NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Best Practices for Developing and Testing Risk Adjustment Models Web Meeting #5

The National Quality Forum (NQF) convened a public web meeting for the Best Practices for Developing and Testing Risk Adjustment Models Technical Expert Panel (TEP) on July 30, 2021.

Welcome, Introductions, and Review of Web Meeting Objectives

Dr. Matthew Pickering, NQF Senior Director, began by welcoming participants to the web meeting. Dr. Pickering provided opening remarks and invited Co-chairs Philip Alberti and Karen Joynt Maddox to provide welcoming remarks. Dr. Pickering conducted roll call and then reviewed the meeting agenda and the following meeting objectives: discuss and adjudicate public comments received for the Technical Guidance report; and finalize the Technical Guidance, which includes finalizing good and emerging best practices for social and/or functional status-related risk adjustment within quality measurement, and finalizing the recommendations and considerations for the standard risk adjustment framework.

Web Meeting #4 Recap

Dr. Pickering provided a recap of web meeting #4, which was held on May 13, 2021. The NQF Scientific Methods Panel (SMP) was also invited to that meeting to provide additional input on the discussion items. Dr. Pickering noted that the main objectives of web meeting #4 were to gather TEP input on the Technical Guidance prior to public comment; continue discussion on the appropriateness of a standard risk adjustment framework, including rationale for or against the use of the same set of risk factors for quality and resource use measures; and to discuss emerging good and best practices for social and functional status-related risk adjustment.

Dr. Pickering thanked the TEP for their feedback on the Technical Guidance. During web meeting #4, the TEP expressed that the Technical Guidance should not be too prescriptive. If the minimum standards for risk adjustment were too strict, it might limit novel methods and/or approaches to measurement. The TEP also agreed that the Technical Guidance should have an increased emphasis on the conceptual model and that the empirical testing approaches of risk factors are not deterministic of inclusion in the final risk model. In order to highlight this, NQF staff updated the report prior to public comment so that empirical testing guidance was moved to an appendix of illustrative examples that developers may consider. In addition, the conceptual model section was updated to reflect the added emphasis of its importance. This also included the creation of a graphic to depict the conceptual model. This graphic will be updated based on TEP's input during web meeting #5 and will subsequently be included in the final version of the Technical Guidance.

During review of the previous web meeting, the TEP members made additional important points about the report. Regarding the description of bias, the report may currently overestimate the ability that developers have to mitigate certain biases. The Technical Guidance should ask two questions of developers: 1) what methods did you use to detect any bias in the data, and 2) were you successful in mitigating the bias? TEP members agreed that the ability to identify this is changing dramatically over

time, but measurement science is not yet at a stage where methods to mitigate all bias is well-understood.

Regarding the use of the term "real variables," the TEP clarified the meaning of this term. This is meant to refer to variables a developer would use based on the construction of the conceptual model. Variables that were identified in the model and can be utilized at the patient-level are "real." Proxy variables are variables meant to measure a similar concept identified in the model and can be used when a "real" variable isn't available. Area-level variables are not a proxy for "real" patient-level variables but can also contribute to the final risk adjustment model in the absence of those "real" variables. The TEP noted that a lack of data availability can make empiric justification or testing for not having variable identified by the conceptual model difficult. In this situation, if the data are not available for analysis, it can be hard to provide empiric evidence of accuracy or inaccuracy. NQF staff have included language in the Technical Guidance to reflect the TEP's input on the use of proxy and area-level variables.

Review and Discuss Public Comments on Technical Guidance

Between June 17 and July 19, the draft Technical Guidance report was available for public comment. NQF Received 11 comments from six stakeholders.

Janaki Panchal, NQF Manager, outlined the following comment themes:

- NQF's approach to addressing health equity
- Alignment of NQF's recommendations with Office of the Assistant Secretary for Planning and Evaluation's (ASPE) Report
- Additional clarification on risk stratification
- Empirical testing requirements
- Guidance vs. being overly prescriptive
- Intended use
- Locus of control
- Widespread support and gratitude for the TEP's work

Prior to the web meeting, the TEP was sent a copy of the public comments. NQF drafted proposed responses to the comments and reviewed them with the TEP during the call.

