Stability of Changes in Health Status
Next Step in Comprehensively Assessing Patient-Reported Outcomes

Chronic diseases can impair patients’ health status due to symptoms, functional limitations, and impaired quality of life. Accordingly, the US Food and Drug Administration (FDA) has endorsed patient-reported health status as a viable outcome for therapy approval. To better interpret the clinical significance of health status changes, it is vital to ascertain the magnitude of within-patient change that is meaningful to patients. This minimally clinically important difference (MCID) based on patient perceptions and values provides a context for interpreting outcomes on scales that are not otherwise intuitive. For instance, for patients with severe chronic obstructive pulmonary disease, a change of approximately 5 points in the Severe Respiratory Insufficiency Questionnaire score has been associated with clinically meaningful patient-perceived changes in health status.

In the setting of heart failure (HF), the 2 patient-reported health status outcomes tools that are reliable, sensitive, and predictive of clinical events such as hospitalization and mortality are the Kansas City Cardiomyopathy Questionnaire (KCCQ) and the Minnesota Living with Heart Failure Questionnaire (MLWHFQ). Both were designed and validated as disease-specific instruments to quantify the health status of patients with HF. The FDA’s Center for Drug Evaluation and Research endorsed the KCCQ and the FDA’s Center for Devices and Radiologic Health endorsed both the KCCQ and MLWHFQ as outcome assessments, which has led to an increasing use of these questionnaires in clinical trials for regulatory approval.

Currently, many clinical trials define an arbitrary time point for analysis. However, patient health status changes, respectively. Although several studies have established MCIDs of different health assessment tools for various diseases such as cancer, currently there is no uniformly accepted threshold by the FDA. For instance, a 10-point change in the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (QLQ-C30) may be considered clinically meaningful.

Leveraging Health Status Assessments to Understand the Treatment Benefits

By collecting and analyzing data from serial health status assessments, it is possible to examine the trajectory of change over time. Research endorsed the KCCQ and the FDA’s Center for Devices and Radiologic Health endorsed both the KCCQ and MLWHFQ as outcome assessments, which has led to an increasing use of these questionnaires in trials for regulatory approval.

Quantifying Clinically Important Improvements in Health Status
Patient-reported outcomes are composed of questions that are organized into domains and transformed on a linear scale. Although population-level mean differences in changes in scores between the intervention and control groups are the most statistically powerful way to compare groups, researchers also need to provide a clinical lens for interpreting these means. This can be done by categorizing groups into proportions of patients who experience small, moderate, or large improvements or deteriorations. For the evaluation of devices, the FDA has accepted previously published clinically important thresholds of the KCCQ changes based on the perspectives of patients, supplemented with trials demonstrating the device’s associations with clinical events and costs. For evaluation of drugs, the FDA has required additional efforts to further support the interpretation of within-patient change and ascertainment of MCID.

These initiatives have created estimates close to the original thresholds of 5, 10, and 15 to 20 points, indicating small to moderate (but clinically important), moderate to large, and large to very large clinical changes, respectively. Although several studies have also established MCIDs of different health assessment tools for various diseases such as cancer, currently there is no uniformly accepted threshold by the FDA. For instance, a 10-point change in the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (QLQ-C30) may be considered clinically meaningful.

Leveraging Health Status Assessments to Understand the Treatment Benefits

Currently, trials are beginning to report not only the mean differences in scores between groups but also the proportion of participants in each study group who have changes of various magnitudes. This is achieved by conducting responder analysis and reporting the proportions of patients who achieve various thresholds of benefit. These can be converted to deduce the number needed to treat (NNT) for a patient to experience a clinically important change in health status. For example, in the DEFINE-HF (Dapagliflozin Effects on Biomarkers, Symptoms and Functional Status in Patients with HF with Reduced Ejection Fraction) trial more patients treated with dapagliflozin than those treated with placebo had a clinically meaningful improvement (>5 points) in KCCQ overall summary scores (absolute rates, 43% vs 33%; odds ratio [OR], 1.70; 95% CI, 0.98-3.10) and clinical summary scores (absolute rates, 47.0% vs 32%; OR, 2.40; 95% CI, 1.31-4.20) corresponding to NNTs of 10 and 7, respectively. Such information may be valuable for clinicians and patients to understand the potential benefits of treatment.

Time Course and Stability of Changes in Health Status
Currently, many clinical trials define an arbitrary time point for analysis. However, patient health status...
changes over time and some treatments might result in acute benefits (eg, structural intervention for valvular disease) whereas others may require more time (eg, benefit mediated by cardiac remodeling). It is thus important to design studies that capture both the acute and long-term benefits of therapy and to include specific metrics and time points for these assessments in study protocols and statistical analysis plans. For example, earlier studies of sodium-glucose cotransporter 2 inhibitors assessed patients’ health status at 3 to 4 months and noted a benefit. More recent studies captured 2-week health status scores and showed benefit at a sooner time point. By collecting and analyzing data from serial health status assessments, it is possible to examine the trajectory of change over time (Figure).

