

# NQF's Proposed Two-Stage CDP Process: Final Evaluation Report

April 2013



# Contents

- Executive Summary..... 3
- Introduction ..... 6
  - NQF's current CDP..... 6
  - Stakeholder recommendation for CDP redesign ..... 8
- Proposed CDP Redesign: A Two-Stage Process..... 10
  - The proposed two-stage process..... 11
  - Key objectives of the proposed two-stage process ..... 13
- Piloting the Two-Stage Process..... 15
  - The Pilot Project: Gastrointestinal / Genitourinary (GI/GU) Measure Endorsement..... 15
- Evaluation Methodology..... 17
- Evaluation Results..... 18
  - Better use of measure development resources ..... 18
  - More efficient use of NQF resources..... 18
  - Earlier and more balanced opportunities for feedback..... 24
  - More frequent endorsement cycles and more predictable schedules..... 24
  - Earlier identification and resolution of issues related to competing measures and measure harmonization..... 25
- Conclusions Based on Evaluation Results ..... 25
- Implications for the Two-Stage Process and the CDP redesign ..... 28
- Considerations for Future CDP Improvement ..... 28
- Next Steps ..... 29
- Appendix A: Concepts submitted to the GI/GU pilot..... 30
- Appendix B: Data Collection Instruments ..... 31
- Appendix C: NQF Staff..... 40

# NQF's Proposed Two-Stage CDP Process: Final Evaluation Report

## Executive Summary

In the fall of 2011, NQF began to formally explore a potential redesign of the Consensus Development Process (CDP) in response to a request from the American Medical Association (AMA), 38 medical specialty societies, and the Centers for Medicare and Medicaid Services (CMS). These stakeholders asked that a potential redesign of the CDP include the following:

- A staged process for measure evaluation that would allow for earlier review of the importance, relevance, and potential focus of measure topics, and later review of measure specifications and testing results; and
- More flexible scheduling so that so that new measures could be evaluated outside of the 3-year cycle typically used for the measure maintenance process.

As NQF staff began exploring the possible redesign of the CDP, several constraints for the process were articulated, including the following:

- Adherence to the key attributes set forth under the 1995 National Technology and Transfer Advancement Act and OMB Circular A-119 (i.e., openness, balance of interest, due process, an appeals process, and consensus).
- Ability to evaluate concepts, newly submitted measures, and measures undergoing endorsement maintenance within the same process
- Balance—and alignment to the extent possible—the interests of all stakeholders in terms of impact on NQF membership, Committee members, measure developers, and staff
- Underlying assumptions of stable and adequate funding to support more frequent CDP projects and use of "standing" steering committees.

To inform the potential redesign of the CDP, NQF obtained feedback on the current CDP from a variety of stakeholders, including measure developers, NQF member organizations, and NQF's cadre of volunteer staff who have served on steering committees, the CSAC, and the Board of Directors. After considering this feedback, the following two key objectives were deemed to be essential elements for any proposed redesign of the CDP:

- Provision of a determination of whether a “measure concept” satisfies the Importance criterion prior to full specification and/or testing of a measure; and
- Provision of more timely and flexible review of fully specified and tested measures.

Using Lean/Six-Sigma workout techniques, NQF staff, bringing input from NQF membership and measure developers, participated in several CDP improvement events to inform the redesign process. These efforts culminated in the proposal of a two-stage CDP. Key objectives of this proposed design include:

- Better use of measure development resources
- More efficient use of NQF resources
- Earlier and more balanced opportunities for stakeholder feedback
- More frequent endorsement cycles
- More predictable schedules
- Earlier identification and resolution of competing measures/harmonization issues

In Stage 1 of the proposed two-stage process, measure concepts would be evaluated against NQF's importance criterion. Potential related and competing measures also would be identified during this stage. All measures, regardless of their stage of development (i.e., concepts, fully specified measures, fully specified measure with testing, measures undergoing maintenance review) would be required to undergo concept review and evaluation against the importance criterion. Once the importance threshold has been met, a measure or concept would be eligible to move to Stage 2 of the process, where fully specified and tested measures would be evaluated against the remaining three evaluation criteria: scientific acceptability, usability, and feasibility. Measure stewards would have up to 18 months to fully specify and complete testing of their approved concepts before bringing the measure forward for Stage 2 review. Measures that are already fully specified and tested could proceed immediately to Stage 2. At the end of Stage 2, measures ratified by the Board would receive endorsement.

Because the proposed two-stage redesign of the CDP represents a substantial departure from the current CDP, NQF leadership recognized the need to pilot both stages of the proposed process prior to recommending its adoption by the NQF Board of Directors, and the GI/GU project was selected to serve as this pilot. The specific goals of the GI/GU pilot were twofold:

1. To assess the process steps, the estimated timeline for projects, and the tools and materials needed to implement core components of the proposed two-stage CDP.
2. To allow NQF to identify and assess the scalability of the proposed 2 two-stage process, including the implications of the redesigned process on staffing and relevant infrastructure supports.

Due to budgetary and internal process constraints, not all of the components and assumptions of the proposed two-stage CDP could be tested directly through the pilot project. Specifically, two key components of the proposed redesign were not implemented: the standing steering committee structure and a process to enable measure developers to submit fully specified and tested measures within 12-18 months of the initial submission and review of their measure concepts. However, the pilot was able to address several of the process components of the proposed two-stage design, including evaluating concepts and importance in a separate stage, providing an upfront technical review, revising the format for measure submissions, and providing additional resources for developers, committees, and staff.

