An item-level response shift study on the change of health state with the rating of asthma-specific quality of life: a report from the PROMIS® Pediatric Asthma Study

Pranav K. Gandhi[1], **Carolyn E. Schwartz**[2,3], **Bryce B. Reeve**[4,5], **Darren A. DeWalt**[5,6], **Heather E. Gross**[5], and **I-Chan Huang**[7]

I-Chan Huang: i-chan.huang@stjude.org

[1]Department of Pharmacy Practice, School of Pharmacy, South College, Knoxville, TN, USA

[2]DeltaQuest Foundation, Concord, MA, USA

[3]Departments of Medicine and Orthopaedic Surgery, Tufts University School of Medicine, Boston, MA, USA

[4]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[5]Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[6]Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[7]Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, 262 Denny Thomas Place, Memphis, TN 38105, USA

Abstract

Purpose—To examine item-level response shift associated with the change in asthma-related health state (i.e., change in asthma control status and global rating of change (GRC) in breathing problems).

Methods—Study sample comprised 238 asthmatic children who were between 8 and 17.9 years and completed the Pediatric Asthma Quality of Life Questionnaire (PAQLQ) symptoms, emotion function, and activity limitation domains at baseline and a follow-up assessment. Structural equation modeling was implemented to assess item-level response shift associated with the change in asthma-related health state with the adjustment for the influence of confounding variables. The

Correspondence to: I-Chan Huang, i-chan.huang@stjude.org.

magnitude of item-level response shift and its influence on the change of domain scores was estimated using Cohen's effect sizes.

Results—We found no instances of item-level response shift. However, two items were identified with measurement bias related to GRC due to breathing problems. Specifically, asthmatic children with better/about the same GRC due to breathing problems reported lower scores for one item in the emotional domain at follow-up compared to those with deteriorated GRC due to breathing problems. In addition, asthmatic children with better/about the same GRC due to breathing problems reported better scores for another item in the symptom domain at baseline compared to those with deteriorated GRC due to breathing problems. The impact of measurement bias was small and did not bias the change of domain scores over time.

Conclusions—No item-level response shift, but two instances of measurement bias, appears in asthmatic children. However, the impact of these measurement issues is negligible.

Keywords

Asthma; Children; Measurement bias; PAQLQ; Response shift

Introduction

The comparison of health-related quality of life (HRQOL) scores over time is based on the premise that the meaning of concepts and the frame of reference for an individual remain consistent over time. However, when an individual's health state changes over time, he/she may also change his/her internal standards, values, and/or conceptualization of HRQOL. This is known as response shift phenomenon, defined as "a change in the meaning of one's self-evaluation of a target construct as a result of: (a) a change in the respondent's internal standards of measurement (i.e., recalibration); (b) a change in the importance of component scales constituting the target construct (e.g., reprioritization); or (c) a redefinition of the target construct (i.e., reconceptualization) [1, 2]." It is important to evaluate the extent to which response shift will bias the interpretation for the changes of HRQOL over time and a HRQOL instrument's ability to detect responsiveness to change [3]. Attempts to identify response shift effect in children and adolescents, although limited, have demonstrated the influence of disease progression on the adjustment and interpretation of HRQOL scores over time [4–6].

Changes in health state and clinical interventions are the primary catalysts for causing the occurrence of response shift [1, 7–10]. The initial statistical methods for detecting response shift effects relied on then-test approach to assess recalibration effects. This method suffers from reliability and validity problems and thus has fallen out of favor [11– 14]. Although there are a number of emerging methods for response shift detection [15, 16], Oort's structural equation modeling (SEM) measurement method [9] is one of the most comprehensive approaches for testing different forms of response shift. This approach addresses other measurement issues simultaneously that include measurement bias and response shift effects in measurement. This framework is presented in Fig. 1, where X represents the observed variables (e.g., PAQLQ items) on the latent construct (denoted by A) such as PAQLQ domain scores, explanatory variables (denoted by E) representing the causes

or predictors (e.g., change in asthma control status) of domain scores, and V representing the confounding variables (e.g., child's age and sex) that may influence domain scores. In Oort's framework, *response shift* effects are identified if the relationships between the item information capturing a specific HRQOL concept are changed over time, as estimated by the change in model parameters over time. *Measurement bias* refers to the inequality in item ratings given the same level of underlying HRQOL between different groups of respondents [9, 10, 17]. In the longitudinal study design, this concept corresponds to *response shift in measurement*, meaning the relationship between the item ratings and subgroups of respondents is not the same over time given the same level of HRQOL [9, 10, 17]. In other words, response shift is a special case of measurement bias.

The majority of previous studies have identified response shift effects using SEM at the domain level. They assessed the impact of response shift based on the change of HRQOL scores, comparing models that adjusted or did not adjust for response shift effects over time [18–21]. This traditional approach at the domain level neglects, however, the reality that response shift effects can take place at the item level. Limited evidence is available on response shift at the item-level and how item-level response shift influences the change in the estimated HRQOL scores at the domain level [8]. As more clinicians utilize short-form scales to capture the same concept as the long-forms in busy clinical practice, it is important to test whether response shift creates any bias in the short forms. In addition, previous studies were designed in the context of "pre-post events" such as the completion of invasive surgery in cancer patients [19, 20]. Few response shift studies have focused on "ongoing health states" (e.g., frequent asthma exacerbation). To the best of our knowledge, a SEM measurement model that determines response shift effects at the item-level after accounting for measurement bias and confounding variables has never been investigated, especially in pediatric populations.

