



# Implications of response shift for micro-, meso-, and macro-level healthcare decision-making using results of patient-reported outcome measures

Richard Sawatzky<sup>1,2,3</sup> · Jae-Yung Kwon<sup>1,4,5</sup> · Ruth Barclay<sup>6</sup> · Cynthia Chauhan<sup>7</sup> · Lori Frank<sup>8</sup> · Wilbert B. van den Hout<sup>9</sup> · Lene Kongsgaard Nielsen<sup>10,11</sup> · Sandra Nolte<sup>12,13</sup> · Mirjam A. G. Sprangers<sup>14</sup> · the Response Shift – in Sync Working Group

Accepted: 11 January 2021 / Published online: 2 March 2021  
© The Author(s) 2021

## Abstract

**Purpose** Results of patient-reported outcome measures (PROMs) are increasingly used to inform healthcare decision-making. Research has shown that response shift can impact PROM results. As part of an international collaboration, our goal is to provide a framework regarding the implications of response shift at the level of patient care (micro), healthcare institute (meso), and healthcare policy (macro).

**Methods** Empirical evidence of response shift that can influence patients' self-reported health and preferences provided the foundation for development of the framework. Measurement validity theory, hermeneutic philosophy, and micro-, meso-, and macro-level healthcare decision-making informed our theoretical analysis.

**Results** At the micro-level, patients' self-reported health needs to be interpreted via dialogue with the clinician to avoid misinterpretation of PROM data due to response shift. It is also important to consider the potential impact of response shift on study results, when these are used to support decisions. At the meso-level, individual-level data should be examined for response shift before aggregating PROM data for decision-making related to quality improvement, performance monitoring, and accreditation. At the macro-level, critical reflection on the conceptualization of health is required to know whether response shift needs to be controlled for when PROM data are used to inform healthcare coverage.

**Conclusion** Given empirical evidence of response shift, there is a critical need for guidelines and knowledge translation to avoid potential misinterpretations of PROM results and consequential biases in decision-making. Our framework with guiding questions provides a structure for developing strategies to address potential impacts of response shift at micro-, meso-, and macro-levels.

**Keywords** Patient-reported outcomes · Response shift · Healthcare decision-making · Patient · Organization · Health policy

## Introduction/background

Patient-reported outcomes (PROs), and by extension patient-reported outcome measures (PROMs), are increasingly used to inform healthcare decision-making. The decisions that PROs inform can be considered at the level of patient care (micro), the healthcare organization (meso), and health policy (macro). There is ample evidence that response shift can affect PROs [1, 2]. Response shift is defined as a change in

the meaning of one's self-evaluation as a result of changes in internal standards (recalibration), values (reprioritization), and/or conceptualization of the target construct (reconceptualization) [3], which may result in measurements at two or more time points not being comparable. Evolving perspectives of measurement validity place increasing emphasis on the inferences, judgements, and decisions based on measurement scores [4]. It is therefore important to consider the potential implications of response shift when measures of change in PROs over time are used to inform decisions. For example, when we assess patients' perceived health status at two time points, we can take their responses to that item at each time point at face value, even if response shift has occurred in the interim. However, if we want to measure

✉ Richard Sawatzky  
rick.sawatzky@twu.ca

Extended author information available on the last page of the article

*change* in perceived health status over time, inferences we draw from the observed change in that item may be invalid, leading to ill-informed decisions. The possibility of response shift must therefore be considered when measuring changes in PROM scores over time and when comparing differences in PROM scores between people (e.g., when comparing people who experienced response shifts with those who did not).

To date, the implications of response shift for healthcare decision-making have rarely been investigated [5]. Our goal is to provide a framework regarding the impacts of response shift at micro-, meso- and macro-levels of healthcare decision-making. In so doing, we intend to raise awareness and provide guidance for PRO researchers who play an important role in informing healthcare providers, healthcare leaders, health policymakers, and regulators regarding the implications of response shift. This work is part of an international, interdisciplinary collaboration (see Appendix for the participating members) to synthesize the research on response shift to date [6–8].

## Analytic frame

Our analysis is informed by perspectives of measurement validity, hermeneutic philosophy, and healthcare decision-making. From a measurement validity point of view, the impact of response shift on healthcare decision-making ultimately pertains to whether the inferences we make are valid, and subsequent actions and decisions made on PROs are well founded. This view of measurement validity is based on the foundational work by Messick [4, 9] and, subsequently, Zumbo [10–13], which has received increasing attention in PRO measurement [14–16]. Response shift threatens the measurement assumption that the process and frame of reference by which a person responds to PROM items are consistent over time [17]. In other words, response shift occurs when observed scores (i.e., responses to PROM

items) do not reflect the construct of interest in the same way at different points in time [18, 19]. As a consequence, the meaning of the PROM score will not be consistent over time. Our inferences, actions, and decisions made on PROs must therefore take into account the possibility that PROM scores may be influenced by response shift.

Building on the theoretical premises of measurement validity, we take a hermeneutic philosophical lens to gain further insight into how the meaning of PROM scores may change over time [20] and the consequent implications for micro-, meso-, and macro-level healthcare decision-making. At a foundational level, the use of PROs involves an interpretive process of understanding the meaning of individuals' responses to PROM items. We specifically drew upon Gadamerian hermeneutics [21, 22] because it draws attention to the dialectical processes (a.k.a. hermeneutic circle) that provide insights into how different stakeholders interpret and use PROM results. Dialectical processes refer to the interplay between patients' prior ideas and experiences and the assumptions held by researchers and clinicians about health. This interplay constitutes a dialectical process of navigating differences in how patients interpret and respond to questions about their health, and how researchers and clinicians interpret the results. We accordingly structure the micro-, meso-, and macro-level implications of response shift based on the following three tenets from Gadamerian hermeneutics: (1) the use of PROMs involves the interpretation of contextual elements; (2) interpretation of PROM results is an ongoing dialectical interaction; and (3) the integration of PROM data in decision-making requires openness and reflexivity. We formulated corresponding key guiding questions to structure our analysis in an effort to critically analyze the implications of using PROMs for healthcare decision-making (see Table 1).

Additionally, our analysis is informed by a recognition that there are many different types of healthcare decisions

**Table 1** Tenets and guiding questions for considering response shift implications based on hermeneutics

Tenet	Guiding questions
Tenet 1. The use of PROMs involves the interpretation of contextual elements	What is the purpose of using PROMs in this context?  Is the context specific to an intervention, program or healthcare model?
Tenet 2. Interpretation of PROM data is an ongoing dialectical interaction	How are assumptions or beliefs influencing the interpretation and use of PROM results?  Is it assumed that an intervention would improve PROM scores?
Tenet 3. The integration of PROM data in decision-making involves ideally openness and reflexivity	Can response shift affect the interpretation of PROM results?  If response shift is identified, how does this affect the inferences about the meaning of PROM scores?

that need to be made for a range of different purposes. We therefore adopted a healthcare system’s orientation towards micro-, meso-, and macro-levels of decision-making that draws attention to the different priorities and perspectives of stakeholders [23, 24]. The micro-level focuses on the use of PROM data in clinical practice for the purposes of informing clinical decisions about a patient’s care. This is particularly pertinent in shared decision-making where clinicians and patients share information about PROs and the best available evidence to make decisions about treatments, goals of care, and continuation or addition of interventions or supportive services, in relation to patients’ preferences [25]. At the meso-level, PROM data are used by healthcare managers and leaders for quality improvement, performance monitoring, and accreditation of different healthcare services and organizations. The macro-level focuses on the overall healthcare system where PROM data are used by government leaders and decision-makers to inform health policy regarding healthcare coverage, including the provision and reimbursement of healthcare services. The levels are interrelated, where the macro-level informs and is informed by impacts at the meso- and micro-levels, and the meso-level informs and is informed by impacts at the micro-level. Figure 1 provides an overview of the different types of decisions made at each level and the concurrent interdependent relationships between these levels. By organizing our examination into micro-, meso-, and macro-levels of decision-making, we also provide a framework for developing recommendations for mitigating potential unintended consequences of response shift at each level.