The consensus development process used in the pilot project is consistent with, but not identical to, NQF's current CDP (version 1.9). Specifically—in addition to the staging of the evaluation process to evaluate measure concepts separately from evaluation of the fully specified and tested measures against remaining criteria—the consensus development process implemented for the pilot differs from the current CDP in the following ways:

- Provision of an upfront technical review
- Earlier opportunity for comment
- Development and use of additional tools for developers and Steering Committee
- Approval of measure concepts.
- Provision of developer checklists to summarize recommendations for improving measure concepts

A total of 18 measure concepts were submitted to Stage 1 of the GI/GU pilot project by eight developer organizations. Of these 18 concepts, ten were concepts only (measure specification and testing not yet started), two were new measures (fully specified and tested), and six were previously endorsed maintenance measures. On August 27-28, a 17-member Steering Committee met in Washington, D.C. to evaluate these concepts. The Steering Committee ultimately recommended 14 concepts for approval, 12 of which were subsequently approved by the CSAC and the NQF Board of Directors. Seven of the approved measure concepts have been submitted for evaluation against the remaining criteria in Stage 2 of the pilot; these measures will be evaluated by the project Steering Committee in April, 2013.

In order to assess the overall scalability and success of the proposed two-stage process, a formal evaluation of both stages of pilot project was planned and initiated. The intent of the evaluation of Stage 1 of the pilot was to capture stakeholders' perspectives of the process midway through the proposed two-stage process, and to assess the effectiveness of both the upfront technical review process and tools developed for use in Stage 1.

The bulk of the evaluation of the pilot was accomplished through a content analysis of qualitative data that were collected via surveys and structured discussions with NQF staff, the project Steering Committee, and measure developers. Some quantitative data also were used to inform the evaluation, including data collected through detailed tracking of project milestones and activities.

Several conclusions regarding the success of the pilot itself, as well as the success of the two-stage process as implemented in the pilot, can be made based on the evaluation of the pilot to date, as follows:

- The goals for the GI/GU pilot were, for the most part, met.
- The proposed two-stage process could enable increased efficiencies for NQF staff and Steering Committees, but could potentially increase the workload for the CSAC and NQF Board.
- Upfront technical review can help to improve the quality of submissions, but the magnitude of that improvement will not be uniform across all developers.
- Upfront technical review is scalable in terms of staff skill sets, but may not be scalable in terms of time requirements.
- The tools developed for the upfront technical review are useful overall, but still need modifications.
- At least some NQF member organizations are willing to provide feedback earlier in process.
- Identification of competing and/or related concepts/measures can be done earlier in the CDP.
- Separating the evaluation process into stages may result in decisions by Steering Committees to recommend concepts that should not go forward.

The implications regarding the proposed two-stage process and the overall CDP redesign effort are as follows:

- While there does seem to have been some value in allowing for review of measure concepts, the complete redesign of the CDP as a two-stage process to facilitate concept review is not advised. Many of the measures that will be evaluated in future CDP projects have been previously endorsed (and thus are already fully specified and tested). For stewards of these measures, the two-stage process will actually increase the workload (i.e., two measure submission processes, two in-person meetings, two rounds of responding to comments), but without any of the attendant savings that could come from not having to specify and/or test measures that do not pass the Importance criterion. Similarly, while there would be savings in

time and resources for NQF staff, steering committees, the CSAC, and Board when subsequent review against the remaining evaluation criteria in Stage 2 is not necessary, those savings would likely be offset due to the duplicative nature of the processes in the two stages, as well as to the required staff resources needed for the upfront technical review process.

- The benefits derived from the formal upfront technical review process do not appear to be a good use of staff time and resources. Although successful to some extent—particularly for developers new to the NQF endorsement process—the desired outcome of higher quality concept submissions did not necessarily materialize as a result of the upfront technical review process. Instead, the success of the process seemed related more to the acceptance of the given feedback rather than to the existence or quality of that feedback.
- The information required for the concept review—and the criteria used to evaluate that information—were not sufficient to provide adequate feedback to measure stewards. Specifically, there are no criteria available under Importance to Measure and Report to justify not recommending the concept when the preliminary specifications are not consistent with the evidence. Also, there were no criteria in place that would allow the Committee to offer consistent and binding feedback regarding measure specifications. Finally, although information regarding the numerator, denominator, and exclusions for the measure concept appeared sufficient to flag potential issues around competing or related measures, it is unknown to what extent these issues can be resolved prior to submission to Stage 2.

**Given the results of the evaluation of the pilot project, NQF will not utilize the two-stage CDP going forward.** Nonetheless, many of the design components from the proposed two-stage CDP were successful and should be incorporated into the existing CDP (i.e., earlier opportunity for multi-stakeholder input; use of the tools developed for the pilot, and more consistent and earlier—but still informal—upfront technical review). Also, NQF should continue to explore ways to allow for concept review at any point in time—both within and outside of scheduled CDPs.

## Introduction

NQF utilizes a formal Consensus Development Process (CDP) to evaluate and endorse consensus standards, including performance measures, best practices, frameworks, and reporting guidelines. The CDP is designed to solicit input from, and carefully consider the interests of, multiple stakeholder groups across the healthcare industry. Over the past 10 years, the procedures that form NQF's CDP and its implementation have evolved to ensure that evaluation of candidate consensus standards continues to follow best practices in performance measurement and standards-setting.

In recent years, the healthcare field has become increasingly focused on performance measurement, quality improvement, and accountability, generating new opportunities for the measurement community, but also creating significant pressures and challenges for all involved. Accordingly, NQF endorsement projects have increased in number and complexity, while simultaneously, stakeholder expectations for the timeliness and effectiveness of the entire measure development, testing, and endorsement process have intensified.