One concern is the stability of observed changes, ie, the proportion of patients with a sustained health status improvement with an intervention. This is important because observed changes at 1 time point may be related to chance, day-to-day variability in health status in chronic diseases, or lack of sustained benefit by the intervention. By evaluating the stability of response over time, researchers could assess how many patients who initially improved remained improved rather than reverting to a neutral or worsened category. Comparing responses generated at 2 times could help confirm that the benefits observed are robust and sustained. For example, a pooled analysis of 2 trials with 760 participants determined the likelihood of improvement or deterioration in health status with ferric carboxymaltose (FCM) vs placebo in patients with HF and measured the stability of the response from baseline to weeks 12 and 24. Results demonstrated that 78.1% of 237 patients who experienced a 5-point improvement or more in the KCCQ with FCM at week 12 sustained it at week 24, whereas 34% of 61 patients who experienced a 5-point deterioration or more at week 12 in the FCM group were no longer reported as experiencing this deterioration at week 24. The 160 patients who had not experienced improvement by week 12 had an approximately 20% chance of doing so by week 24. This suggests that the favorable response observed with FCM remained stable over time.

These examples highlight the potential to leverage health status outcomes to provide intuitive ways to quantify the benefits of therapeutic changes observed with an intervention. Describing the magnitude of change both early and later can help establish expectations for patients, inform clinicians whether to continue or abandon a therapy if patients do not initially respond, and ensure that the results are unlikely due to reasons other than the benefit from an intervention. As patient-reported outcomes gain momentum, it is important to leverage clinically important thresholds of change and to analyze results over time to provide clinically intuitive understanding of the treatment benefits.

### ARTICLE INFORMATION

Published Online: August 18, 2022.

doi:10.1001/jama.2022.13551

Conflict of Interest Disclosures: Dr Butler has served as a consultant to Abbott, Adrenomed, American Regent, Amgen, Applied Therapeutics, AstraZeneca, Bayer, Boehringer Ingelheim, Cardior, CVRx, Eli Lilly, G3 Pharma, Imbria, Impulse Dynamics, Innolife, Janssen, LivaNova, Medtronic, Merck, Novartis, Novo Nordisk, Occlutech, Pfizer, Sequana Medical, and Vifor. Dr Shazheeb Khan reported no disclosures.

Additional Contributions: We thank John Spertus, MD, MPH, St Luke’s Mid America Heart Institute, University of Missouri-Kansas City, for his insightful comments on the manuscript. No compensation was given.

### REFERENCES


### Figure. Stability of Response in Health-Related Quality-of-Life Trials

<table>
<thead>
<tr>
<th>Potential advantages of reporting stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deduce sustained effect of intervention over extended follow-up</td>
</tr>
<tr>
<td>Show new changes in health status in patients who may have shown no initial change</td>
</tr>
<tr>
<td>Eliminate the interference of chance factors</td>
</tr>
<tr>
<td>Eliminate intrapatient and disease state variability</td>
</tr>
<tr>
<td>Eliminate instrument and testing variability</td>
</tr>
<tr>
<td>Provide stability to the estimates generated</td>
</tr>
</tbody>
</table>

- One concern is the stability of observed changes, ie, the proportion of patients with a sustained health status improvement with an intervention. This is important because observed changes at 1 time point may be related to chance, day-to-day variability in health status in chronic diseases, or lack of sustained benefit by the intervention. By evaluating the stability of response over time, researchers could assess how many patients who initially improved remained improved rather than reverting to a neutral or worsened category. Comparing responses generated at 2 times could help confirm that the benefits observed are robust and sustained. For example, a pooled analysis of 2 trials with 760 participants determined the likelihood of improvement or deterioration in health status with ferric carboxymaltose (FCM) vs placebo in patients with HF and measured the stability of the response from baseline to weeks 12 and 24. Results demonstrated that 78.1% of 237 patients who experienced a 5-point improvement or more in the KCCQ with FCM at week 12 sustained it at week 24, whereas 34% of 61 patients who experienced a 5-point deterioration or more at week 12 in the FCM group were no longer reported as experiencing this deterioration at week 24. The 160 patients who had not experienced improvement by week 12 had an approximately 20% chance of doing so by week 24. This suggests that the favorable response observed with FCM remained stable over time.

- These examples highlight the potential to leverage health status outcomes to provide intuitive ways to quantify the benefits of therapeutic changes observed with an intervention. Describing the magnitude of change both early and later can help establish expectations for patients, inform clinicians whether to continue or abandon a therapy if patients do not initially respond, and ensure that the results are unlikely due to reasons other than the benefit from an intervention. As patient-reported outcomes gain momentum, it is important to leverage clinically important thresholds of change and to analyze results over time to provide clinically intuitive understanding of the treatment benefits.