NQF's approach to addressing health equity

One commenter requested additional information on NQF's broader work and approach to addressing health equity and how that work fits into this Risk Adjustment Technical Guidance. The commenter noted that since measurement is used for improvement for more equitable care for all, health equity should be incorporated into the report more. To directly confront discrimination in American healthcare, NQF is applying an equity lens to all aspects of ongoing work, with the goal of empowering healthcare stakeholders to take meaningful and measurable action to achieve health equity. NQF recognizes the commenter's concerns about masking disparities via risk adjustment and so gave suggestions to address this in the Technical Guidance. For example, the report states that social or functional risk factors should only be included in a final risk adjustment model if the conceptual model makes the need clear and the measured entity does not have the ability to improve the risk factor. The TEP agreed with this approach. Some TEP members mentioned that stratification is essential to ensure differences are visible. NQF will ensure that additional work by the National Academy of Medicine (NAM), NQF, and ASPE on this parallel approach of risk adjustment and stratification is cited. The TEP also raised that continuous reporting of stratified results is not commonly available to the public, but should be. One TEP member pointed out that stratification could mean something different to different people and so clarity is needed. For example, it could mean measure results in one hospital stratified by

race, but it could also mean how black patients in a specific hospital do compared to a national average or to another hospital. There was agreement from the TEP for the complementary use of stratification and risk adjustment together.

Alignment of NQF's recommendations with ASPE

One comment received during the commenting period questioned how the recommendations regarding risk adjustment in the Technical Guidance align with the recent ASPE recommendations. In response to this comment, the proposed response highlights that NQF is an independent standards setting organization that convenes expert stakeholder groups to make independent recommendations and assessment. NQF also noted that, similar to the ASPE report, this Technical Guidance acknowledges the importance to risk adjust certain outcome and cost/resource use measures, where appropriate. This Technical Guidance also aligns with the ASPE report in asserting that there is a need to improve risk adjustment overall to meet the demands of a changing healthcare landscape. While ASPE's report focuses on Value-based Purchasing programs, the TEP members agreed that NQF's Technical Guidance has a very specific intent; it is focused on producing guidance for measures to attain NQF endorsement. To that regard, this Technical Guidance describes a framework of minimum standards that developers should consider for social and/or functional risk adjustment within quality measurement. Outcome and cost/resource use measures need to consider the conceptual relationship of social and/or functional risk factors at the start of care, provider actions, the provider locus of control to impact the social risk factor, and the intended use (or the incentives/resources allocated to impact the social/functional risk factors) in determining whether or not social and functional risk adjustment is appropriate. The TEP agreed with this proposed response.

Additional clarification on risk stratification

Another comment sought additional clarification on guidance for stratification and NQF gathered further input from the TEP regarding the risk stratification requirement presented in the Technical Guidance. As stated currently, the report requests that if risk adjustment is deemed appropriate, stratification without adjustment would also be recommended. However, some TEP members raised concerns that this guidance may be prescriptive in a way that it shouldn't be. For example, one TEP member raised that stratification prior to risk adjustment may be problematic. If results are stratified by a variable prior to adjustment, then any apparent differences would appear to be due to the stratified variable, which is not truly associated with a risk factor. Stratification and risk adjustment are not to be thought of as parallel or complementary because there are numerous technical questions that remain unanswered prior to prescribing a standard methodology. The TEP agreed that if a developer does adjust, there can be value in also stratifying to highlight where the issues are. This helps to balance the notion of risk adjustment as masking disparities. One TEP member noted that stratification has been incorporated into payment models (such as the Medicare Shared Savings Program) to allocate funds by strata as well, bringing back the notion that stratification can be interpreted in different ways.

NQF will modify the language of the report to present this consensus – that the most informative (which may mean risk adjusted) stratified results should be presented for understanding disparities, inequities, and/or differences. Further work is needed to better operationalize this minimum standard, but this will take place during the option year, if awarded, for this project and cannot yet be incorporated into the Technical Guidance. The TEP agree with this approach.

Empirical testing requirements

One commenter requested additional information on some specific empirical testing requirements (e.g., consistently providing data on shifting performance of accountable entities and ordering of variables). Although NQF recognizes that standardized empirical testing approaches are beneficial, the intent of

this guidance is not to be prescriptive to the types of empirical testing that the developer should conduct. Instead, the TEP recommended that although not deterministic, developers should examine the empirical evidence in conjunction with the conceptual model. TEP members stressed that measure should still have the greatest validity and reliability as possible, while also thinking about the conceptual model. Each measure will require significant balancing. However, TEP members agreed that this question is focused on operationalizing the Technical Guidance report and is more appropriate to find solutions for these issues in the option year, if awarded. The TEP agreed that these are essential factors to think through, although may differ measure by measure. NQF will clarify this in the Technical Guidance.