### NQF's current CDP

NQF's current CDP (version 1.9) involves eight principal steps. Each contains several sub-steps and is associated with specific actions. The eight principle steps are:

- 1. Call for Nominations.** As a consensus development project begins, NQF issues a call for nominations for the project's steering committee. The steering committee typically includes a variety of stakeholders (e.g., consumers, purchasers, providers, professionals, plans, and healthcare quality experts) who have relevant knowledge and/or proficiency in quality measurement, clinical practice, disparities, and/or other care of vulnerable populations needed to evaluate. The main function of the steering committee is to evaluate candidate standards against NQF's measure evaluation criteria and make recommendations for endorsement.
- 2. Call for Candidate Standards.** Prior to the official start of a consensus development project, NQF issues a formal call for candidate standards (typically these are performance measures). Each candidate measure has a measure steward who assumes responsibility for the submission of the measure for potential endorsement to NQF. The measure steward is responsible for making the necessary updates to the measure, informing NQF about any changes that are made to the measure on an annual basis, and providing the required measure information for the measure maintenance process that occurs approximately every three years. Several conditions must be met before proposed measures may be considered and evaluated for potential NQF endorsement.
- 3. Candidate Consensus Standards Review.** After the close of a call for candidate standards, the steering committee conducts a detailed review of all submitted standards, sometimes with the help of a technical advisory panel. The duration of a steering committee's review of the candidate consensus standards for a given project can vary depending on the scope of the project, the number of standards under review, and the relative complexity of the standards. During this review process, the steering committee may meet several times, via conference calls and/or in-person meetings, to discuss and evaluate the candidate measures. At the conclusion of this review period, the steering committee is expected to achieve consensus as to whether a candidate measure is suitable for NQF endorsement.
- 4. Public and Member Comment.** After the steering committee completes its initial review of the submitted candidate standards, a draft report detailing the committee's recommendations is posted on the NQF website for review and comment by members of NQF and the public. All submitted comments are compiled and made available to the steering committee for consideration. The steering committee may revise its recommendations in direct response to a specific comment or series of comments that are submitted during this phase of the CDP.
- 5. Member voting.** Once the steering committee has reviewed all comments and has made any desired revisions to their recommendations, NQF members vote on those candidate standards that are recommended by the committee.
- 6. Consensus Standards Approval Committee (CSAC) Decision.** The work of the CSAC focuses on the approval of proposed consensus standards and the ongoing enhancement of NQF's CDP. The CSAC reviews the recommendations of steering committees, the comments received, and the results of NQF member votes. After detailed review of a candidate standard, the CSAC determines if consensus has been reached across the various NQF Member Councils, seeks further input from Council Leaders if there is a lack of consensus, and, on some occasions, requests a second round of Member voting on a particular candidate standard or set of standards. The CSAC can grant full endorsement, time-limited endorsement, or deny endorsement of a candidate standard.

**7. Board Ratification.** CSAC decisions regarding consensus standards are submitted to NQF's Board of Directors. The Board can affirm or deny a CSAC decision. All consensus standards that are approved by the CSAC must be ratified by the Board for endorsement.

**8. Appeals.** After a consensus standard has been formally endorsed by NQF, any interested party may file an appeal of the endorsement decision with the NQF Board of Directors. An appeal may only be filed in response to NQF endorsement of a candidate standard or set of standards (i.e., appeals regarding the decision to not endorse a candidate standard cannot be filed).

In the current CDP, to receive NQF endorsement, a measure submitted to NQF must satisfy the four main evaluation criteria:

- **Importance to Measure and Report.** This must-pass criterion focuses on the extent to which a measure is evidence-based, has variable or overall less-than-optimal performance, and addresses a specific high-impact aspect of healthcare.
- **Scientific Acceptability of the Measure Properties.** This must-pass criterion focuses on the extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care.
- **Usability.** This criterion focuses on the extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.<sup>1</sup>
- **Feasibility.** This criterion focuses on the extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement

Further, if there are competing measures, an endorsed measure must be judged to be “best in class” or have justification for multiple measures; if there are related measures, an endorsed measure must be harmonized or have justification for differences in specification.

The time required for the current CDP is approximately nine months. Scheduling of CDP projects is contingent on available funding; generally, however, both condition-specific (e.g., neurological disorders) and cross-cutting (e.g., patient safety) CDP projects run on a three-year cycle. Typically, measures eligible for evaluation within the current CDP must be fully specified and tested for reliability and validity.<sup>2</sup> Both new and previously endorsed measures (which are periodically re-evaluated as part of a measure maintenance process) can be evaluated in the current CDP; both types of measures go through the same process and must meet the same evaluation criteria.

### Stakeholder recommendation for CDP redesign

In the fall of 2011, the American Medical Association (AMA), 38 medical specialty societies, and the Centers for Medicare and Medicaid Services (CMS) formally requested that NQF consider revising the current CDP so as to incorporate a staged process for measure evaluation that would allow for earlier

---

<sup>1</sup> This Usability criterion was in effect during the period when redesign of the CDP was considered. The revised Usability criterion, renamed Usability and Use, was approved by the NQF Board of Directors in the summer of 2012 and will be implemented in 2013.

<sup>2</sup> Occasionally, untested measures will be accepted for consideration for potential time-limited (one year) endorsement, if an incumbent measure does not address the specific topic of interest in the proposed measure, a critical timeline (e.g., legislative mandate) must be met, and if the measure is not complex (e.g., not a composite measure or an outcome measure requiring risk adjustment).



review of the importance, relevance, and potential focus of measure topics and later review of measure specifications and testing results. An underlying motivation behind this request is that measures often do not make it past the importance criterion. However, because measure developers are currently required to submit fully specified and tested measures for evaluation, developers may make costly investments of time and other resources to specify and test a measure that may not achieve NQF endorsement.

The above organizations also asked that NQF consider more flexible CDP scheduling so that new measures could be evaluated outside of the 3-year cycle typically used for the measure maintenance process. Under the current CDP timelines, developers could wait two or more years to submit a measure for the next maintenance cycle if testing data are not available by the project submission deadline.