Asthma is a common chronic condition in children [22] and the prevalence of poorly controlled asthma status varies between 32 and 64 % in asthmatic children [23–25]. Poor asthma control status is a major factor associated with impairments in different domains of HRQOL [22, 26–30]. The purpose of this study was to identify response shift associated with the change of health states in asthmatic children using the PAQLQ, an asthma-specific HRQOL instrument. Oort's modified SEM measurement model was applied to assess item-level response shift in asthma-specific HRQOL associated with health states measured by the change in asthma control status and global rating of change (GRC) in breathing problems. The impact of response shift was investigated by the change of HRQOL scores with and without adjusting for the response shift effects over time. We hypothesized that response shift in asthma-specific HRQOL can be detected at the item level due to its association with the change in asthma control status and breathing problems. However, the impact of response shift on the change of HRQOL domain scores will be small since acute asthma attack or an acute flare episode may be less significant life events.

Methods

Source of data

PROMIS Pediatric Asthma Study—The Patient-Reported Outcomes Measurement Information System® (PROMIS®) Pediatrics Asthma Study is a NIH-funded project which was designed to validate PROMIS Pediatric Short Forms and a legacy measure, the PAQLQ.

Enrollment criteria—Potential participants were identified from the Florida Medicaid and State Children's Health Insurance Program (SCHIP); 238 dyads of asthmatic children and their parents agreed to participate in this study. The enrollment criteria included children between 8 and 17.9 years old and   18 years for parents, having continuous enrollment (6 months) in Medicaid and SCHIP, having a diagnosis of asthma (ICD-9-CM: 493.1, 493.2, or other 493.x), experiencing at least two asthma-related health care visits during the past year, and having access to the internet and telephone services in the past 6 weeks. After children and parents enrolled into this study, a research package was sent to parents for introducing the study purpose and procedures.

Data collection

A dynamic patient-centered approach was used to collect longitudinal HRQOL data (Fig. 2), and this approach assumes that individuals' HRQOL will be changed in different time frames per the change of underlying health status (i.e., asthma control). Asthma control status, peak flow values, nighttime sleep quality and quantity, and school functioning were reported weekly (26 weeks in total across 2 years) by parents through a research website: weeks 1–13 in the first year and weeks 14–26 in the second year. Pediatric HRQOL data were collected through telephone interviews with children at the first year baseline (T1), the first year follow-up (T2), the second year baseline (T3), and the second year follow-up (T4). The research team evaluated the change of asthma control status by comparing asthma control status reported in week 1 to a particular week between weeks 2–13 of the first year, and asthma control status reported in week 14 to a particular week between weeks 15–26 of the second year. If a change in asthma control status were identified, research coordinators scheduled a telephone interview with children to collect HRQOL data (T2 and T4). If asthma control status remained the same during the 13-week window, a telephone interview was scheduled at the end of the observational period to assess a child's HRQOL. In this study, only data collected from the first year (T1 and T2) were used for investigating response shift.

Measure

HRQOL—The PAQLQ was developed to evaluate asthma-specific HRQOL for children and adolescents between 8 and 17.9 years old. The questionnaire comprises 23 items covering three domains: symptoms (10 items), activity limitation (5 items), and emotional function (8 items). A seven-point response category for each item is utilized (from 7 = "not bothered at all" to 1 = "extremely bothered"). The specific domain scores are calculated by summing the corresponding item scores and dividing by the number of items [31, 32].

Explanatory variables

Asthma control and communication instrument (ACCI): Asthma control status was measured using the asthma control and communication instrument (ACCI), which is a well-validated instrument to measure asthma control status [33]. This instrument was developed on the basis of the 2007 National Asthma Education Prevention Program (NAEPP) Expert Panel Report-3 (EPR-3) [34]. On the ACCI, 11 items assess 5 domains of asthma status including 5 items for asthma control; 3 items for short-term asthma-related health care; 1 item for direction of asthma symptoms; 1 item for adherence to daily asthma medication; 1 item for asthma concern; and 1 open-ended question for measuring patient and physician communication. Per the scoring guidelines, a child's asthma control status was classified as well-controlled or poorly controlled. The ACCI has demonstrated satisfactory psychometric properties including concurrent validity and discriminant and known-group validity [35].

Global rating of change (GRC): GRC due to breathing problems was measured during the follow-up (T2) telephone interview by asking each child "Are your breathing problems better, worse, or about the same as the last time we did this survey?" GRC due to breathing problems was classified as better/about the same or worse.

Confounding variables—Several important covariates collected from the T1 assessment which can potentially influence a child's HRQOL were included in the analyses, including the child's age (a continuous variable), gender (male or female), race/ethnicity (white or non-white), and the number of comorbid conditions (a continuous variable ranging from 0 to 6).

Statistical analysis

A two-step procedure was conducted to first assess response shift for the PAQLQ, followed by measurement bias and response shift in measurement. This sequence of testing emphasizes the importance of identifying potential instances of response shift and investigating whether other measurement bias issues related to explanatory and confounding variables affect the results of response shift [9, 17].

Step 1: establishing an appropriate measurement model

Step 1 was to establish the measurement model for the PAQLQ (Fig. 1). This is an important step because lack of fit of the measurement model to the data can lead to erroneous identification of response shift, measurement bias, and response shift in measurement. A pre-specified construct of the PAQLQ reported in the previous publication [35] was used as a framework to identify the appropriate factor structure for this study. In this step, the factor loadings and intercepts were not constrained to be equal across the two time points. In Step 1a, explanatory variables and four confounding variables (child's age, race, gender, and comorbid conditions) were further added to Step 1.

A variety of fit indices were adopted to assess the appropriateness of the measurement model, including the goodness-of-fit index Chi-square (a nonsignificant Chi-square indicates good model fit) and root-mean-square error of approximation (RMSEA: values below 0.08 indicate a satisfactory model fit and values below 0.05 indicate a close fit) [36].