### When does response shift occur?

Response shifts can result from any event or experience that causes a person to think differently about internal standards, values, or conceptualization of their health. Response shifts can occur as a result of sudden *acute conditions*, such as cerebral hemorrhage, head injury, and spinal cord injury [26–28]. Response shifts can also result from experiences in living with potentially progressive *chronic conditions*, including multiple sclerosis, cancer, diabetes, HIV/AIDS, chronic obstructive pulmonary disease, and sleep apnea [1, 29–33]. Chronically ill patients with stable health are less likely to experience response shifts [34]. Response shifts have also been demonstrated in patients whose health improved, for example in patients with pain undergoing total knee arthroplasty, hearing impaired adults receiving hearing aid fitting, and edentulous patients receiving denture treatment [35–37]. *Healthcare interventions* could also lead to response shifts [38]. Some interventions are designed to promote reframing how one thinks about one’s health and are particularly likely to produce response shifts, in fact, as a desired outcome [39, 40]. For other interventions, response shifts may be an unintended consequence (e.g., invasive surgery or toxic cancer treatment [41]). Finally, response shifts could occur as a result of *life events* that are not defined by specific diseases or chronic conditions (e.g., experiences associated with human development and aging) [42, 43]. Another example includes informal caregivers who may experience response shifts as a result of the challenges and experiences associated with caregiving [44]. In all of the above

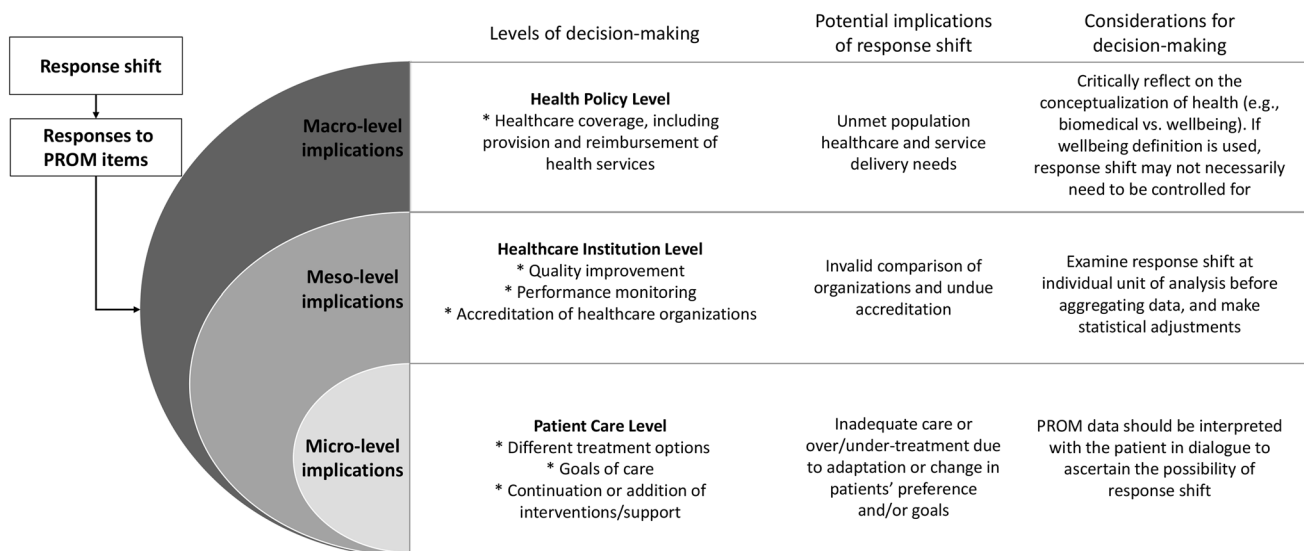


Fig. 1 Potential response shift implications for use of PROM data at micro-, meso-, and macro-levels of healthcare decision-making

situations, it is important to consider that some people may be more susceptible to response shifts than others.

While various methods exist for evaluating the occurrence and magnitude of response, the “then-test” has been used most frequently (see Sébille et al. for explanation of methods and limitations [8]). The “then-test” entails a retrospective assessment administered at follow-up asking for a re-evaluation of one’s functioning at the time of the first assessment (pre-test) [17]. However, although this test explicitly operationalizes the notion of response shift, it is also sensitive to recall bias. A meta-analysis by Schwartz et al. (conducted in 2006) of 19 studies using the “then-test” indicated that the magnitudes of response shift effects are generally small and vary by the type of health outcome being measured (e.g., symptoms versus physical function) [45]. In fact, Rapkin and Schwartz have linked response shift to a particular type of self-reported outcome, i.e., evaluation-based outcomes, as opposed to performance-based or perception-based outcomes [46]. Nonetheless, it should be noted that even a small response shift effect may change conclusions [41].

It is also important to consider unanticipated response shift effects that are often overlooked, which can be exemplified for four basic research designs (Table 2). In cross-sectional observational designs, response shifts may have taken place prior to study entry. For example, differential exposure to factors leading to response shifts (e.g., length of time living with a chronic condition) could result in differences between respondents in their frames of reference for interpreting the PROM items and response scales. In longitudinal interventional studies with or without cost-effectiveness analyses, treatments that are compared may induce response shifts in different outcomes, or with different magnitudes, and/or directions of response shift. In these instances, erroneous conclusions may be drawn regarding the differences in PROM results between groups, changes in PROM results over time, and the preference for and cost-effectiveness of one treatment over another.

## Implications of response shift at different levels of healthcare decision-making

### Micro-level decision-making

At the micro-level, PROM data are used to inform decisions about different treatment options, goals of care, and the need for continuation or addition of interventions or supportive services (see Fig. 1 and Table 3). These types of decisions may be informed by PROMs data from individual patients [47, 48] as well as aggregate PROMs data about groups or populations [49].

When a choice needs to be made between *different treatment options*, the patient’s health status, preferences, and values need to be taken into account, in a process of shared decision-making, if the patient so desires. It is possible that the relative importance that patients place on different PRO domains changes over time (i.e., reprioritization response shift). For example, response shift analyses of the Quality of Life in Epilepsy Inventory-31 (QOLIE-31) comparing surgical and medical treatments for epilepsy suggest that patients with epilepsy who receive surgical treatment place relatively less importance on “seizure worry” and more on the social function domain over the course of time in their illness trajectory in relation to the two treatments [50] (see Table 3). These results suggest that social function is an important consideration when deciding between surgical and medical treatments. From a hermeneutic point of view, consideration of “dialectical interaction” can help to further explore the possibility of a shift in relative importance by reaching a shared understanding between the patient and the clinician in the context of shared decision-making. For example, if this response shift is ignored, clinicians might erroneously focus on seizure management by prescribing drugs that adversely affect patients’ social function, which has now been identified as being more important. Aggregated PROM data may inform decisions between different treatment options based on published results of intervention studies or patient registries that used PROMs with comparable patients undergoing the target intervention. As PROM results have been shown to be predictive of other outcomes such as re-operations [51], re-occurrence of index events [52], and mortality [53, 54], their use for individual decision-making is likely to increase. A particular challenge regarding the use of such aggregated PROM data is that response shift may have occurred but not taken into account (e.g., by statistically adjusting for the occurrence of response shift), thereby potentially misinforming individual decision-making for patients who are likely to undergo response shift.

When decisions need to be made about the *goals of care*, it is important to realize that such goals may change over time, depending on the course of the disease trajectory and the preferences of the patient. At different stages, patients may find different life domains relevant (reconceptualization) and prioritize those differently (reprioritization). For example, Ahmed et al. [29] found that the relative importance of different health domains in the Patient-Generated Index (PGI) changed at different stages of recovery following a stroke. When such response shifts are ignored, inappropriate goal setting that is not aligned with the patient’s current priorities may occur (Table 3). To prevent inappropriate goal setting, hermeneutics can be used to ask open and reflexive questions with the patient such as: How has the patient’s thinking changed, and what would the patient do differently now? In addition, if a condition cannot be cured