Because NQF strives to continually improve its own systems, policies, and processes and is committed to seeking feedback on its processes and responding to stakeholders needs, many *incremental* changes to various processes and procedures within the CDP have been implemented over the past several years. In response to the above-mentioned letter, however, NQF began to formally explore a potential *redesign* of the CDP in the fall of 2011.

### *Constraints on a CDP redesign*

As a voluntary consensus standards-setting organization, NQF's process must demonstrate the key attributes set forth under the 1995 National Technology and Transfer Advancement Act and OMB Circular A-119 (i.e., openness, balance of interest, due process, an appeals process, and consensus). These requirements are addressed through various elements of the CDP, such as multi-stakeholder committees, open public meetings and calls, opportunities for public comment, and appeals.

Because NQF's CDP is organized to address measures within topic areas, NQF staff determined that any redesign of the CDP must accommodate concepts, newly submitted measures, and measures undergoing endorsement maintenance. Also, NQF prioritized its commitment to balance its responsibilities to all stakeholders—not only by fulfilling its obligations under federal laws and regulations, but also in terms of impact on NQF membership, Committee members, measure developers, and staff—and to ensure that the interests of these groups are aligned to the greatest extent possible. Some initial proposals for streamlining the CDP were ultimately judged to be unworkable because they would conflict with either the federal requirements and/or the constraints on the scope/balance of any redesign of the CDP.

Finally, NQF staff articulated two underlying assumptions necessary to fully implement any proposed redesign of the CDP: stable and adequate funding to support more frequent (at least yearly) CDP projects in the future and use of "standing" steering committees.<sup>3</sup> Ideally, these standing committees would consider measure concepts and measures on a preset schedule throughout the year.

---

<sup>3</sup> The plan to utilize standing steering committees was devised prior to considering a redesign of the CDP. A key objective behind the use of standing committees was to increase the consistency of the evaluation process across topics and over time.

### *Additional stakeholder feedback on current CDP*

To inform the potential redesign of the CDP, NQF obtained feedback on the current CDP from a variety of other stakeholders, including measure developers, NQF member organizations, and NQF's cadre of volunteer staff who have served on steering committees, the CSAC, and the NQF Board of Directors.

In general, other measure developers echoed the AMA/CMS/specialty society criticisms, again noting the limited opportunities to submit measures given the typical three-year cycle for submissions and investment of resources in testing without guarantee of endorsement. Developers also noted their concern with consistency of steering committees across projects and topics.

Feedback from NQF membership revolved around two themes: unpredictable project schedules and desire for earlier input. Specifically, members noted that the unpredictable schedules of projects themselves, as well as the attendant unpredictability associated with the nominations and comment periods, leads to challenges for the membership to be responsive and up-to-date on endorsement activities. Members (particularly consumers and purchasers) also articulated a desire to provide earlier input into the process.

Members of steering committees, the CSAC, and the Board voiced their unease that resources are spent on measures that do not achieve endorsement and/or are not focused on areas of greatest need to consumers and purchasers. They also noted that often significant time and effort is spent on reviewing and evaluating measures that are not ready for NQF endorsement, and alluded to the fact that issues with individual measures can delay entire projects. Finally, these stakeholders called attention to the difficulty in addressing harmonization and competing measures in a coordinated and timely manner.

After considering this feedback, the following two key objectives were deemed to be essential elements for any proposed redesign of the CDP:

- Provision of a determination of whether a “measure concept” satisfies the Importance criterion prior to full specification and/or testing of a measure; and
- Provision of more timely and flexible review of fully specified and tested measures.

### **Proposed CDP Redesign: A Two-Stage Process**

To accomplish these key objectives, NQF staff embarked on a 6-month endeavor to redesign the CDP. Staff used the AMA/CMS/ medical society recommendation for a staged process as the basis for the proposed redesign.

Using Lean/Six-Sigma workout techniques, NQF staff, bringing input from NQF membership and measure developers, participated in several CDP improvement events that focused on, among many others, the following topics:

- Benefits and potential risks of a CDP redesign
- Organization of standing steering committees
- Initial design of a two-stage process
- Revisions to existing materials used for review and reporting
- Development of new materials for use in the redesign
- Roles and responsibilities of staff teams
- Quantification of potential impact of new process across the membership, developers, committees, and CSAC/Board

- Planning of a pilot to test the redesigned process
- Development of department infrastructure to support roll-out of new processes
- Development of communication plans for various stakeholder groups

These efforts culminated in the proposal of a two-stage CDP, shown in Figure 1, and are more fully described below.

### The proposed two-stage process

In Stage 1 of the proposed two-stage process, measure concepts would be evaluated against NQF's importance criterion<sup>4</sup>. Potential related and competing measures also would be identified during this stage. All measures, regardless of their stage of development (i.e., concepts, fully specified measures, fully specified measure with testing, measures undergoing maintenance review) would be required to undergo concept review and evaluation against the importance criterion. Information required for the concept review would include the following:

- Description of the measure concept, including:
  - Measure Title
  - Brief Description
  - Numerator statement
  - Preliminary numerator details (not coding)
  - Denominator statement
  - Preliminary denominator details (not coding)
  - Exclusions under consideration
  - Preliminary exclusion details (not coding)
  - Risk adjustment variables under consideration
  - Proposed levels of analysis, data source, settings of care, topic area (taxonomy)
  - For outcome measures, proposed risk adjustment/stratification methodology
- Information to demonstrate that the criteria for importance to measure and report have been met
  - High impact (Importance subcriterion 1a)
  - Opportunity for improvement (Importance subcriterion 1b)
  - Evidence supporting the measure focus (Importance subcriterion 1c)
- Planned use/current use
- Identification of related and competing measures