Step 2: detecting different types of response shift

In Step 2, explanatory variables (change in asthma control status and GRC in breathing problems) with direct effects on the latent factors (i.e., domain scores) were included in the model. The analyses also adjusted for the influence of four confounding variables. All confounding variables were associated with explanatory variables and the latent factors; they were not assumed to directly affect the observed variables (i.e., items) (Fig. 1). Response shift was tested when comparing the model in Step 1a (parameters freely estimated with the inclusion of explanatory and confounding variables) to the model in Step 2 (parameters fully constrained with the inclusion of explanatory and confounding variables) using the Chi-square tests. If a statistically significant difference between Step 1a and Step 2 were found, the subsequent analyses was to identify a specific type of response shift (reconceptualization, reprioritization, recalibration) by testing the difference between the model with a relaxation on some constrained parameters and the model with full constrained parameters (Step 2a).

To identify specific types of response shift, the constraints on parameters of item factor loadings and item intercepts were sequentially relaxed (Step 2a) [18, 37]. For factor loadings and intercepts, one parameter of an individual item was relaxed at a time, and all other parameters were constrained over time. First, equality constraints were released on the factor loading of an individual item while imposing equality constraints on factor loadings of the remaining items. After inspecting each factor loading parameter, a similar process was conducted by releasing equality constraints on the intercepts of an individual item while imposing equality constraints on the remaining intercepts and factor loadings when response shift was not identified.

Reconceptualization response shift is indicated if a change in the matrix pattern containing all factor loadings at T1 differs from the matrix pattern of factor loadings at T2. *Reprioritization* response shift has occurred if the factor loadings of individual items in a specific domain changed over time. *Recalibration* response shift is indicated if the intercept of individual items in a specific domain changes over time. Recalibration response shift implies that subjects may adjust their perception to all response options in the same direction and to the same extent [8, 10, 18, 19]. Sequential analyses to identify different types of response shift were guided by the changes in the modification index values and Chi-square difference test (Chi-square difference of 3.84 with df(1); $p < 0.05$) [10, 17].

Step 3: detecting measurement bias and response shift in measurement

Subsequent to the identification of different types of item-level response shift, measurement bias and response shift in measurement were investigated in Step 3 by adjusting for the influence of explanatory and confounding variables on individual items. Measurement bias is operationalized as a significant association between confounding variables and the response to individual HRQOL items at T1 and T2, respectively, given the same underlying HRQOL. Response shift in measurement is operationalized as the inequality in the magnitude of measurement bias across T1 and T2. In the modeling process, a total of 414 modification indices were calculated (92 (46×2) direct effects of 2 explanatory variables constrained at zero, 184 (46×4) direct effects of 4 confounding variables constrained at

zero, 138 (46 × 3) factor loadings constrained over time, and 46 intercepts constrained over time). Due to the large number of tests, a Bonferroni-adjusted F value [38, 39] of 15.08 (associated with a probability of 0.05/414) was used to control for Type I error.

Two criteria were applied to identify the instances of measurement bias and response shift in measurement, where specific item parameters related to a modification index >15.08 were freely estimated by adjusting for the influence of explanatory and confounding variables, and these parameters only remained freed if the overall model fit indicated by a similar change in Chi-square value. The items with the highest modification indices and difference in Chi-square value >15.08 were the first to be freely estimated followed by the items with the second highest modification indices. This process was continued until all modification index values were <15.08. When the associations of explanatory and confounding variables with specific PAQLQ items were not equal across the two time points, response shift in measurement was identified.

Parameters estimated from Step 3 were used to calculate the effect size of the true change and the response shift. The absolute difference in the estimates between the model that accounted for response shift (Step 3) and the model that did not account for response shift (Step 2) represent the response shift effects. Cohen's effect size d with the values <0.2, 0.2–0.49, 0.5–0.79, and   0.8 were considered to be negligible, small, medium, and large, respectively [21].

LISREL 8.8 [40] was used to test the SEM, and SAS 9.1 software [41] was used for the remaining analyses. Based on the RMSEA values of 0.05 and 0.08, the present study had almost 100 % statistical power to reject the hypothesis that the model does not fit data [38]. Because the scores on the majority of the PAQLQ items were non-normally distributed per Kolmogorov–Smirnov and Shapiro–Wilk tests ($p < 0.001$) (statistics and graphs are available upon request), a robust maximum likelihood (RML) estimation was used in SEM analyses. The attrition rate from T1 to T2 was 6.7 % (16/238) and the incomplete answer to the survey was approximately 0 %. Given the acceptable missing data, we decided to not adjust for the missingness in the statistical analyses.

Results

Description of the sample

Table 1 shows that nearly 60 % ($n = 142$) of the children were male and 38 % ($n = 91$) were Caucasian. The mean age of children was 12.25 years (SD = 2.58). Table 2 shows the mean and SDs of the PAQLQ items at T1 and T2. Paired t tests indicated statistically significant improvement in 12 out of 23 items from T1 to T2 ($p < 0.05$).

Identification of response shift, measurement bias, and response shift in measurement

Step 1: establishing an appropriate measurement model—Parameters for factor loadings and intercepts of individual items were freely estimated between T1 and T2. Model fit statistics indicated satisfactory results with RMSEA = 0.050 (Step 1, Table 3) that allowed for testing different types of response shift, measurement bias, and response shift in measurement in Step 2 and Step 3.

Step 2: different types of response shift—To identify different types of response shift corresponding to the change of asthma-related health states, the factor loadings and intercepts of all items were constrained to be equal between T1 and T2 (Step 2). Response shift was tested when comparing the model in Step 1a (parameters freely estimated with the inclusion of explanatory and confounding variables) to the model in Step 2 (parameters fully constrained with the inclusion of explanatory and confounding variables) using the Chi-square tests. There was a statistically significant difference between Step 1a and Step 2; therefore, subsequent analyses were conducted to identify a specific type of response shift (i.e., reconceptualization, reprioritization, or recalibration). However, we found no instances of types of item-level response shift (Step 2a).