**Table 2** Implications of response shift for study results

Research design	Occurrence of response shift	Effect on results	Example
<p><b>Cross-sectional observational:</b> comparing groups at the same time point</p>	<p>Prior response shifts could result in different standards, values, or interpretations (meaning) of PROs between groups. For example, events in the past may have induced changed perspectives (earlier response shifts), resulting in either higher or lower levels of PROM results than would be expected based on health status.</p>	<p>Response shifts that differ between groups may affect conclusions about differences in PROM results, e.g., insignificant (mitigated) differences where differences exist or significant (amplified) differences where none exist.</p>	<p>Breast cancer patients were found to report comparable or superior PROM results (including anxiety and depression) in comparison with healthy women [90].</p>
	<p><b>Longitudinal observational:</b> assessing one group over time</p>	<p>Health <i>deterioration</i> leading to lowered standards, changed priorities, and conceptualizations. Health <i>improvement</i> leading to raised standards, changed priorities, and conceptualizations.</p>	<p>This type of response shift may result in higher levels of QoL than would be expected based on health status. Such response shift may result in lower levels of PROM results than would be expected based on health status.</p>
<p><b>Longitudinal interventional</b></p>	<p>Treatment may induce response shifts when patients need to adapt to the changing health state. For example, treatments that are compared may induce response shifts in different outcomes, or magnitude and/ or direction in the same outcomes.</p>	<p>The comparison of outcomes that are differentially affected between groups by response shifts may lead to erroneous conclusions, e.g., insignificant (mitigated) differences where differences exist or significant (amplified) differences where none exist, also when randomization has taken place.</p>	<p>Patients with myelofibrosis treated with ruxolitinib reported improvement in functional domains and global QoL whereas patients receiving placebo-reported deterioration [92]. Longer follow-up indicated a small decline in QoL in the treatment condition; the placebo condition was not followed up. Whereas not measured, authors hypothesized that the decline might be related to response shift [93]. The difference between treatment and placebo conditions might have been amplified by the occurrence of response shift, an effect that may have decreased over time due to adaptation.</p>
	<p><b>Longitudinal interventional, including cost-effective analysis</b></p>	<p>Same as longitudinal interventional.</p>	<p>Differences in response shifts between two treatments may lead to incorrect cost-effective analyses and an incorrect preference for one treatment over another, also when randomization has taken place.</p>



**Table 3** Examples and potential implications of response shift at the individual patient (micro)-level based on Gadamerian hermeneutics

Type of decision	Response shift example	Contextual elements: purposes for using PROMs	Dialectical processes: assumptions influencing interpretation	Openness and reflexivity: impact of response shift on inferences about meaning of the scores	Potential consequences: How response shift affects decision-making
Choice between different treatment options	Patients with temporal lobe epilepsy participating in a surgical randomized trial were found to reprioritize the importance of domains constituting QoL [50].	The Quality of Life in Epilepsy Inventory-31 (QOLIE-31), a 31-item questionnaire that comprised of 7 domains (overall quality of life, social function, fatigue, emotional well-being, seizure worry, memory, and medication effect) was used to assess surgical and medical treatment.	The traditional assumption in evaluating treatment effectiveness in individuals with epilepsy is that freedom from seizures and number of antiepileptic drugs prescribed continues to be the highest priority before and after surgery.	Over time, patients with epilepsy placed less importance on seizure worry and more on social function items.	If response shift (reprioritization) is ignored, patients may be prescribed interventions (e.g., drugs) that could have adverse effect on patient's social function, which may be as, or even more, important than seizure management.
Decision about goals of care	Patients at 6- and 24-weeks post-stroke who were discharged home from an acute hospital admission were found to experience reconceptualization and reprioritization response shifts [29].	The Patient-Generated Index (PGI) was used. This individual-level measure asks patients to identify five important life domains and weigh each domain according to its importance. Patients were additionally interviewed post-stroke to assess whether they indicated to have experienced response shift.	Persons with stroke experience a sudden loss of function, which gradually recovers with time. The implicit assumption of this study was that at different stages of recovery, patients may find different life domains important (reconceptualization) and/or how change how important different domains are relative to one another (reprioritization).	Of the 46 interviewed patients, 13 (28%) expressed verbalizations reflecting a response shift. The majority selected entirely different life domains at 24 weeks and changed their importance accordingly.	Response shift may occur during recovery (rehabilitation intervention and natural recovery). When ignored, this could lead to inappropriate goal setting that is not aligned with the domains of importance to the patient nor with his/her priorities of that moment.
Decisions about the need for continuation or addition of interventions or supportive services	In patients diagnosed with major depressive disorder participating in a randomized trial comparing psychotherapy with anti-depressant medication, response shift was found, with this response shift effect being larger in the psychotherapy group [39].	The Beck Depression Inventory (BDI), a 21-item questionnaire designed to measure behavioral symptoms of depression, was used.	The implicit assumption of this study is that psychotherapy includes education about depression, which could change patients' concept of the disorder and the way they view their symptoms, and in turn influence the way in which patients respond to PROM items.	Patients who received psychotherapy, compared to the medication group, were better at assessing their level of symptomology, viewed depression as a more unified concept, and had become more aware of their depressive symptoms.	Patients may not receive interventions/services such as psychotherapy because clinicians may believe that the worsened PROM scores were the result of the intervention rather than response shift, which allowed patients to be more aware of their condition and its symptoms.

(i.e., most chronic conditions), patients can be taught to change their values and frame of reference as part of a process of adaptation. In those situations, response shift itself is the goal of treatment, as is frequently the case in rehabilitation [38, 55, 56], patient education (see, for example, Kiresuk and Sherman [57], Reuben and Jennings [58], and Nolte et al. [59]), and most psychological treatments (e.g., cognitive behavioral therapy) [60].

When a decision needs to be made about *continuation or addition of interventions or supportive services*, patients' self-reports on their health status can be used to assess the changes they may have experienced. Routine collection of PROM data in clinical practice has been recommended for purposes of monitoring the impact of health conditions and interventions from the patient's point of view. Examples include the increasing use of PROM data in electronic medical records of healthcare providers, as well as personal health records that are maintained by patients themselves [61]. The PROM data can be used to determine whether continuation or addition of interventions or supportive services are required to reduce symptoms or better address functional problems or concerns of the patient. A particular risk with this type of use of PROM data is that ignoring response shifts can lead to erroneous conclusions about whether interventions and supportive services are achieving the desired outcomes. For example, education about depression is often provided as part of psychotherapy treatment aimed at changing a patient's perspective on depression and how they view depressive symptoms. A study by Fokkema et al. [39] showed that the resulting reconceptualization response shift could result in deterioration, rather than improvement in PROM depression scores, due to increased awareness and recognition of their depressive symptoms (see also Table 3). Subsequently, a clinician may attribute this deterioration to the treatment itself (e.g., education with psychotherapy) and may therefore no longer continue the treatment even though the patient has experienced an improvement in depressive symptoms. This situation illustrates the importance of interpreting PROM depression scores in dialogue with the patient as the basis for making an accurate assessment of the change of the patients' health status over time.

### Meso-level of decision-making

At the meso-level, the focus is on the use of PROM data for decisions made by healthcare managers and administrators for quality improvement, performance monitoring, reimbursement, and accreditation [62, 63] (see Fig. 1 and Table 4). Meso-level decision-making uses aggregated PROM data. Translating individual-level data to group-level decision-making could lead to biases and skewed inferences for multiple reasons, including failure to account for individual-level response shifts.

Decisions about *quality improvement* initiatives involve patients' evaluations of healthcare systems and use of PROM data. Often these efforts are aimed at improving efficiency of healthcare delivery, to optimize limited resources and control costs for the health system. These within-system efforts are usually directed and used by administration. Response shift is among the threats to valid inferences about impacts of such initiatives, but is rarely examined. For example, a study by Osborne et al. [40] reveals that participants of a self-management program may respond differently to evaluative questions about their health posed after the introduction of the program compared to before. Results suggest that 87% of the participants experienced response shift (see Table 4). For example, response shift occurred when participants realized that their health before the program was worse than they thought. This would result in an underestimation of the improvements in PROM scores over time. If the possibility of response shift is ignored, this could lead to an ill-informed meso-level decision to cancel or alter the self-management program. To prevent the possibility of this ill-informed decision, hermeneutics can be used to ask participants about contextual factors that could have altered their perspectives and to seek a balanced appreciation of the diversity of perspectives when meso-level decisions are being made based on PROMs results.