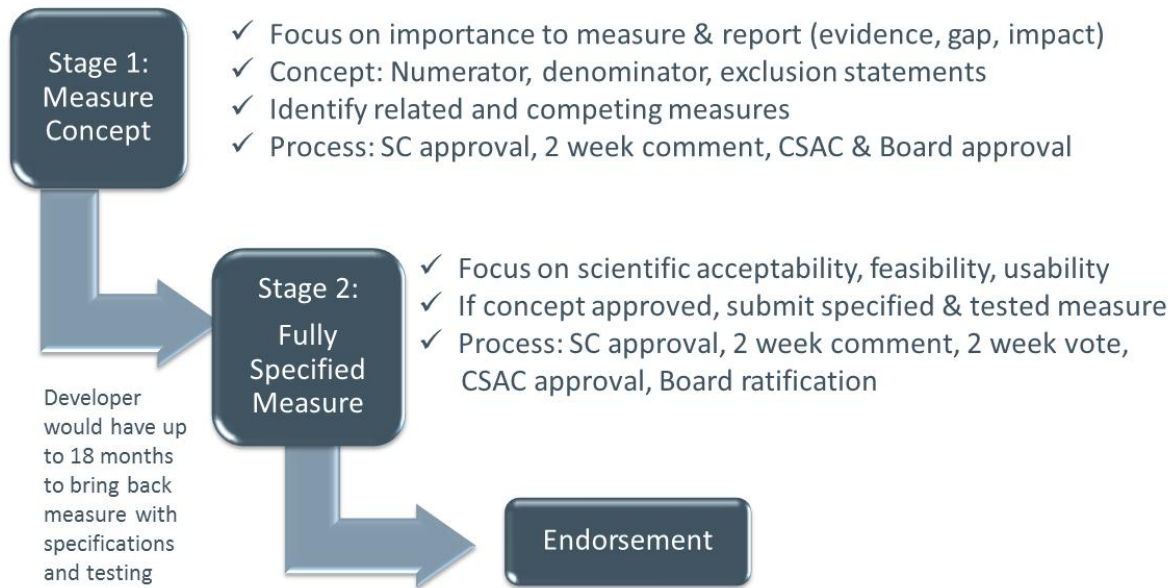
Once the importance threshold has been met, a measure or concept would be eligible to move to Stage 2 of the process. In Stage 2 of the process, fully specified and tested measures would be evaluated against the remaining three evaluation criteria: scientific acceptability, usability, and feasibility. Measure stewards would have up to 18 months to fully specify and complete testing of their approved concepts before bringing the measure forward for Stage 2 review. Measures that are already fully specified and tested could proceed immediately to Stage 2. At the end of Stage 2, measures ratified by the Board would receive endorsement.

---

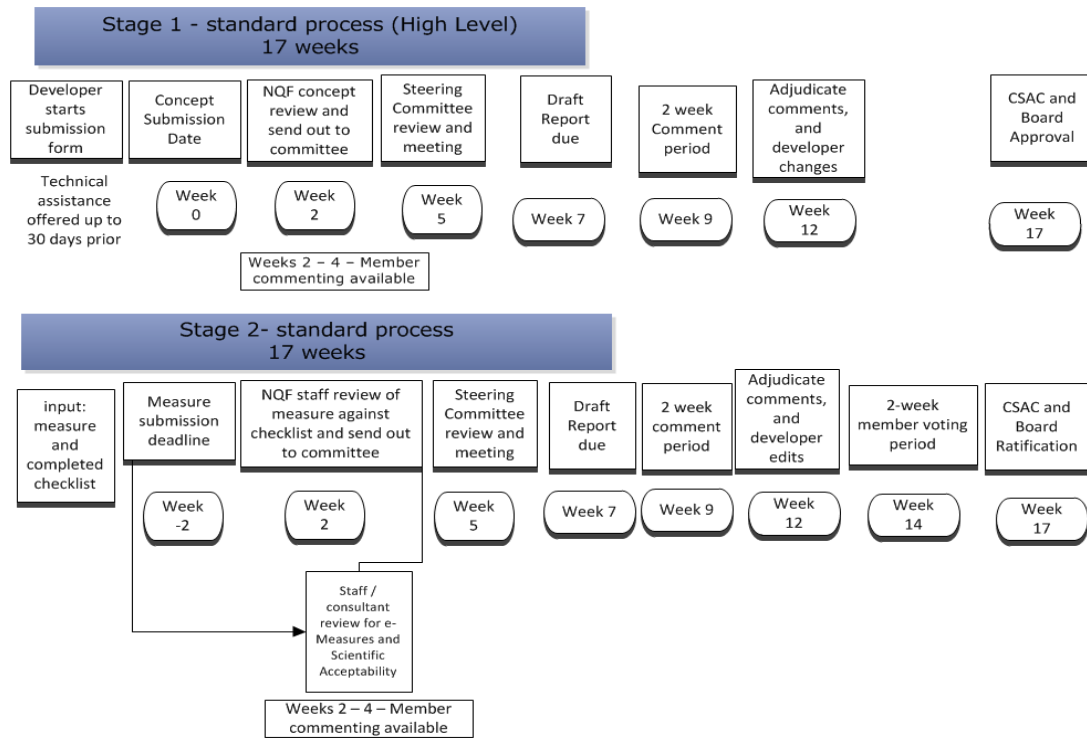
<sup>4</sup> The three subcriteria under the Importance to Measure and Report criterion included—at the time of the CDP redesign—the following: high impact (1a); performance gap/opportunity for improvement (1b); and evidence supports measure focus (1c).

Each stage of the two-stage process would take 17 weeks for completion (see Figure 2). Initiation of CDP projects would be staggered to ensure adequate staffing for projects and feasible timelines for external stakeholders (e.g., public and member commenting, voting).