Step 3: measurement bias and response shift in measurement—Following the investigation of different types of response shift at the item level, Step 3 investigated the influence of explanatory and confounding variables on the PAQLQ items by testing measurement bias and response shift in measurement. Modification index values >15.08 at T1 and/ or T2 indicated that the model fit could be further improved by accounting for measurement bias and/or response shift in measurement. The parameters for the item with highest value of modification index >15.08 were freely estimated followed by the item with the second highest modification index >15.08. These steps were continued until all modification index values were <15.08.

The relationship between GRC due to breathing problems and item #21 was not fully determined by their relationships with the latent trait of emotional domain. The modification index value was >15.08, and a direct relationship between GRC due to breathing problems and item #21 was included (estimated at −0.267). The violation of measurement invariance was consistent across T1 and T2, which indicated that children and adolescents with better/ about the same GRC due to breathing problems reported lower scores on this item than those with deteriorated GRC due to breathing problems, conditioning on the same latent trait of emotional domain. After freely estimating the parameter, the overall model fit indicated by Chi-square difference value was >15.08 (Step 3a, Table 3). Next, the relationship between GRC due to breathing problems and item #14 was not fully determined by their relationships with the latent trait of symptom domain. The violation of measurement invariance was consistent across T1 and T2 and a direct relationship between GRC due to breathing problems and item #14 was included (estimated at 0.336) indicating that children and adolescents with better/about the same GRC due to breathing problems had higher scores on this item in the symptom domain than those with deteriorated GRC due to breathing problems. After freely estimating the parameter, the overall model fit indicated by Chi-square difference value was >15.08 (Step 3b, Table 3). We found a positive effect of GRC due to breathing problems on item #14 (0.336 at T1). Neither measurement bias nor response shift in measurement was found to be associated with another explanatory variable (change in asthma control) and four confounding variables.

After testing the influence of explanatory and confounding variables on items, the final model showed improvement and close fit, χ^2 (1231) = 1916.925 and RMSEA = 0.049 (90 % CI 0.044–0.053).

Impact of response shift at the domain level

The impact of response shift on domain scores led to a negligible increase in mean latent scores of the symptom (ES = 0.017) and emotional function (ES = 0.019) domains, whereas a negligible decline in mean latent scores of the activity limitation (ES = 0.010) domain. The ES was estimated by testing the change of domain-level scores with and without accounting for response shift and measurement bias.

Discussion

Using modified Oort's SEM approach [9], we found no instances of item-level response shift in asthmatic children based on the PAQLQ. We also tested the association of specific asthma-related health states (i.e., change in asthma control and GRC in breathing problems) with the PAQLQ items. In support to our hypothesis, GRC due to breathing problems was found to influence PAQLQ items, after accounting for measurement bias and confounding variables. Two instances of measurement bias were identified where there were relationships between GRC due to breathing problems on one item in the symptom domain and another item in the emotional domain. However, the impact of measurement bias is small and will not bias the change of domain scores over time.

Past pediatric studies have frequently used a design approach such as then-test to detect response shift effects [42–44]. Researchers have found divergent results across the then-test and the SEM when evaluating response shift [21], whereas others have found similar findings across the two approaches [45]. The discrepancies can be attributed to the level of analysis, where the SEM approach and the then-test approaches identify response shift at the group and at the individual level, respectively [21]. Ours is the first study to use the modified method of Oort et al. [9] to investigate the presence of item-level response shift in a pediatric population. Response shift at the domain or group level will only be detected when a substantial number of participants are affected [18–21]. The use of domain-level SEM may also mask the item-level response shift, especially recalibration response shift, because domain-level approach tends to neglect information at item level (e.g., item intercepts or thresholds).

We found that asthmatic children with better/about the same GRC due to breathing problems reported lower scores for item #21 of the emotional domain at T2 compared to those with deteriorated GRC due to breathing problems. In addition, asthmatic children with better/about the same GRC due to breathing problems reported better scores for item #14 of the symptom domain at T1 compared to those with deteriorated GRC due to breathing problems. Researchers note that response shift may be present in attitudes or emotional domains rather than in the symptom domains (e.g., fatigue or nausea) [46]. It has also been suggested that children are likely to undergo changes in their life perception compared with adults when the domain of interest appears to be changing over time [46]. Consistent with our hypothesis, our findings suggest that the change in GRC in breathing problems led to different ratings of two PAQLQ items given the same underlying construct. Several studies have found significant effects of response shift in pediatric populations with cancer, diabetes, and otitis media [42–44]. The possible interpretations are, first, the type of illness or health states, or the duration and severity of the disease experiences matter to the presence

of response shift [1, 46]. In this context, response shift may be less likely to occur due to an acute asthma attack or an acute flare episode compared with significant life-threatening events such as cancer [6, 42]. Second, our study samples were not newly diagnosed asthma patients and were likely to have adapted to the disease progression; any response shift, if any, would have already occurred, which also may explain no evidence of response shift effects.

The Oort SEM methodology is useful for testing influence of catalysts on HRQOL by accounting for measurement bias and confounding variables [9, 10, 17]. Previous studies have not addressed the issues of response shift, measurement bias, and response shift in measurement together in a higher-order construct of HRQOL composed of domains and related items [10, 17]. Application of statistical methods using item response theory or Oort's proposed SEM approach for discrete data provides alternative methods for detecting response shift at the item level [47, 48]. In our study, it was feasible to evaluate the impact of multiple catalysts (e.g., change in asthma control status and global rating of change in breathing problems) on response shift in asthmatic children. Response shift was evaluated anytime in a three-month window whenever asthma control status changed and the frequency of health state changes over time may have influenced the occurrence of response shift. Future work should include more than two time points to enable assessment of multiple changes in health states and evaluate its impact on the identification of potential response shift.