Other meso-level decisions relate to *performance monitoring* [64], which involves between-system comparisons based on PROM data. For example, the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey is administered by the Centers for Medicaid and Medicare (CMS) in the US as part of their public reporting or reimbursement programs. The survey provides information on the quality of health services at multiple levels of the healthcare system and allows comparisons across healthcare providers. Patient-reported "experience" variables are an important component of this comparison, and randomly selected hospital patients are surveyed cross-sectionally or longitudinally. However, comparisons between systems may potentially provide misleading results if response shift is ignored. For example, a study by Feldman et al. [65] comparing different primary healthcare organizational models found that PROM scores were equivalent despite differences in complex health problems presented in those organizations (for a similar attempt of large scale PROM use in the UK, see [66]). A Hermeneutic perspective of openness and reflexivity points to the importance of exploring the possibility of response shift in patients with complex problems. This may, for example, occur as a result of patients adapting to their limitations or living with comorbidities for a prolonged length of time.

Similarly, *accreditation of healthcare organizations* that relies on evaluation-based outcomes may be impacted by response shift. In addition to PROs, accreditation also often

**Table 4** Examples and potential implications of response shift at the healthcare institution (meso) level based on Gadamerian hermeneutics

Type of decision	Response shift example	Contextual elements: purposes for using PROMs	Dialectical interaction: assumptions influencing interpretation	Openness and reflexivity: impact of response shift on inferences about meaning of the scores	Potential consequences: How response shift affects decision-making
Decision about quality improvement initiatives	The arthritis foundation assessed their chronic disease self-management program and found that participants responded differently to questions posed after the program than they did to the same questions posed before [40].	The Health Education Impact Questionnaire (HEI-Q) was used to assess response shift in measuring the benefits of self-management programs.	Traditional evaluation of health intervention programs assumes that participants do not change the way they perceive their health over the period of the program.	The majority of participants (~87%) had an altered internal standard of measurement on at least 1 item, which may obscure self-management programs as having small or no effects.	Quality improvement initiatives of self-management programs may not show any difference unless response shift is used as an outcome or at least is taken into account.
Decisions based on performance monitoring	In a study comparing primary healthcare organizational models and PROs, PROM scores over a period of 18 months of patients in various clinics were equivalent despite differences in complex health problems [65].	The Medical Outcomes Short Form-36 (SF-36), the Health Assessment Questionnaire (HAQ), the Minnesota Living with Heart Failure Questionnaire, the Chronic Respiratory Questionnaire, and the Audit of Diabetes-Dependent Quality of Life (ADDQoL) were used to assess primary healthcare organizational models.	The implicit assumption of the study is that certain organizational models (e.g., community-oriented model) may be more effective in improving health outcomes in patients with chronic disease.	The largest portion of the variance in PROM scores was attributed to individual factors (e.g., number of comorbidities), whereas only 4.3% was accounted for by organizational factors. There was not much change in health status and QoL over the study period, which the authors attributed to the slow progress of the conditions, attrition of the worst cases, and the possible occurrence of response shift as a result of adaption to limitations.	Ignoring response shift could lead to PROM scores of community practices being equivalent to other types of primary healthcare organizations, even when community practices have more patients with complex health needs
Accreditation of healthcare organizations	Residents participating in an adolescent medicine rotation training program aimed at improving their ability to manage adolescents' health issues demonstrated a significant increase in self-perceived skill levels for all assessed domains. Participation in didactic instruction did not yield significant additional benefit for any of the assessed domains. The program was required by Accreditation Council for Graduate Medical Education [67].	Survey-based resident self-assessment was used to compare differential gain in self-assessed skills and confidence in residents exposed to a didactic curriculum to those who participated in the clinical rotation-only group. Assessments took place after the program and were administered as a post- and then-test.	Post-then comparison was assumed to reduce response shift bias seen in traditional pre- and post-test design particularly with self-assessment tools.	There were no differences in self-assessed skills and confidence in residents exposed to a didactic curriculum to those who participated in the clinical rotation-only group as the post-then comparisons took possible response shift effects into account.	If the possibility of response shift is ignored, groups that should differ may have comparable PROM scores (conversely, differences could be amplified where none exist), which could result in undue accreditation.



relies on other evaluation-based outcomes (e.g., self-perceived competence or skill) in different populations (e.g., healthcare professionals). For example, Ruedinger et al. [67] chose a different approach in an accreditation study for the adolescent medicine training program. They compared a didactic curriculum intervention group with a clinical rotation-only group at one point in time following these interventions. They administered a post- and then-test to examine retrospectively the change pediatric resident trainees may have undergone. While recognizing other types of biases present (e.g., recall bias and social desirability bias), the authors noted that this study design helped to reduce response shift bias and allowed for comparable self-assessment scores. With many organizations (e.g., National Quality Forum) supporting increased collection of PROM data and evaluation-based outcomes of other populations (e.g., healthcare professionals) for quality measurement [68], questions regarding how to best account for response shift in quality rating, reimbursement, and accreditation decisions are likely to grow.

### Macro-level of decision-making

At the macro-level, implications of response shift focus on the use of PROMs by governments for health policy purposes [69, 70] (See Fig. 1, and Table 5). The effect of response shift on PROM results readily translates from the micro- and meso-level to the macro-level decisions regarding healthcare coverage. For example, a study of the September 11 terror attacks illustrates how a population-level calamity may result in response shifts resulting in reprioritization of relationships, compassion, and spirituality in a large segment of the population [71]. Consequently, when ignoring response shift, decisions about healthcare coverage may not fully address the need for services corresponding with these changing priorities. Decisions about healthcare coverage, including *provision and reimbursement of healthcare*, are increasingly informed by PROM data. Health technology assessment supports such decision-making by synthesizing and evaluating diverse types of evidence, typically by aggregation of individual responses to patient preference measures [72]. In addition to evidence of effectiveness, healthcare needs and cost-effectiveness are important criteria for provision and reimbursement of care. Whereas treatment effectiveness (e.g., as investigated with randomized clinical trials) and healthcare needs (see micro-level) are typically assessed in patients, there is debate about whether values in cost-effectiveness analyses should preferably be provided by patients or by the “healthy population” or public. The possible occurrence of response shift in patients adds importance to this debate, as the public may not be representative for patient populations. Quality-adjusted life years (QALYs) are often used as an outcome measure in

cost-effective analyses that provides a uniform economic reference framework across healthcare. QALYs are derived from patients’ responses to PROMs by combining length of life with quality of life (QoL). As such, QALYs have an evaluative component that is particularly susceptible to response shift. Here, quality is measured by utility, which reflects the value of QoL on a scale anchored at 100% for perfect health and 0% for health that is as bad as dead. Economically, larger QALY gains justify higher costs. Guidelines in many countries prefer utilities to be assessed from a societal perspective [71, 73, 74]. As taxes and health insurance premiums are mostly paid by the general public, they should also have a say in what is important. This societal approach has the additional advantage that it reduces the impact of response shift. If patients adapt to a particular condition and no longer experience the burden, then relieving that condition would not represent a QALY gain, thus reducing the extent to which costs are acceptable. The general public mostly does not experience the particular condition, so they cannot adapt to it either. As a result, the public generally gives more weight to conditions than patients undergoing the condition, with the exception of depression [75].

Since societal valuations still need to be linked to measurements from patients who are experiencing the conditions and treatments, utility scores may be attached to disease-specific questionnaires. This may be a relatively objective and sensitive approach, but tends to overestimate utility in poor health states [76]. Instead, generic questionnaires have been developed that span a wide range of health and QoL domains, and for which societal valuations are available to calculate utility scores [77–80]. These questionnaires may reduce the impact of response shift by asking patients to provide a more factual *description* of their health, whereas the more subjective *valuation* is provided by the non-adapting general public. Over the past decades, utility scores have been estimated in many countries, acknowledging that calibration, prioritization, and conceptualization may differ over place and time [81]. As utility questionnaires have been criticized for favoring the cure sector, current research includes attempts to capture reprioritization and reconceptualization for the care sector and in particular end-of-life care [82, 83].