**Figure 1: Proposed Two-Stage Endorsement Process**



**Figure 2. Timeframe for Stage 1 and Stage 2**



Subsequent to finalizing the two-stage process as detailed above, NQF staff presented preliminary plans to the NQF Board, the CSAC, and other stakeholders, including various NQF councils. The general tenor of the feedback was positive, with the proposed redesign seen as “more than a good faith effort” to align the process with needs of the field. However, stakeholders did voice some concerns regarding the proposed process (e.g., adequacy of a 2-week commenting period; confusion about whether the upfront technical review process would reflect “triage”, “coaching”, or both; added time and activities in two stages for endorsement of maintenance measures).

### Key objectives of the proposed two-stage process

The proposed two-stage process had six key objectives, each of which can be categorized under the two key CDP redesign objectives (see Table 1). Each of these six objectives is described more fully below.

**Table 1. Objectives of the Two-Stage CDP redesign**

		Key objectives of CDP redesign	
		Evaluation of measure concept prior to full measure specification and/or testing	More timely and flexible review of fully specified and tested measures
Key objectives of the two-stage process	Better use of measure development resources	More frequent endorsement cycles	
	More efficient use of NQF resources	More predictable schedules	
	Earlier and more balanced opportunities for stakeholder feedback	Earlier identification and resolution of competing measures/harmonization issues	

*Better use of measure development resources*

Measure stewards who avail themselves of the measure concept approval option prior to fully specifying and testing a measure would not have to expend time, effort, and financial resources on specification and testing of measures that do not pass the importance criterion.

*More efficient use of NQF resources*

For those concepts that do not pass the importance criterion, review and evaluation of the fully specified and tested measure against the remaining criteria would not be necessary, thus saving time and effort for both NQF staff and volunteers (i.e., Steering Committee members, NQF members, CSAC members, the Board, and members of the public).

*Earlier and more balanced opportunities for feedback*

Review of measure concepts would allow for multi-stakeholder input (e.g., NQF expert panels and other key stakeholders such as consumers and purchasers) earlier in the measure development process. Such feedback could potentially aid stewards and developers to refine their measures prior to full evaluation and may increase the likelihood that such a measure would be endorsed.

*More frequent endorsement cycles*

Instead of NQF’s current three-year cycle of projects in measure endorsement topic areas, there would be at least two opportunities for measure stewards to submit measure concepts and/or measures into a topic area each year. More frequent endorsement cycles would also eliminate the need for time-limited endorsement of untested measures.

*More predictable schedules*

Use of standing steering committees would facilitate the establishment of regular, fixed timelines for concept/measure evaluation, thus increasing the predictability of all facets of the endorsement process. Measure developers would have a greater ability to plan their development efforts to correspond with the appropriate endorsement cycles; NQF members and the public would have a better sense of when measures will be available for comment and voting; and Steering Committee members, who participate

on a volunteer basis and must balance their contributions to NQF projects with their already-busy work schedules, would be able to plan for in-person meetings, conference calls, and other NQF-related activities well in advance.

### *Earlier identification and resolution of issues related to competing measures and measure harmonization*

As initially conceptualized, harmonization and competing measure issues would be identified in Stage 1 of the proposed two-stage process, and developers would be expected to address these issues prior to entering Stage 2. For example, if competing measures or concepts were identified, the measure steward could decide to suspend further development of the measure, thus conserving developer resources needed for measure specification and testing. Earlier identification would also allow for earlier collaboration, if desired, between developers for the purposes of measure harmonization or joint measure development.

## **Piloting the Two-Stage Process**

Because the proposed two-stage redesign of the CDP represents a substantial departure from the current CDP, NQF leadership recognized the need to pilot both stages of the proposed process prior to recommending its adoption by the NQF Board of Directors. Piloting the proposed two-stage process has allowed NQF to obtain tangible experience with many of the new processes from the perspectives of measure developers, Steering Committee members, membership, NQF staff, and other stakeholders.

Although Stage 2 of the pilot project is still underway,<sup>5</sup> a formal evaluation of the pilot to date has helped to inform NQF leadership regarding the potential for fulfilling the objectives of the proposed two-stage process and, more broadly, for fulfilling the overall objectives of the CDP redesign effort.

### **The Pilot Project: Gastrointestinal / Genitourinary (GI/GU) Measure Endorsement**

While several previously-scheduled endorsement projects could have served as the pilot project for the proposed two-stage process, the GI/GU project was chosen because it was scheduled to begin soon after the two-stage process was formulated (i.e., in the summer of 2012), both concepts and fully specified and tested measures were available for review within the project, and the project had a relatively small number of concepts/compared to other scheduled projects.

The specific goals of the GI/GU pilot were twofold:

1. To assess the process steps, the estimated timeline for projects, and the tools and materials needed to implement core components of the proposed two-stage CDP.
2. To allow NQF to identify and assess the scalability of the proposed two-stage process, including the implications of the redesigned process on staffing and relevant infrastructure supports.

Unfortunately, because of budgetary and internal process constraints, not all of the components and assumptions of the proposed two-stage CDP could be tested directly through the pilot project. Specifically, two key components of the proposed redesign were not implemented: the standing steering committee structure and a process to enable measure developers to submit fully specified and

---

<sup>5</sup> Due to funding limitations, the timeline for Stage 2 of the pilot has been delayed; current projections anticipate the conclusion of Stage 2 by October 31, 2013.

tested measures within 12-18 months of the initial submission and review of their measure concepts. However, the pilot was able to address several of the process components of the proposed two-stage design, including evaluating concepts and importance in a separate stage, providing an upfront technical review, revising the format for measure submissions, and providing additional resources for developers, committees, and staff.