There are several limitations to consider when interpreting our results. First, the generalizability of the findings is limited due to the use of participants who were enrolled in Medicaid/SCHIP programs. Second, it is plausible that certain unmeasured catalysts, for instance, the change in lung functioning measured by forced expiratory volume in 1 s and treatment strategies (e.g., inhaled corticosteroids) may affect response shift. Future work is needed to investigate the role of other potential catalysts to cause item-level and domain-level response shift among asthmatic children. Third, measurement model parameters for each domain were specified with item intercepts rather than item thresholds. This approach was applied to accommodate the small sample size in our study. A larger sample size is needed when using item thresholds for testing response shift.

Conclusion

No item-level response shift appears in asthmatic children based on the PAQLQ. However, two items of the PAQLQ emerge measurement bias related to GRC due to breathing problems. The impact of measurement bias is small and will not bias the change of domain scores over time.

Acknowledgments

References

1. Sprangers MA, Schwartz CE. Integrating response shift into health-related quality of life research: A theoretical model. Social Science and Medicine. 1999; 48:1507–1515. [PubMed: 10400253]

2. Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality of life research. Social Science and Medicine. 1999; 48:1531–1548. [PubMed: 10400255]

3. Hamidou Z, Dabakuyo TS, Bonnetain F. Impact of response shift on longitudinal quality-of-life assessment in cancer clinical trials. Expert Review of Pharmacoeconomics & Outcomes Research. 2011; 11(5):549–559. [PubMed: 21958100]

4. Gillison F, Skevington S, Standage M. Exploring response shift in the quality of life of healthy adolescents over 1 year. Quality of Life Research. 2008; 17:997–1008. [PubMed: 18633730]

5. De Civita M, Regier D, Alamgir AH, Anis AH, FitzGerald MJ, et al. Evaluating health-related quality of life studies pediatric populations. Pharmacoeconomics. 2005; 23(7):659–685. [PubMed: 15987225]

6. Schwartz CE, Feinberg RG, Jilinskaia E, Applegate JC. An evaluation of psychosocial intervention for survivors of childhood cancer: Paradoxical effects of response shift over time. Pysho-Oncology. 1999; 8:344–354.

7. Rapkin BD, Schwartz CE. Toward a theoretical of quality-of-life appraisal: Implications of findings from studies of response shift. Health and Quality of Life Outcomes. 2004; 2:14. [PubMed: 15023229]

8. Ahmed S, Sawatzky R, Levesque J-F, Ehrmann-Feldman D, Schwartz CE. Minimal evidence of response shift in the absence of a catalyst. Quality of Life Research. 2014; 23(9):2421–2430. [PubMed: 24899546]

9. Oort FJ, Visser MRM, Sprangers MAG. Formal definitions of measurement bias and explanation bias clarify measurement and conceptual perspectives on response shift. Journal of Clinical Epidemiology. 2009; 62:1126–1137. [PubMed: 19540722]

10. Gandhi PK, Ried LD, Kimberlin C, Kauf TK, Huang I-C. Influence of explanatory and confounding variables on HRQoL after controlling for measurement bias and response shift in measurement. Expert Reviews of Pharmacoeconomics & Outcomes Research. 2013; 13(6):841–851.

11. Schwartz CE, Sprangers MAG, Carey A, Reed G. Exploring response shift in longitudinal data. Psychology and Health. 2004; 19(1):51–69.

12. Schwartz CE, Rapkin BD. Understanding appraisal processes underlying the then test: A mixed methods investigation. Quality of Life Research. 2012; 21(3):381–388. [PubMed: 21969088]

13. Ahmed S, Mayo NE, Corbiere M, Wood-Dauphinee S, Hanley J, Cohen R. Change in quality of life in people with stroke over time: true change or response shift? Quality of Life Research. 2005; 14:611–627. [PubMed: 16022056]

14. Schwartz CE, Sprangers MAG. Guidelines for improving the stringency of response shift research using the then-test. Quality of Life Research. 2010; 19:455–464. [PubMed: 20084465]

15. Schwartz CE, Sprangers MAG, Ahmed S, Bode R, Li Y, Oort FJ, et al. Response shift in patients with multiple sclerosis: An application of three statistical techniques. Quality of Life Research. 2011; 20(10):1561–1572. [PubMed: 22081216]

16. Schwartz CE, Ahmed S, Sawatsky R, Sajobi T, Mayo N, Finkelstein JA, et al. Guidelines for secondary analysis in search of response shift. Quality of Life Research. 2013; 22(22):2663–2673. [PubMed: 23572398]

17. King-Kallimanis BL, Oort FJ, Visser MRM, Sprangers MAG. Structural equation modeling of health-related quality-of-life data illustrates the measurement and conceptual perspectives on response shift. Journal of Clinical Epidemiology. 2009; 62:1157–1164. [PubMed: 19595574]

18. Gandhi PK, Ried LD, Huang IC, Kimberlin CL, Kauf TL. Assessment of response shift using two structural equation modeling techniques. Quality of Life Research. 2012; 22(3):461–471. [PubMed: 22476611]

19. Oort FJ, Visser MRM, Sprangers MAG. An application of structural equation modeling to detect response shifts and true change in quality of life data from cancer patients undergoing invasive surgery. Quality of Life Research. 2005; 14:599–609. [PubMed: 16022055]

20. Visser M, Oort F, Sprangers M. Methods to detect response shift in quality of life data: A convergent validity study. Quality of Life Research. 2005; 14:629–639. [PubMed: 16022057]

21. Ahmed S, Mayo NE, Wood-Dauphinee S, Hanley JA, Cohen SR. The structural equation modeling technique did not show a response shift, contrary to the results of the then test and the individualized approaches. Journal of Clinical Epidemiology. 2005; 58(11):1125–1133. [PubMed: 16223655]

22. Gandhi PK, Kenzik KM, Thompson LA, DeWalt D, Revicki DA, Shenkman EA, et al. Exploring factors influencing asthma control and asthma-specific health-related quality of life among children. Respiratory Research. 2013; 14(1):26. [PubMed: 23432913]