### Considerations for decision-making

Two key questions need to be considered for healthcare decision-making at the micro-, meso-, and macro-level. First, should decision-making be based on adjusted data (where response shift is taken into account), or unadjusted data (where response shift affects the scores)? Second, will the decision-making be informed by retrospective data that may be affected by response shift, or by the possible occurrence of response shift in the future, using prospective data? The answers to these questions will depend on the particular

**Table 5** Examples and potential implications of response shift at the health policy (macro) level based on Gadamerian hermeneutics

Type of decision	Response shift example	Contextual elements: purposes for using PROMs	Dialectical interaction: assumptions influencing interpretation	Openness and reflexivity: impact of response shift on inferences about meaning of the scores	Potential consequences: How response shift affects decision-making
Healthcare coverage	Following the September 11 terror attacks, Manhattan residents were assessed for post-traumatic stress disorder. While most residents reported one or more stress symptoms, positive effects were also reported, including closer relationships with others, and increased compassion and spirituality [71].	Various measures were used to assess the psychological trauma of September 11, including: the 20-item Center for Epidemiological Studies-Depression Scale, the eight-item Hope Scale, and the 20-item Spiritual Meaning-Long Form.	Response shift has generally been assumed to be an individual-level phenomenon, triggered by a perceived change in an individual's health status as a result of treatment or disease, rather than a population-level phenomenon.	The event triggered a reprioritization of values with close relationships, compassion, and spirituality leading respondents to perceive PROM items independent of any actual change in functional status.	If population response shift is ignored, large segments of the population may face unmet needs (e.g., spiritual care) that are not covered by healthcare or addressed in the current treatment guidelines.
Provision and reimbursement of healthcare services	The objective of this study was to describe a framework for understanding the potential mechanisms that play a role in health state valuations when making reimbursement decisions [94].	Health state utilities (e.g., quality-adjusted life years) are used for various health-care decisions, including reimbursement of treatments based on their cost-utility.	Health state utility measure is assumed to be relatively similar in outcomes for comparative cost/utility analyses between the general public and patients.	The general public who are asked to imagine experiencing health states assign lower utilities to those states than do patients who are actually experiencing these states.	If differences in values provided by respondent groups (e.g., public vs. patients) are ignored such that patients with serious disabilities report high utility scores due to response shift, funding for treatment may be withdrawn based on inaccurate assessment of utility.

decision at stake, the specific patient, organization, or population under consideration, and the nature and type of the health definition used. Additional considerations pertaining to each level of decision-making also need to be taken into account.

First, at the micro-level, healthcare providers are advised to interpret a patient's self-reported outcomes data in dialogue with the patient [84] to ascertain the possibility of response shift, and let this inform the decision that is to be made; otherwise, a simplistic routine use of PROMs may do more harm than good. However, when dialogue with the patient is not possible, secondary sources of information could also be considered in determining whether adaptation, or other factors influencing response shift, may be taking place. For example, when patients have been exposed to interventions focused on facilitating an adaptive process (e.g., developing better coping skills after chemotherapy), an improvement in PROM scores may be interpreted as evidence of adaptation irrespective of there being any change in a person's actual health status [55, 85].

Second, at the meso-level, it is important to consider that decision-making is often based on cross-sectional data, or on system-level longitudinal data, in which within-subject (individual-level) longitudinal change is not explicitly measured. In these cases, response shift may not be identified even if it occurred. Where possible, we therefore recommend obtaining and analyzing longitudinal individual-level data to examine implications of response shift, and making statistical adjustments to account for the possibility of differential response shifts in some groups of people (e.g., by using latent variable methods [86], while taking limitations of current approaches into account [8]). Aggregating these adjusted PROM data for the purposes of quality improvement, performance monitoring, or accreditation will provide a sound basis for meso-level decision-making.

Third, at the macro-level, it is particularly important to critically reflect on the conceptualization of health and its implications for healthcare policy. For example, if the biomedical definition of health as the absence of pathology is used [87], then response shift must be controlled for to distinguish a change in individuals' reported health status from their biomedical health status. In contrast, if a broader and positively phrased conception of health is used to include "physical, mental, and social well-being" [88] and the "ability to adapt and self manage" [89], then response shift may not necessarily need to be controlled for.

## Epilogue

The key message of this paper is that response shifts could either reduce or increase the size of the change in PROM results among people with the same health state, thereby

providing misleading information, which in turn may affect the quality of healthcare decision-making at the micro-, meso-, and macro-levels. If response shift is measured and taken into account [86], the results are expected to be more sensitive by teasing out the effect of adaptation leading to response shift from the actual health change, thereby providing a sound basis for medical decision-making. This is one of the first papers that started to outline possible influences of response shift on healthcare decision-making [5]. The currently described impacts may not be exhaustive nor sufficiently nuanced. Given the novelty of considering response shift in the context of healthcare decision-making, we need empirical studies to examine under what circumstances response shift affects which types of decisions to what extent. Such studies may also teach us how response shift effects relate to other known biases due to, for example, other forms of response bias, representativeness, and missing data. With more empirical data available, we also expect to improve our understanding of how to account for response shift to obtain a more stringent basis for decision-making. It is hoped that increased awareness of the potential, heretofore neglected influence of response shift, will improve the quality of healthcare decision-making at all levels.

**Supplementary Information** The online version of this article (<https://doi.org/10.1007/s11136-021-02766-9>) contains supplementary material, which is available to authorized users.

**Funding** Open access funding provided by University of Gothenburg.. This analysis was undertaken, in part, thanks to funding from the Canada Research Chairs (CRC) program supporting Dr. Sawatzky's CRC in Person-Centered Outcomes.

## Compliance with ethical standards

**Conflict of interest** All authors declare that they have no conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ilie, G., Bradfield, J., Moodie, L., Lawen, T., Ilie, A., Lawen, Z., ..., Rutledge, R. D. H. (2019). The role of response-shift in studies assessing quality of life outcomes among cancer patients:

- A systematic review. *Frontiers in Oncology*, 9, 783. <https://doi.org/10.3389/fonc.2019.00783>.
2. Sajobi, T. T., Brahmabatt, R., Lix, L. M., Zumbo, B. D., & Sawatzky, R. (2018). Scoping review of response shift methods: Current reporting practices and recommendations. *Quality of Life Research*, 27(5), 1133–1146. <https://doi.org/10.1007/s11136-017-1751-x>.
  3. Sprangers, M. A., & Schwartz, C. E. (1999). The challenge of response shift for quality-of-life-based clinical oncology research. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, 10(7), 747–749.
  4. Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
  5. Hartog, I. D., Willems, D. L., van den Hout, W. B., Scherer-Rath, M., Oreel, T. H., Henriques, J. P. S., ..., Sprangers, M. A. G. (2019). Influence of response shift and disposition on patient-reported outcomes may lead to suboptimal medical decisions: A medical ethics perspective. *BMC Medical Ethics*, 20(1), 61. <https://doi.org/10.1186/s12910-019-0397-3>.
  6. Sprangers, M. A. G., Sajobi, T., Vanier, A., Mayo, N. E., Sawatzky, R., Lix, L. M., ... Response Shift—in Sync Working Group (2021). Response shift: Synthesis of the work to date and future directions. *Quality of Life Research*. <https://doi.org/10.1007/s11136-020-02747-4>.
  7. Vanier, A., Oort, F., McClimans, L., Ow, N., Gulek, B. G., Böhnke, J. R., ... Response Shift—in Sync Working Group. (n.d.). *Response shift in patient-reported outcomes measures: A formal definition and a revised model*. Manuscript, under review (same issue).
  8. Sébille, V., Lix, L. M., Ayilara, O. F., Sajobi, T. T., Janssens, A. C. J. W., Sawatzky, R., Response Shift—in Sync Working Group (2021). Critical examination of current response shift methods and proposal for advancing new methods. *Quality of Life Research*. <https://doi.org/10.1007/s11136-020-02755-4>.
  9. Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research*, 45(1), 35–44. <https://doi.org/10.1023/A:1006964925094>.
  10. Hubley, A. M., & Zumbo, B. D. (2013). Psychometric characteristics of assessment procedures: An overview. In *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology* (pp. 3–19). Washington, DC, US: American Psychological Association. <https://doi.org/10.1037/14047-001>.
  11. Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 65–82). Charlotte, NC: IAP Information Age Publishing.
  12. Zumbo, B. D., & Chan, E. K. H. (Eds.). (2014). *Validity and validation in social, behavioral, and health sciences*. Springer International Publishing. Retrieved from <http://www.springer.com/gp/book/9783319077932>.
  13. Zumbo, B. D., & Hubley, A. M. (Eds.). (2017). *Understanding and investigating response processes in validation research*. Springer International Publishing. Retrieved from <http://www.springer.com/gp/book/9783319561288>.
  14. Hawkins, M., Elsworth, G. R., & Osborne, R. H. (2018). Application of validity theory and methodology to patient-reported outcome measures (PROMs): Building an argument for validity. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(7), 1695–1710. <https://doi.org/10.1007/s11136-018-1815-6>.
  15. Edwards, M. C., Slagle, A., Rubright, J. D., & Wirth, R. J. (2018). Fit for purpose and modern validity theory in clinical outcomes assessment. *Quality of Life Research*, 27(7), 1711–1720. <https://doi.org/10.1007/s11136-017-1644-z>.
  16. Sawatzky, R., Chan, E. K. H., Zumbo, B. D., Ahmed, S., Bartlett, S. J., Bingham, C. O., ..., Lix, L. M. (2017). Montreal Accord on Patient-Reported Outcomes (PROs) use series-Paper 7: Modern perspectives of measurement validation emphasize justification of inferences based on patient reported outcome scores. *Journal of Clinical Epidemiology*, 89, 154–159. <https://doi.org/10.1016/j.jclinepi.2016.12.002>.
  17. Howard, G. S., Ralph, K. M., Gulanic, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychological Measurement*, 3(1), 1–23. <https://doi.org/10.1177/014662167900300101>.
  18. Oort, F. J. (2005). Towards a formal definition of response shift (in reply to G.W. Donaldson). *Quality of Life Research*, 14(10), 2353–2355.
  19. Oort, F. J., Visser, M. R. M., & Sprangers, M. A. G. (2009). Formal definitions of measurement bias and explanation bias clarify measurement and conceptual perspectives on response shift. *Journal of Clinical Epidemiology*, 62(11), 1126–1137. <https://doi.org/10.1016/j.jclinepi.2009.03.013>.
  20. Kwon, J. Y., Thorne, S., & Sawatzky, R. (2018). Interpretation and use of patient-reported outcome measures through a philosophical lens. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*. <https://doi.org/10.1007/s11136-018-2051-9>.
  21. Debesay, J., Näden, D., & Slettebø, A. (2008). How do we close the hermeneutic circle? A Gadamerian approach to justification in interpretation in qualitative studies. *Nursing Inquiry*, 15(1), 57–66. <https://doi.org/10.1111/j.1440-1800.2008.00390.x>.
  22. Gadamer, H. G. (1998). *Truth and Method* (2nd ed.). New York: Continuum.
  23. Krawczyk, M., Sawatzky, R., Schick-Makaroff, K., Stajduhar, K., Öhlen, J., Reimer-Kirkham, S., ... Cohen, R. (2018). Micro-meso-macro practice tensions in using patient-reported outcome and experience measures in hospital palliative care. *Qualitative Health Research*. <https://doi.org/10.1177/1049732318761366>.
  24. Sutherland, H. J., & Till, J. E. (1993). Quality of life assessments and levels of decision making: Differentiating objectives. *Quality of Life Research*, 2(4), 297–303. <https://doi.org/10.1007/BF00434801>.
  25. Elwyn, G., Dehlendorf, C., Epstein, R. M., Marrin, K., White, J., & Frosch, D. L. (2014). Shared decision making and motivational interviewing: Achieving patient-centred care across the spectrum of health care problems. *Annals of Family Medicine*, 12, 270–275.
  26. Schwartz, C. E., Stucky, B., Rivers, C. S., Noonan, V. K., & Finkelstein, J. A. (2018). Quality of life and adaptation in people with spinal cord injury: Response shift effects from 1 to 5 years postinjury. *Archives of Physical Medicine and Rehabilitation*, 99(8), 1599–1608.e1. <https://doi.org/10.1016/j.apmr.2018.01.028>.
  27. Ahmed, S., Mayo, N. E., Corbiere, M., Wood-Dauphinee, S., Hanley, J., & Cohen, R. (2005). Change in quality of life of people with stroke over time: True change or response shift? *Quality of Life Research*, 14(3), 611–627. <https://doi.org/10.1007/s11136-004-3708-0>.
  28. Blair, H., Wilson, L., Gouick, J., & Gentleman, D. (2010). Individualized vs. global assessments of quality of life after head injury and their susceptibility to response shift. *Brain Injury*, 24(6), 833–843. <https://doi.org/10.3109/02699051003789203>.
  29. Ahmed, S., Mayo, N. E., Wood-Dauphinee, S., Hanley, J. A., & Cohen, S. R. (2005). Using the patient generated index to evaluate response shift post-stroke. *Quality of Life Research*, 14(10), 2247–2257. <https://doi.org/10.1007/s11136-005-8118-4>.
  30. Chin, K., Fukuhara, S., Takahashi, K., Sumi, K., Nakamura, T., Matsumoto, H., ..., Nakamura, T. (2004). Response shift in