The consensus development process used in the pilot project is consistent with, but not identical to, NQF's current CDP (version 1.9). Specifically—in addition to the staging of the evaluation process to evaluate measure concepts separately from evaluation of the fully specified and tested measures against remaining criteria—the consensus development process implemented for the pilot differs from the current CDP in the following ways:

- **Upfront technical review.** NQF staff provided an upfront technical review<sup>6</sup> of concept/measure submissions (for both stages) prior to making the materials available to the Steering Committee. These reviews are intended to ensure that submission materials are complete and responsive to the questions needed for concept/measure evaluation and to provide individualized support to those developers who need assistance with the submission process.
- **Earlier opportunity for comment.** A two-week comment period was opened to NQF members prior to the Stage 1 in-person Committee meeting, in order to allow members an opportunity to submit comments for consideration and discussion by the Committee at the in-person meeting. A similar two-week member commenting period will be opened for Stage 2 of the pilot.
- **Development of additional tools.** Several tools were developed to provide additional information and guidance to developers and the Steering Committee regarding NQF's evaluation criteria. The tools used in Stage 1 included a Developer Guidebook, a Steering Committee Guidebook, a concept submission form, an evidence attachment, and a measure tracking log (the latter for internal staff use). In addition, a measure testing attachment was developed for Stage 2 of the pilot. The tools developed for the pilot are described more fully below.
- **Approval of measure concepts.** For Stage 1 of pilot, the Steering Committee recommended “approval” (or not) of measure concepts (rather than a recommendation regarding endorsement as in the current CDP). Actual endorsement of a measure will not be granted until a measure has passed all four NQF criteria at the end of Stage 2.<sup>7</sup>
- **Provision of developer checklists.** A checklist was provided to each developer following the Committee and CSAC review of concepts in Stage 1. These checklists summarize the recommendations for improving the concept(s) and considerations for specifying and testing the measure that must be addressed prior to submission to Stage 2 of the pilot.<sup>8</sup>

---

<sup>6</sup> Although the upfront technical review process was intended to be optional going forward, the review was required for the pilot. Specifically, for Stage 1, measure stewards were asked to submit at least half of their concepts for review; for Stage 2, they were asked to submit at one of each measure type submitted (i.e., process or outcome). The upfront technical review consisted of two levels of review (L1=assessment of completeness of submission and L2=assessment of responsiveness of submission).

<sup>7</sup> Per the proposed two-stage process, stewards whose concepts are approved in Stage 1 have up to 18 months to fully develop, specify, and test their measure(s) before submitting it for Stage 2 review against the remaining NQF criteria. However, in the pilot, Stage 2 commenced immediately after conclusion of Stage 1, and thus the utility of the potential 18-month “gap period” cannot be evaluated.

<sup>8</sup> The developer checklists can be considered either tools or infrastructure supports for Stage 2 of the proposed redesign process; the utility of these checklists will be assessed after the conclusion of Stage 2 of the pilot.



Also, the pilot was staffed somewhat differently than typical CDP projects, with five staff assigned to the pilot rather than three as is currently done. Two staff members (one senior director and one senior project manager) were assigned responsibility for the upfront technical review, while three staff members (one senior director, one senior project manager, and one project analyst) handled the remaining aspects of the CDP.

A total of 18 measure concepts were submitted to Stage 1 of the GI/GU pilot project by eight developer organizations. Of these 18 concepts, ten were concepts only (measure specification and testing not yet started), two were new measures (fully specified and tested), and six were previously endorsed maintenance measures. On August 27-28, a 17-member Steering Committee met in Washington, D.C. to evaluate these concepts. The Steering Committee ultimately recommended 14 concepts for approval, 12 of which were subsequently approved by the CSAC and the Board of Directors. Seven of the approved measure concepts have been submitted for evaluation against the remaining criteria in Stage 2 of the pilot; these measures will be evaluated by the project Steering Committee in April, 2013.

## Evaluation Methodology

In order to assess the overall scalability and success of the proposed two-stage process, a formal evaluation of both stages of the pilot project was planned and initiated. However, as mentioned earlier, not all of the components, assumptions, and processes of the proposed two-stage CDP could be tested directly through the pilot project.

The intent of the evaluation of Stage 1 of the pilot was to capture stakeholders' perspectives of the process midway through the proposed two-stage process, and to assess the effectiveness of both the upfront technical review process and tools developed for use in Stage 1. The bulk of the evaluation of Stage 1 of the pilot was accomplished through a content analysis of qualitative data that were collected via surveys and structured discussions with NQF staff, the project Steering Committee, and measure developers, as follows:

- **NQF staff: Upfront technical review team.** Both team members tasked with upfront technical review participated in two 30-minute individual interviews and one 45-minute group interview.
- **NQF staff: CDP activities team.** All three team members participated in three 30-minute individual interviews and one 45-minute group interview.
- **Project Steering Committee.** Fourteen of the 17 Steering Committee members responded to a written survey and then participated in a 45-minute structured discussion during day two of the pilot project's in-person meeting.
- **Measure developers.** An online survey was distributed to those developer organizations that submitted concepts for the upfront technical review process. All developer staff members involved in the review process were invited to participate in the survey (a total of 30 individuals from seven developer organizations). A total of 10 developers responded to the survey; these individuals represented six of the seven developers who submitted measure concepts to Stage 1 of the pilot.

All individual and group interviews with NQF staff were conducted by an experienced qualitative researcher external to the Performance Measures group. The structured discussion with the project Steering Committee was led by the senior director from the CDP activities team. All data collection instruments used for the Stage 1 evaluation are included in Appendix B.

Quantitative data also were used to inform the evaluation of Stage 1 of the pilot. These data were collected through detailed tracking of project milestones and activities and through the surveys described above.

Data on the measure testing attachment and on the upfront technical review during Stage 2 of the pilot were collected in a telephone survey of four measure developer groups who are participating in Stage 2. This telephone interview was conducted by two NQF staff external to the Performance Measures group. The survey instrument used in this data collection effort also is included in Appendix B.<sup>9</sup>

## Evaluation Results

The results of the evaluation of the pilot to date are presented using the organizational framework of objectives presented in table 1.