23. Smith LA, Bokhour B, Hohman KH, Miroshnik I, Kleinman KP, Cohn E, et al. Modifiable risk factors for suboptimal control and controller medication underuse among children with asthma. Pediatrics. 2008; 122(4):760–769. [PubMed: 18829799]

24. Lozano P, Finkelstein JA, Hecht J, Shulruff R, Weiss KB. Asthma medication use and disease burden in children in a primary care population. Archives of Pediatrics and Adolescent Medicine. 2003; 157(1):81–88. [PubMed: 12517200]

25. Koster ES, Raaijmakers JA, Vijverberg SJ, Koenderman L, Postma DS, Koppelman GH, et al. Limited agreement between current and long-term asthma control in children: The PACMAN cohort study. Pediatric Allergy and Immunology. 2011; 22(8):776–783. [PubMed: 21749459]

26. Chen H, Gould MK, Blanc PD, Miller DP, Kamath TV, Lee JH, et al. Asthma control, severity, and quality of life: Quantifying the effect of uncontrolled disease. Journal of Allergy and Clinical Immunology. 2007; 120:396–402. [PubMed: 17561244]

27. Zheng, Li; Huang, I-C.; Thompson, L.; Tuli, S.; Huang, S-W.; DeWalt, D., et al. The relationships between asthma control, daytime sleepiness, and quality of life among children with asthma: A path analysis. Sleep Medicine. 2013; 14(7):641–647. [PubMed: 23684939]

28. Dean BB, Calimlim BM, Kindermann SL, Khandker RK, Tinkelman D. The impact of uncontrolled asthma on absenteeism and health-related quality of life. The Journal of Asthma. 2009; 46:861–866. [PubMed: 19905909]

29. Pont LG, Van der Molen T, Denig P, Van der Werf GT, Haaijer-Ruskamp FM. Relationship between guideline treatment and health-related quality of life in asthma. European Respiratory Journal. 2004; 23(5):718–722. [PubMed: 15176686]

30. Guilbert TW, Garris C, Jhingran P, Bonafede M, Tomaszewski KJ, Bonus T, et al. Asthma that is not well-controlled is associated with increased healthcare utilization and decreased quality of life. The Journal of Asthma. 2011; 48(2):126–132. [PubMed: 21128880]

31. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffin LE, Townsend M. Measuring quality of life in children with asthma. Quality of Life Research. 1996; 5:35–46. [PubMed: 8901365]

32. Raat H, Bueving HJ, de Jongste JC, Grol MH, Juniper EF, van der Wouden JC. Responsiveness, longitudinal- and cross-sectional construct validity of the Pediatric Asthma Quality of Life Questionnaire (PAQLQ) in Dutch children with asthma. Quality of Life Research. 2005; 14(1): 265–272. [PubMed: 15789960]

33. Patino CM, Okelo SO, Rand CS, Riekert KA, Krishnan JA, Thompson K, et al. The asthma control and communication instrument: A clinical tool developed for ethnically diverse populations. The Journal of Allergy and Clinical Immunology. 2008; 122(5):936.e6–943.e6. [PubMed: 18848721]

34. National Asthma Education and Prevention Program. Expert panel report 3 (EPR-3): Guidelines for the diagnosis and management of asthma-summary report 2007. Journal of Allergy and Clinical Immunology. 2007; 120(5 Suppl):S94–S138. [PubMed: 17983880]

35. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. European Respiratory Journal. 1997; 10(10):2285–2294. [PubMed: 9387955]

36. Brown, TA. Confirmatory factor analysis for applied research. New York, London: Guilford Press; 2006.

37. Barclay-Goddard R, Lix LM, Tate R, Weinberg L, Mayo NE. Response shift was identified over multiple occasions with a structural equation modeling framework. Journal of Clinical Epidemiology. 2009; 62:1181–1188. [PubMed: 19615858]

38. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychological Methods. 1996; 1(2):130–149.

39. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Statistics in Medicine. 1990; 9:811–818. [PubMed: 2218183]

40. Jöreskog, K.; Sörbom, D. LISREL 8. Lincolnwood, Il: Scientific Software Internatonal Inc; 1993.

41. SAS Institute. SAS users guide, version 9.1. Cary, NC: SAS Institute, Inc; 2004.

42. Brinksma A, Tissing WJ, Sulkers E, Kamps WA, Roodbol PF, Sanderman R. Exploring the response shift phenomenon in childhood patients with cancer and its effect on health-related quality of life. Oncology Nursing Forum. 2014; 41(1):48–56. [PubMed: 24195842]

43. Wagner JA. Response shift and glycemic control in children with diabetes. Health and Quality of Life Outcomes. 2005; 3:38. [PubMed: 15955236]

44. Timmerman AA, Anteunis LJ, Meesters CM. Response-shift bias and parent-reported quality of life in children with otitis media. Archives of Otolaryngology - Head and Neck Surgery. 2003; 129(9):987–991. [PubMed: 12975273]

45. Verdam MGE, Oort FJ, Visser MRM, Sprangers MAG. Response shift detection through then-test and structural equation modeling: Decomposing observed change and testing tacit assumptions. Netherlands Journal of Psychology. 2012; 67(3):58–67.