- perception of sleepiness in obstructive sleep apnea-hypopnea syndrome before and after treatment with nasal CPAP. *Sleep*, 27(3), 490–493.
31. Li, Y., & Rapkin, B. (2009). Classification and regression tree uncovered hierarchy of psychosocial determinants underlying quality of life response shift in HIV/AIDS. *Journal of Clinical Epidemiology*, 62(11), 1138–1147. <https://doi.org/10.1016/j.jclinepi.2009.03.021>.
  32. Schwartz, C. E., Sprangers, M. A. G., Oort, F. J., Ahmed, S., Bode, R., Li, Y., & Vollmer, T. (2011). Response shift in patients with multiple sclerosis: An application of three statistical techniques. *Quality of Life Research*, 20(10), 1561–1572. <https://doi.org/10.1007/s11136-011-0056-8>.
  33. Wagner, J. A. (2005). Response shift and glycemic control in children with diabetes. *Health and Quality of Life Outcomes*, 3, 38. <https://doi.org/10.1186/1477-7525-3-38>.
  34. Ahmed, S., Sawatzky, R., Levesque, J.-F., Ehrmann-Feldman, D., & Schwartz, C. E. (2014). Minimal evidence of response shift in the absence of a catalyst. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 23(9), 2421–2430. <https://doi.org/10.1007/s11136-014-0699-3>.
  35. Joore, M. A., Potjewijd, J., Timmerman, A. A., & Anteunis, L. J. C. (2002). Response shift in the measurement of quality of life in hearing impaired adults after hearing aid fitting. *Quality of Life Research*, 11(4), 299–307.
  36. Razmjou, H., Schwartz, C. E., Yee, A., & Finkelstein, J. A. (2009). Traditional assessment of health outcome following total knee arthroplasty was confounded by response shift phenomenon. *Journal of Clinical Epidemiology*, 62(1), 91–96. <https://doi.org/10.1016/j.jclinepi.2008.08.004>.
  37. Ring, L., Höfer, S., Heuston, F., Harris, D., & O'Boyle, C. A. (2005). Response shift masks the treatment impact on patient reported outcomes (PROs): The example of individual quality of life in edentulous patients. *Health and Quality of Life Outcomes*, 3, 55. <https://doi.org/10.1186/1477-7525-3-55>.
  38. Wilson, I. B. (1999). Clinical understanding and clinical implications of response shift. *Social Science & Medicine* (1982), 48(11), 1577–1588. [https://doi.org/10.1016/s0277-9536\(99\)00050-7](https://doi.org/10.1016/s0277-9536(99)00050-7).
  39. Fokkema, M., Smits, N., Kelderman, H., & Cuijpers, P. (2013). Response shifts in mental health interventions: An illustration of longitudinal measurement invariance. *Psychological Assessment*, 25(2), 520–531. <https://doi.org/10.1037/a0031669>.
  40. Osborne, R. H., Hawkins, M., & Sprangers, M. A. G. (2006). Change of perspective: a measurable and desired outcome of chronic disease self-management intervention programs that violates the premise of preintervention/postintervention assessment. *Arthritis and Rheumatism*, 55(3), 458–465. <https://doi.org/10.1002/art.21982>.
  41. Visser, M. R. M., Oort, F. J., & Sprangers, M. A. G. (2005). Methods to detect response shift in quality of life data: A convergent validity study. *Quality of Life Research*, 14(3), 629–639.
  42. Barclay, R., & Tate, R. B. (2014). Response shift recalibration and reprioritization in health-related quality of life was identified prospectively in older men with and without stroke. *Journal of Clinical Epidemiology*, 67(5), 500–507. <https://doi.org/10.1016/j.jclinepi.2013.12.003>.
  43. Galenkamp, H., Braam, A. W., Huisman, M., & Deeg, D. J. H. (2013). Seventeen-year time trend in poor self-rated health in older adults: Changing contributions of chronic diseases and disability. *European Journal of Public Health*, 23(3), 511–517. <https://doi.org/10.1093/eurpub/cks031>.
  44. Currow, D. C., Farquhar, M., Ward, A. M., Crawford, G. B., & Abernethy, A. P. (2011). Caregivers' perceived adequacy of support in end-stage lung disease: Results of a population survey. *BMC Pulmonary Medicine*, 11(1), 55. <https://doi.org/10.1186/1471-2466-11-55>.
  45. Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A. G., & Fayers, P. M. (2006). The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research*, 15(9), 1533–1550. <https://doi.org/10.1007/s11136-006-0025-9>.
  46. Schwartz, C. E., & Rapkin, B. D. (2004). Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. *Health and Quality of Life Outcomes*, 2, 16. <https://doi.org/10.1186/1477-7525-2-16>.
  47. Dibb, B., & Yardley, L. (2006). How does social comparison within a self-help group influence adjustment to chronic illness? A longitudinal study. *Social Science & Medicine*, 63(6), 1602–1613. <https://doi.org/10.1016/j.socscimed.2006.03.031>.
  48. von Blanckenburg, P., Seifart, U., Conrad, N., Exner, C., Rief, W., & Nestoriuc, Y. (2014). Quality of life in cancer rehabilitation: The role of life goal adjustment. *Psycho-Oncology*, 23(10), 1149–1156. <https://doi.org/10.1002/pon.3538>.
  49. Greenhalgh, J., Gooding, K., Gibbons, E., Dalkin, S., Wright, J., Valderas, J., & Black, N. (2018). How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. *Journal of Patient-Reported Outcomes*, 2, 42. <https://doi.org/10.1186/s41687-018-0061-6>.
  50. Sajobi, T. T., Fiest, K. M., & Wiebe, S. (2014). Changes in quality of life after epilepsy surgery: The role of reprioritization response shift. *Epilepsia*, 55(9), 1331–1338. <https://doi.org/10.1111/epi.12697>.
  51. Eneqvist, T., Nemes, S., Bülow, E., Mohaddes, M., & Rolfson, O. (2018). Can patient-reported outcomes predict re-operations after total hip replacement? *International Orthopaedics*, 42(2), 273–279. <https://doi.org/10.1007/s00264-017-3711-z>.
  52. Pinheiro, L. C., Reshetnyak, E., Sterling, M. R., Richman, J. S., Kern, L. M., & Safford, M. M. (2019). Using health-related quality of life to predict cardiovascular disease events. *Quality of Life Research*, 28(6), 1465–1475. <https://doi.org/10.1007/s11136-019-02103-1>.
  53. Brown, D. S., Thompson, W. W., Zack, M. M., Arnold, S. E., & Barile, J. P. (2015). Associations between health-related quality of life and mortality in older adults. *Prevention Science*, 16(1), 21–30. <https://doi.org/10.1007/s11121-013-0437-z>.
  54. Andreasen, J., Gobbens, R. J. J., Eriksen, H. H., & Overvad, K. (2019). Health-related quality of life at hospital discharge as a predictor for 6-month unplanned readmission and all-cause mortality of acutely admitted older medical patients. *Quality of Life Research*, 28(11), 3015–3024. <https://doi.org/10.1007/s11136-019-02259-w>.
  55. Barclay-Goddard, R., King, J., Dubouloz, C.-J., Schwartz, C. E., & Response Shift Think Tank Working Group. (2012). Building on transformative learning and response shift theory to investigate health-related quality of life changes over time in individuals with chronic health conditions and disability. *Archives of Physical Medicine and Rehabilitation*, 93(2), 214–220. <https://doi.org/10.1016/j.apmr.2011.09.010>.
  56. van Rijn, T. (2009). A physiatrist's view of response shift. *Journal of Clinical Epidemiology*, 62(11), 1191–1195. <https://doi.org/10.1016/j.jclinepi.2009.01.023>.
  57. Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal*, 4(6), 443–453. <https://doi.org/10.1007/BF01530764>.
  58. Reuben, D. B., & Jennings, L. A. (2019). Putting goal-oriented patient care into practice. *Journal of the American Geriatrics Society*, 67(7), 1342–1344. <https://doi.org/10.1111/jgs.15885>.




59. Nolte, S., Elsworth, G. R., Newman, S., & Osborne, R. H. (2013). Measurement issues in the evaluation of chronic disease self-management programs. *Quality of Life Research*, 22(7), 1655–1664. <https://doi.org/10.1007/s11136-012-0317-1>.
60. Miner, A. S., Schueller, S. M., Lattie, E. G., & Mohr, D. C. (2015). Creation and validation of the cognitive and behavioral response to stress scale in a depression trial. *Psychiatry Research*, 230(3), 819–825. <https://doi.org/10.1016/j.psychres.2015.10.033>.
61. Gensheimer, S. G., Wu, A. W., Snyder, C. F., & PRO-EHR Users' Guide Steering Group, & PRO-EHR Users' Guide Working Group. (2018). Oh, the places we'll go: Patient-reported outcomes and electronic health records. *The Patient*, 11(6), 591–598. <https://doi.org/10.1007/s40271-018-0321-9>.
62. Greenhalgh, J., Dalkin, S., Gibbons, E., Wright, J., Valderas, J. M., Meads, D., & Black, N. (2018). How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. *Journal of Health Services Research & Policy*, 23(1), 57–65. <https://doi.org/10.1177/1355819617740925>.
63. Peterson, K., Anderson, J., & Bourne, D. (2018). *Evidence brief: Use of patient reported outcome measures for measurement based care in mental health shared decision-making*. Washington (DC): Department of Veterans Affairs (US). Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK536143/>
64. Cella, D., Hahn, E. A., Jensen, S. E., Butt, Z., Nowinski, C. J., Rothrock, N., & Lohr, K. N. (2015). *Patient-reported outcomes in performance measurement*. Research Triangle Park (NC): RTI Press. Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK424378/>
65. Feldman, D. E., Lévesque, J.-F., Lemieux, V., Tourigny, A., Lavoie, J.-P., & Tousignant, P. (2012). Primary healthcare organization and quality-of-life outcomes for persons with chronic disease. *Healthcare Policy*, 7(3), 59–72.
66. Kyte, D., Cockwell, P., Lencioni, M., Skrybant, M., von Hildebrand, M., Price, G., ..., Calvert, M. (2016). Reflections on the national patient-reported outcome measures (PROMs) programme: Where do we go from here? *Journal of the Royal Society of Medicine*, 109(12), 441–445. <https://doi.org/10.1177/0141076816677856>.
67. Ruedinger, E., Carlin, K., Inwards-Breland, D., & McCarty, C. A. (2019). Effectiveness of the adolescent medicine rotation in improving pediatric residents self-assessed skill and confidence caring for youth. *Journal of Adolescent Health*, 64(4), 530–536. <https://doi.org/10.1016/j.jadohealth.2018.10.007>.
68. National Quality Forum. (2013). Patient reported outcomes (PROs) in performance measurement. National Quality Forum. Retrieved from [http://www.qualityforum.org/Projects/nrPatient-Reported\\_Outcomes/Patient-Reported\\_Outcomes.aspx](http://www.qualityforum.org/Projects/nrPatient-Reported_Outcomes/Patient-Reported_Outcomes.aspx)
69. Black, N., Burke, L., Forrest, C. B., Sieberer, U. H. R., Ahmed, S., Valderas, J. M., ..., Alonso, J. (2016). Patient-reported outcomes: Pathways to better health, better services, and better societies. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 25(5), 1103–1112. <https://doi.org/10.1007/s11136-015-1168-3>.
70. Black, N., & Tan, S. (2013). Use of national clinical databases for informing and for evaluating health care policies. *Health Policy (Amsterdam, Netherlands)*, 109(2), 131–136. <https://doi.org/10.1016/j.healthpol.2012.10.007>.
71. Lau, D., Agborsangaya, C., Sayah, F. A., Wu, X., Ohinmaa, A., & Johnson, J. A. (2012). Population-level response shift: Novel implications for research. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 21(9), 1495–1498. <https://doi.org/10.1007/s11136-011-0064-8>.
72. Garrison, L. P., Neumann, P. J., Willke, R. J., Basu, A., Danzon, P. M., Doshi, J. A., ..., Weinstein, M. C. (2018). A health economics approach to US value assessment frameworks-summary and recommendations of the ISPOR special task force report [7]. *Value in Health*, 21(2), 161–165. <https://doi.org/10.1016/j.jval.2017.12.009>.
73. Bernhard, J., Lowy, A., Maibach, R., Hüry, C., & Swiss Group for Clinical Cancer Research (SAKK) and the Swiss Institute for Applied Cancer Research (SIAC), Bern, Switzerland. (2001). Response shift in the perception of health for utility evaluation. An explorative investigation. *European Journal of Cancer (Oxford England: 1990)*, 37(14), 1729–1735.
74. Helgesson, G., Ernstsson, O., Åström, M., & Burström, K. (2020). Whom should we ask? A systematic literature review of the arguments regarding the most accurate source of information for valuation of health states. *Quality of Life Research*. <https://doi.org/10.1007/s11136-020-02426-4>.
75. Papageorgiou, K., Vermeulen, K. M., Schroevers, M. J., Stiggelbout, A. M., Buskens, E., Krabbe, P. F. M., ..., Ranchor, A. V. (2015). Do individuals with and without depression value depression differently? And if so, why? *Quality of Life Research*, 24(11), 2565–2575. <https://doi.org/10.1007/s11136-015-1018-3>.
76. Mereaglia, M., Whittal, A., Nicod, E., & Drummond, M. (2020). “Mapping” health state utility values from non-preference-based measures: A systematic literature review in rare diseases. *Pharmacoeconomics*. <https://doi.org/10.1007/s40273-020-00897-4>.
77. Al-Janabi, H., Flynn, T. N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. *Quality of Life Research*, 21(1), 167–176. <https://doi.org/10.1007/s11136-011-9927-2>.
78. Brazier, J. E., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care*, 42(9), 851–859. <https://doi.org/10.1097/01.mlr.0000135827.18610.0d>.
79. EuroQol Group. (1990). EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*, 16(3), 199–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9).
80. Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., ..., Boyle, M. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical Care*, 40(2), 113–128. <https://doi.org/10.1097/00005650-200202000-00006>.
81. Roudijk, B., Donders, A. R. T., & Stalmeier, P. F. M. (2019). Cultural values: Can they explain differences in health utilities between countries? *Medical Decision Making*, 39(5), 605–616. <https://doi.org/10.1177/0272989X19841587>.
82. Revicki, D. A., & Lenderking, W. R. (2012). Methods and issues associated with the use of quality-adjusted life-years. *Expert Review of Pharmacoeconomics & Outcomes Research*, 12(1), 105–114. <https://doi.org/10.1586/erp.11.100>.
83. Coast, J., Bailey, C., Orlando, R., Armour, K., Perry, R., Jones, L., & Kinghorn, P. (2018). Adaptation, acceptance and adaptive preferences in health and capability well-being measurement amongst those approaching end of life. *The Patient*, 11(5), 539–546. <https://doi.org/10.1007/s40271-018-0310-z>.
84. Mejdahl, C. T., Schougaard, L. M. V., Hjollund, N. H., Riiskjær, E., Thorne, S., & Lomborg, K. (2018). PRO-based follow-up as a means of self-management support—an interpretive description of the patient perspective. *Journal of Patient-Reported Outcomes*. <https://doi.org/10.1186/s41687-018-0067-0>.
85. Westerman, M. J., The, A.-M., Sprangers, M. A. G., Groen, H. J. M., van der Wal, G., & Hak, T. (2007). Small-cell lung cancer patients are just “a little bit” tired: Response shift and self-presentation in the measurement of fatigue. *Quality of Life Research*, 16(5), 853–861. <https://doi.org/10.1007/s11136-007-9178-4>.
86. Oort, F. J. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment*,

