changes	3.8 (1.0)	4.4 (0.9)	0.6 (1.4)	6.4***	0.60	0.51**
That your device delivers the correct, full dose of your medication	3.7 (1.0)	3.8 (1.1)	0.2 (1.4)	1.7 ($P = 0.094$)	0.10	0.34**
That you are using the device properly	4.0 (0.9)	4.2 (1.0)	0.2 (1.2)	2.8**	0.22	0.30**
Device Bother ($n = 214$)	73.2 (23.3)	83.3 (19.5)	10.1 (26.1)	5.6***	0.43	0.64**
Size of your device	4.3 (0.9)	4.4 (0.9)	0.1 (1.2)	1.8 ($P = 0.076$)	0.11	0.40**
Physical discomfort related to using your device	3.9 (1.0)	4.4 (0.9)	0.5 (1.2)	5.5***	0.50	0.55**
Using your device in public	3.6 (1.4)	4.2 (1.1)	0.6 (1.4)	6.1***	0.43	0.57**

TRIM Treatment Related Impact Measure, *SD* Standard deviation, *ITSQ* Insulin treatment satisfaction questionnaire

[a] Effect size = mean change in score divided by the standard deviation of mean baseline score

*** Correlation is significant at the 0.001 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 3 TRIM-Diabetes and TRIM-Diabetes Device measurement model properties

	n	CFI	RMSEA	Chi-Square (Sig.)	*df*
TRIM-Diabetes total (28 items)	222	0.955	0.031	3602.4 ($P < 0.001$)	378
Treatment Burden (6 items)	235	0.972	0.020	763.7 ($P < 0.001$)	15
Daily Life (5 items)	235	0.818	0.072	447.6 ($P < 0.001$)	10
Diabetes Management (5 items)	235	0.888	0.051	498.2 ($P < 0.001$)	10
Compliance (4 items)	235	0.988	0.018	310.9 ($P < 0.001$)	6
Psychological (8 items)	229	0.948	0.037	923.5 ($P < 0.001$)	28

Table 3 continued

	n	CFI	RMSEA	Chi-Square (Sig.)	*df*
TRIM-Diabetes Device total (8 items)	226	1.000	0.000	722.6 ($P < 0.001$)	28
Device Function (5 items)	226	1.000	0.000	476.9 ($P < 0.001$)	10
Device Bother (3 items)	227	1.000	0.000	181.7 ($P < 0.001$)	3

CFI Comparative fit index, *RMSEA* Root mean square error of approximation, *df* degrees of freedom

Internal consistency

All alphas for the TRIM-D and TRIM-DD (overall score and all domains) were above 0.70 indicating acceptable internal consistency. Additionally, the confirmatory RCT sample alphas were similar to the development coefficients (within 0.1).

Discussion

These analyses found that the TRIMs total scores as well as all domain scores were significantly responsive over time and had the ability to differ between levels of change of an external criterion. Thus, internal and external responsiveness for the TRIM-D and TRIM-DD have been confirmed in an RCT sample. The measurement model was confirmed for all domains with lower than expected fit statistics for the Daily Life and Diabetes Management domains. Given that these domains were shown to have a strong factor structure in the development of the measures [3], this finding may be specific to this trial design or sample. Further testing the TRIM-D domain structure in other trials is warranted to confirm these findings.

The total score and all domain scores of the TRIMs were significantly responsive over time with the Treatment Burden domain showing the greatest responsiveness and the Psychological Health domain the least responsiveness. Additionally, the greatest number of individual items which were not responsive over time came from the Psychological Health domain. These findings should be interpreted in light of the study's nature. Given that all patients received the same insulin treatment, it is understandable that the psychological component of treatment, which is often driven by treatment efficacy, would be the least responsive. However, the fact that the overall Psychological Health domain was still significant as an overall concept and suggests that insulin pen delivery system does contribute positively to the psychological impact of treatment.

As expected, given that the study was a device cross-over with non-superiority for drug effect, the Treatment Burden domain, the domain which should be most impacted by delivery mode, was the most responsive domain. These findings underscore the importance of understanding the independent contribution of domains, given the specific study design and hypotheses, in order to optimally identify, a priori, domains of a measure which will be responsive to change. As the TRIMs were developed and validated for stand-alone use of each domain as well as the total score, future use of the TRIMs can and should take independent domain responsiveness into consideration when making these a priori hypotheses.

Certain study limitations should be considered in interpreting results. To assess external validity, the ITSQ, a PRO measure rather than a clinical measure, was used as the reference value. It was not possible to use a clinical reference value due to two factors. First, HbA1c $\leq$9% was a study eligibility criterion and the majority of patients entered the study in good or adequate HbA1c control (61%, <7.5). Thus, there could only be a limited number of patients who could change from inadequate to adequate glucose control. In fact, in this sample, there were only 11 patients (4.8%) who changed from randomization poor control (>7.0%) to adequate control over the 12-week period (<7.0%). Second, the study was designed as a non-inferiority trial to examine difference in insulin delivery mode rather than drug treatment efficacy, and all patients received the same insulin treatment during the study. Thus, no differences in glucose control were expected or found. As a result of these design features, there was not an adequate size sample of patients who had a significant improvement or worsening of HbA1c to conduct responsiveness analyses using a clinical reference value. Further, the fact that a majority of these patients were in good control at study start may limit the external generalizability of findings.

Validation is an iterative process. This study continues that process for the TRIM-Diabetes and TRIM-Diabetes Device measures. To date, all evidence supports the use of these measures in future clinical trials.

Acknowledgments This study was funded by Novo Nordisk. Dr. Brod and Mr. Bushnell are advisors/paid consultants to Novo Nordisk. Mr. Christensen, Ms. Hammer and Ms. Busk are employees and stakeholders of Novo Nordisk.

References

1. Food and Drug Administration (US). (2009). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
2. Beaton, D. E., Bombardier, C., Katz, J. N., & Wright, J. G. (2001). A taxonomy for responsiveness. *Journal of Clinical Epidemiology, 54*(12), 1204–1217.
3. Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: A critical review and recommendations. *Journal of Clinical Epidemiology, 53*(5), 459–468.
4. Matza, L. S., Johnston, J. A., Faries, D. E., Malley, K. G., & Brod, M. (2007). Responsiveness of the Adult Attention-Deficit/Hyperactivity Disorder Quality of Life Scale (AAQoL). *Quality of Life Research, 16*(9), 1511–1520.
5. Brod, M., Hammer, M., Christensen, T., Lessard, S., & Bushnell, D. M. (2009). Understanding and assessing the impact of treatment in diabetes: the Treatment-Related Impact Measures for Diabetes and Devices (TRIM-Diabetes and TRIM-Diabetes Device). *Health and Quality of Life Outcomes, 7*, 83.
6. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: N.J.:L. Erlbaum Associates.
7. Anderson, R. T., Skovlund, S. E., Marrero, D., Levine, D. W., Meadows, K., Brod, M., et al. (2004). Development and validation of the insulin treatment satisfaction questionnaire. *Clinical Therapeutics, 26*(4), 565–578.
8. Brod, M., Christensen, T., & Bushnell, D. (2007). Maximizing the value of validation findings to better understand treatment satisfaction issues for diabetes. *Quality of Life Research, 16*(6), 1053–1063.
9. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.
10. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.