46. Brossart DF, Clay DL, Willson VL. Methodological and statistical considerations for threats to internal validity in pediatric outcome data: Response shift in self-report outcomes. Journal of Pediatric Psychology. 2002; 27(1):97–107. [PubMed: 11726684]

47. Verdam MGE, Oort FJ, Sprangers MAG. Using structural equation modeling to detect response shifts and true change in discrete variables: An application to the items of the SF-36. Quality of Life Research. 2015

48. Guilleux A, Blanchin M, Vanier A, Guillemin F, Falissard B, Schwartz CE, et al. Response shift algorithm in item response theory (ROSALI) for response shift detection with missing data in longitudinal patient-reported outcome studies. Quality of Life Research. 2015; 24:553–564. [PubMed: 25477228]
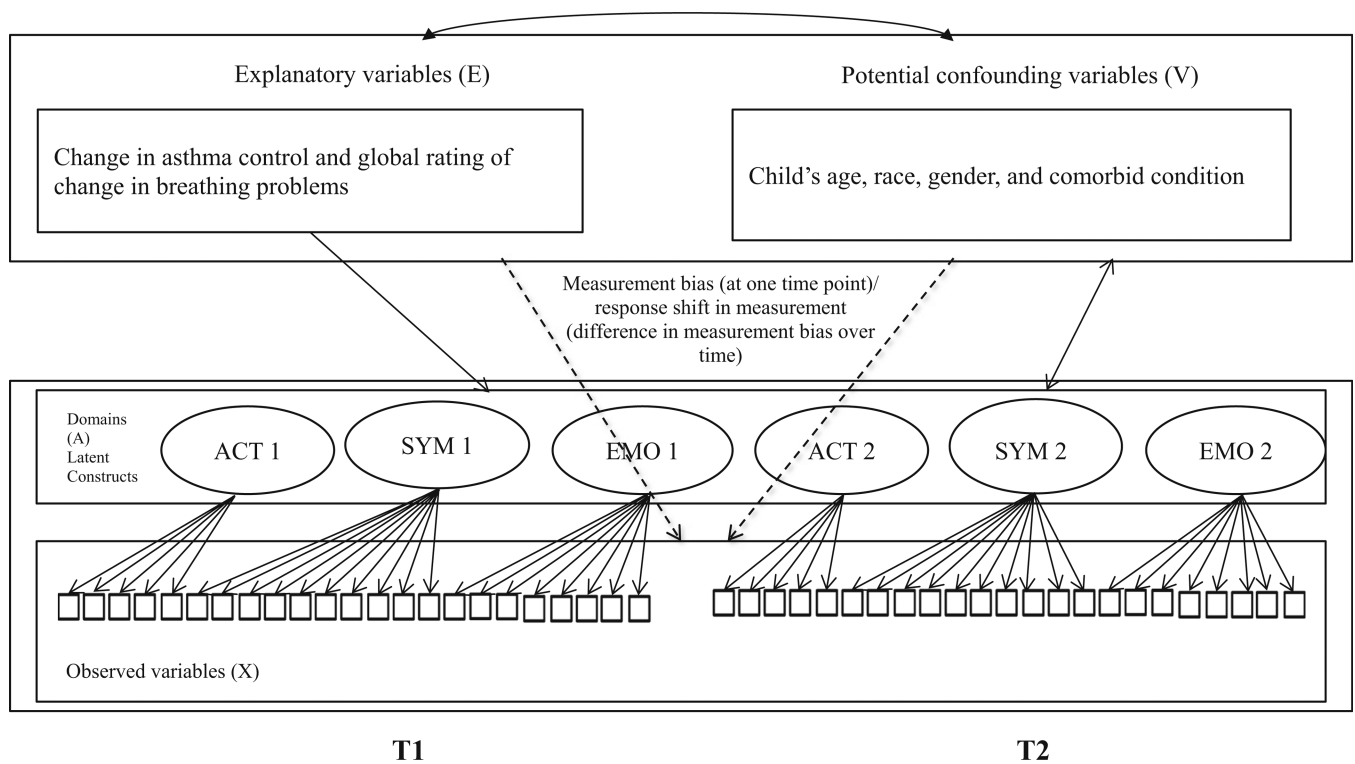
Fig. 1.
Graphical representation of the model to identify response shift, measurement bias, and response shift in measurement. *ACT* activity limitation, *SYM* symptom, *EMO* emotional function. *Note* Although not shown in the figure, all domains *A* are correlated with each other, all explanatory variables *E* are correlated with each other. The *double-headed arrow* represents correlations between all *E* and *V* variables and correlations between all *V* variables and *A*. *Dashed arrows* represent measurement bias/response shift. *Single-headed arrow* represents direct effects of *E* on *A*. *T1* and *T2* indicates the two time points in this study, i.e., baseline and anytime in a 3-month window whenever asthma control status changed
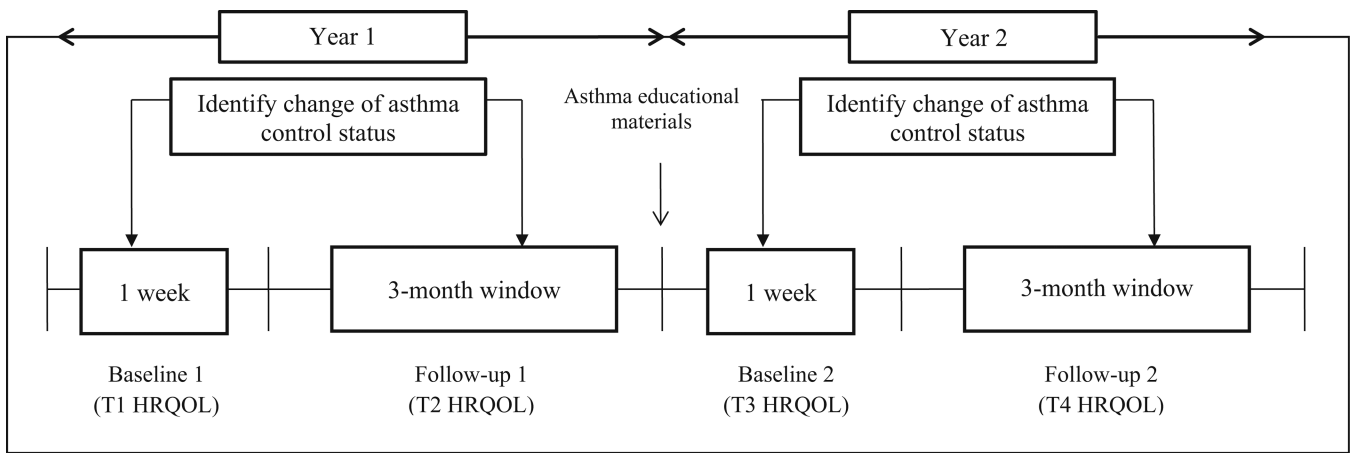
Fig. 2.
Approach to observe change of asthma control and patient-reported outcomes