- Care and Rehabilitation*, 14(3), 587–598. <https://doi.org/10.1007/s11136-004-0830-y>.
87. Boorse, C. (2014). A second rebuttal on health. *The Journal of Medicine and Philosophy*, 39(6), 683–724. <https://doi.org/10.1093/jmp/jhu035>.
  88. World Health Organization. (1946). Preamble to the Constitution of WHO as adopted by the International Health Conference, New York, 19 June–22 July 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of WHO, no. 2, p. 100) and entered into force on 7 April 1948. Retrieved from <http://www.who.int/suggestions/faq/en/>
  89. Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., ..., Smid, H. (2011). How should we define health? *BMJ*. <https://doi.org/10.1136/bmj.d4163>.
  90. Groenvold, M., Fayers, P. M., Sprangers, M. A., Bjorner, J. B., Klee, M. C., Aaronson, N. K., ..., Mouridsen, H. T. (1999). Anxiety and depression in breast cancer patients at low risk of recurrence compared with the general population: A valid comparison? *Journal of Clinical Epidemiology*, 52(6), 523–530.
  91. Andrykowski, M. A., Donovan, K. A., & Jacobsen, P. B. (2009). Magnitude and correlates of response shift in fatigue ratings in women undergoing adjuvant therapy for breast cancer. *Journal of Pain and Symptom Management*, 37(3), 341–351. <https://doi.org/10.1016/j.jpainsymman.2008.03.01>.
  92. Mesa, R. A., Gotlib, J., Gupta, V., Catalano, J. V., Deininger, M. W., Shields, A. L., ..., Verstovsek, S. (2013). Effect of ruxolitinib therapy on myelofibrosis-related symptoms and other patient-reported outcomes in COMFORT-I: A randomized, double-blind, placebo-controlled trial. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 31(10), 1285–1292. <https://doi.org/10.1200/JCO.2012.44.4489>.
  93. Verstovsek, S., Mesa, R. A., Gotlib, J., Levy, R. S., Gupta, V., DiPersio, J. F., ..., COMFORT-I investigators. (2015). Efficacy, safety, and survival with ruxolitinib in patients with myelofibrosis: Results of a median 3-year follow-up of COMFORT-I. *Haematologica*, 100(4), 479–488. <https://doi.org/10.3324/haematol.2014.115840>.
  94. Stiggelbout, A. M., & de Vogel-Voogt, E. (2008). Health state utilities: A framework for studying the gap between the imagined and the real. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 11(1), 76–87. <https://doi.org/10.1111/j.1524-4733.2007.00216.x>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Richard Sawatzky<sup>1,2,3</sup>  · Jae-Yung Kwon<sup>1,4,5</sup> · Ruth Barclay<sup>6</sup> · Cynthia Chauhan<sup>7</sup> · Lori Frank<sup>8</sup> · Wilbert B. van den Hout<sup>9</sup> · Lene Kongsgaard Nielsen<sup>10,11</sup> · Sandra Nolte<sup>12,13</sup> · Mirjam A. G. Sprangers<sup>14</sup> · the Response Shift – in Sync Working Group

<sup>1</sup> School of Nursing, Trinity Western University, 7600 Glover Road, Langley, BC V2Y 1Y1, Canada

<sup>2</sup> Centre for Health Evaluation and Outcome Sciences, Providence Health Care Research Institute, Vancouver, Canada

<sup>3</sup> Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

<sup>4</sup> School of Nursing, University of British Columbia, Vancouver, Canada

<sup>5</sup> School of Nursing, University of Victoria, Victoria, Canada

<sup>6</sup> Department of Physical Therapy, College of Rehabilitation Sciences, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Canada

<sup>7</sup> Patient Representative, Wichita, KS, USA

<sup>8</sup> Behavioral & Policy Sciences, RAND Corporation, Arlington, VA, USA

<sup>9</sup> Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

<sup>10</sup> Department of Haematology, Quality of Life Research Center, Odense University Hospital, Odense, Denmark

<sup>11</sup> Department of Internal Medicine and Cardiology, Regional Hospital Viborg, Viborg, Denmark

<sup>12</sup> ICON GmbH, Munich, Germany

<sup>13</sup> Corporate Member of Freie Universität Berlin, Humboldt-Universität Zu Berlin, and Berlin Institute of Health, Medical Department, Division of Psychosomatic Medicine, Charité – Universitätsmedizin Berlin, Berlin, Germany

<sup>14</sup> Department of Medical Psychology, Amsterdam University Medical Centers, Research Institute Amsterdam Public Health, Amsterdam, The Netherlands