Guidance vs. being overly prescriptive

Some comments noted that the Technical Guidance is too rigid, particularly on the position that bivariate testing should not be deterministic. NQF would like to clarify that the intent of this Technical Guidance report is to provide several illustrative examples of empirical testing approaches that developers may consider (Appendix D in the Technical Guidance). The TEP also recommended that although not deterministic, developers should examine the empirical evidence in conjunction with the conceptual model. Developers should also describe the statistical methods used and the results and interpretation of the analyses. Developers should be transparent about their approach and their interpretation of the results. NQF recognizes that more work is needed to operationalize the Technical Guidance further, and NQF hopes to accomplish that in an option year, if awarded. The TEP did not raise any concerns with this response.

Intended use and locus of control

There were some comments on expanding the intended use portion of the conceptual model section to be clearer on when measures should and shouldn't be used in particular programs. NQF proposed the addition of a statement in the intended use section to ask developers to include the applications, if known, where the measure should not be used. The TEP noted that this might be too prescriptive as it is hard to say how a measure shouldn't be used. The TEP requested the developer lay out the context and implications for use as much as possible in their submission. They should identify the potential for bias, locus of control, or other areas of discomfort that measure users ought to consider before deciding to use the measure. One TEP member noted that depending on the funding stream, it can be hard to make explicit statements like that in an NQF submission. The TEP members agreed that they should consider some more specific questions before setting a requirement like this. However, at a minimum, developers should re-evaluate social and/or functional risk adjustment when adapting measures for other uses.

Calibration

One commenter sought additional guidance on situations where miscalibration is identified. The commenter was specifically worried about disentangling miscalibration from a true difference in outcomes. TEP members noted that when building a model, measure developers should test whether the model under- or over-predicts outcomes within a specific subgroup to test its appropriate fit for that subgroup. One TEP member gave an example for consideration: when calibrating a risk model that does not include race, a developer may find that the model was mis-calibrated for race. This has implications for how the model can be used. In this situation, the model cannot be used to make determinations about whether the quality of an entity varies based on the racial composition of their patient population. Developers need to look at the observed to expected ratio to review how the model is performing. One TEP member noted that this test may not identify why the calibration is inaccurate, but it is a red flag, nonetheless. Developers need to think carefully about applying the model appropriately

in this situation. The conceptual model can help developers to think about testing certain subgroups identified earlier in the development process. The TEP further noted that it is usually unwise to build separate models on subgroups unless you have a very large population.

Finalizing Technical Guidance

The TEP then discussed some final items for completing the Technical Guidance report. Hannah Ingber, NQF Senior Analyst, presented a conceptual model figure for the TEP's reactions. TEP members agreed that a graphical depiction is helpful to measure developers and noted a few additional edits. Patient factors present before the start of care should include medical comorbidities. "Functional vulnerabilities" is an odd phrase and so should be changed to "functional factors." TEP members also mentioned that the model does not conform completely to some epidemiological figure conventions. For example, mediators usually sit on the line of the causal pathway between process and outcome rather than above it. TEP members agreed that the model also does not necessarily capture some nuances of care provision. Namely, that there may be steps and actions, depending on the level of analysis, that the accountable entity can take to help mitigate the impact of the risk factor moving forward, not just to improve the measured outcome. This may get lost in the model if it only focuses on the measured outcome. The TEP agreed that a framework for thinking through these factors could include whether a measured entity can influence, mitigate, and/or eliminate a risk factor while developing the conceptual model. NQF staff will incorporate the suggested revisions to the model in preparation for the final Technical Guidance.

Public Comment

Ms. Panchal opened the web meeting to allow for public comment. No public comments were offered.

Next Steps

Ms. Panchal noted that the NQF staff will update the Technical Guidance to reflect the TEP's final discussion and the final version of the Technical Guidance will be posted on September 13, 2021. NQF staff shared the contact information for the project and adjourned the meeting by thanking the TEP for their participation and engagement throughout the project.