Table 1

Subject characteristics ($n = 238$)

Characteristics	N (%) or mean (SD)
Child age, years	12.25 (2.58)
Child gender, %	
Male	142 (59.70)
Female	95 (39.90)
Child race/ethnicity, %	
White	91 (38.20)
Black	61 (25.60)
Hispanic	64 (26.9)
Other	
Asthma control change, %	
Deteriorated	58 (24.40)
Same/improved	137 (57.60)
Global rating of change in breathing problems, %	
Deteriorated	30 (12.60)
Same/improved	167 (70.20)

Numbers for each variable may not equal the sample size due to missing data

SD standard deviation

Table 2

Means, standard deviations, and pre-post Cohen's "d" effect sizes for PAQLQ items

Item[a]	Baseline Mean	Baseline SD	Follow-up Mean	Follow-up SD	Cohen's d[b]
Activity domain					
How much have you been bothered by your asthma in physical activities (such as running, swimming, sports, walking uphill/upstairs and bicycling)? (item #1)	5.204	1.668	5.355	1.640	0.091
How much have you been bothered by your asthma in being with animals (such as playing with pets and looking after animals)? (item #2)	5.805	1.722	6.090	1.524	0.176
How much have you been bothered by your asthma in activities with friends and family (such as playing at recess and doing things with your friends and family)? (item #3)	5.724	1.561	6.032	1.377	0.210*
How often did you feel you couldn't keep up with others because of your asthma? (item #19)	5.272	1.890	5.615	1.650	0.194*
Think about all the activities that you did in the past week. How much were you bothered by your asthma doing these activities? (item #22)	5.213	1.785	5.500	1.723	0.163
Symptom domain					
How much did coughing bother you? (item #4)	5.059	1.973	5.284	1.898	0.116
How often did your asthma make you feel tired? (item #6)	5.136	1.812	5.693	1.706	0.317**
How much did asthma attacks bother you? (item #8)	5.860	1.743	6.129	1.600	0.161
How much did wheezing bother you? (item #10)	5.484	1.731	5.777	1.636	0.174
How much did tightness in your chest bother you? (item #12)	5.502	1.734	5.732	1.727	0.133
How much did shortness of breath bother you? (item #14)	5.014	1.878	5.581	1.708	0.316**
How often did your asthma wake you up during the night? (item #16)	5.837	1.654	5.969	1.543	0.083
How often did you feel out of breath? (item #18)	5.000	1.791	5.466	1.712	0.266**
How often did you have trouble sleeping at night, because of your asthma? (item #20)	5.851	1.649	5.980	1.457	0.083
How often did you have difficulty taking a deep breath? (item #23)	5.465	1.724	5.814	1.581	0.211*
Emotional domain					
How often did your asthma make you feel frustrated? (item #5)	5.448	1.709	5.774	1.627	0.195*
How often did you feel worried, concerned, or troubled because of your asthma? (item #7)	5.805	1.588	6.120	1.356	0.214*
How often did your asthma make you feel angry? (item #9)	5.968	1.611	6.219	1.428	0.165
How often did your asthma make you feel irritable? (item #11)	5.828	1.498	6.095	1.435	0.182*
How often did you feel different or left out because of your asthma? (item #13)	5.846	1.715	6.291	1.280	0.297**
How often did you feel frustrated because you couldn't keep up with others? (item #15)	5.367	1.803	5.793	1.648	0.247**

Item[a]	Baseline		Follow-up		Cohen's d[b]
	Mean	SD	Mean	SD	
How often did you feel uncomfortable because of your asthma? (item #17)	5.664	1.705	5.844	1.550	0.111
How often did you feel frightened by an asthma attack? (item #21)	6.036	1.609	6.357	1.233	0.226*

[a] All PAQLQ items end with "during the past week"

[b] Cohen's effect size: <0.2, 0.2–0.49, 0.5–0.79, and 0.8 indicate 'negligible', 'small', 'moderate', and 'large' differences

*
$p < 0.05$;

**
$p < 0.01$ in paired t test

Table 3

Goodness of fit of models in measurement bias and response shift in measurement detection procedure when controlling for asthma health states and confounding variables

Step	Specification	Df	χ^2	RMSEA (90 % CI)
Step 1	All parameters[a] freely estimated (excluding explanatory and confounding variables)	947	1465.781	0.048 (0.043–0.053)
Step 1a	All parameters freely estimated (including explanatory and confounding variables)	1190	1901.142	0.050 (0.046–0.054)
Step 2	All parameters constrained (including explanatory and confounding variables)	1233	1957.146	0.050 (0.046–0.054)
Step 2a	Detect types of response shift			
	No specific types of response shift were identified			
Step 3	Detect measurement bias and/or response shift in measurement			
Step 3a	Global rating of change in breathing problems—item #21 at T2	1232	1941.502	0.049 (0.045–0.053)
Step 3b	Global rating of change in breathing problems—item #14 at T1	1231	1916.925	0.049 (0.044–0.053)

RMSEA, root-mean-square error of approximation; RMSEA 0.05 close fit, 0.05–0.08 reasonable fit, 0.10 poor fit

[a]The parameter estimates include 46 factor loadings, 46 intercepts, 73 residual covariances (that includes between the same item across time, among the same item at each time point, and a few instances of covariances across different items), 21 common factor (co)variances and 6 common factor means resulting in a total of 192 parameter estimates