### Better use of measure development resources

Assessment of this objective was not included as part of the evaluation of Stage 1 of the pilot.

### More efficient use of NQF resources

Three CDP redesign components address the objective of more efficient use of NQF resources: the actual two-stage review process, wherein the measure concept is evaluated against the importance criteria separately from evaluation of the remaining criteria, implementation of an upfront technical review; and scalability of the overall process.

#### *Two-stage review process*

Ideally, evaluating a measure concept against the Importance criterion prior to, and separate from, evaluation of a full measure against the remaining criteria will result in more efficient review by both staff, project Steering Committees, NQF members, the CSAC, and the Board. This assumption cannot be completely evaluated without considering the cycle time for both stages of the pilot and examining, in-depth, the review process in both stages of the pilot. Nonetheless, results from Stage 1 of the pilot are useful because, for those concepts not approved to move forward to Stage 2 of the process, neither staff nor Committee members will have to spend time reviewing, discussing, and evaluating information submitted in reference to the Scientific Acceptability, Usability, and Feasibility criteria.

#### **Will separate concept review result in increased efficiencies for NQF staff and Steering Committees?**

In Stage 1 of the pilot project, only 12 of the 18 submitted measures were ultimately approved by the CSAC and the Board of Directors. Thus, the time and resources required by staff and the Steering Committee for subsequent review against the remaining evaluation criteria will not be required for one-third of the concepts/measures initially submitted to the project. Typically, reviewers spend the majority of their time on the Importance and Scientific Acceptability criteria; conservatively, therefore, the rejection of one-third of the measure concepts could result in a "time savings" of 16 percent for the complete review of measures by staff and Steering Committees over the life of the project.

---

<sup>9</sup> Because of the decision not to implement the two-stage CDP, the formal evaluation of Stage 2 of the pilot has been discontinued. Thus, results of this survey are the only evaluation data for Stage 2 that are included in this report.

### **Will separate concept review result in increased efficiencies for the CSAC and Board?**

As designed, the two-stage CDP called for the review and approval of measure concepts by both the CSAC and the Board at the end of Stage 1 and review and ratification of the associated measures at the end of Stage 2. For those six measure concepts that were not recommended for approval by the project Steering Committee, total time spent on review by these two groups may be somewhat less than what would be required for a full review. In contrast, for the 12 measure concepts that were approved by the CSAC and Board, a second round of review may be needed (if those measures are recommended by the Steering Committee in Stage 2 of the process). While theoretically the CSAC and Board would be reviewing different facets of the concept/measure in the two stages, at least some overlap would be expected (e.g., getting oriented to the concept/measure).

### *Upfront technical review*

The rationale for implementing an upfront technical review of concept/measure submissions was to identify incomplete submissions or submissions that are not responsive to the items needed for evaluation. Ideally, once identified, the measure steward would amend the submission per the advice of NQF staff and thus improve the overall quality of the submission. Improved submissions should theoretically result in more efficient use of NQF resources because more complete and responsive submissions would be easier and quicker for Committees to review.<sup>10</sup>

### **Did upfront technical review help improve quality of submissions?**

In Stage 1 of the pilot, 14 concepts were submitted for upfront technical review. Staff members participating in the upfront technical review were somewhat ambivalent about the utility of the technical review process. While they considered the overall experience to be very positive, and believed that the one-on-one approach in interacting with and educating measure developers was valuable, they did not think that the upfront technical review necessarily improved the quality of the submissions or, ultimately, helped to better inform Committee deliberations. Arguably this perceived lack of success of the upfront technical review process by the review team was due to their perceptions of developer response to the review. Specifically, team members noted that while some developers seemed appreciative of and responsive to the feedback, others seemed not to value the feedback and appeared unengaged with the process.

These perceptions of developer responses corresponded in part to the feedback provided by developers. When asked about the overall value of the upfront technical review process for Stage 1, 90% of the individual respondents said that it was valuable or somewhat valuable and 80% said they would utilize an upfront technical review option if offered in the future. Further, representatives from all six responding developers said that they revised their submissions based on the feedback. In general, developers seemed to value the actual feedback from staff somewhat more than the initial measure submission guidance that was provided during a pre-submission conference call.<sup>11</sup> Interestingly, opinions of the individual respondents within the developer groups were not necessarily identical. For example, in regard to the perceived overall value of the upfront technical review process, ratings within

---

<sup>10</sup> As initially conceptualized in the proposed two-stage process, incomplete or non-responsive submissions that are not revised in a timely manner would be pulled from the current evaluation cycle but accepted in a later evaluation cycle that would begin within 4-6 months. However, this component of the two-stage process could not be tested in the GI/GU pilot.

<sup>11</sup> The purpose of this one-on-one call with each developer was to orient developers to the pilot project and to introduce developers to the new tools developed for Stage 1 of the pilot.

developer groups typically ranged from "somewhat valuable" to "valuable"; however, for one developer group, one respondent rated the process as "somewhat valuable" while the other rated it as "not valuable".

Developer feedback on the upfront review process for Stage 2 of the pilot was similar to that for Stage 1. Responding developers saw some value in the process, specifically regarding actual errors in their submission, a need to repeat information in multiple sections of the form, a need to move information within the form, or a need for more clarification. While the developers reported revising their submissions to address these types of feedback, they also noted their unwillingness to revise other parts of their submissions, even though feedback was given. In general, the developers did not favor a mandatory upfront technical review; one developer suggested that a partial review (e.g., of particular sections as requested by the developer) might be useful. Two of the four developer groups indicated a willingness to take advantage of the upfront review process again if offered. However, the necessity of submitting measures earlier so as to allow time for the review process was cited by the other two developer groups for not participating in the process in the future. (One noted that, due to their practice of having multiple staff members contribute to the NQF submission, they cannot meet the upfront technical review deadline; the other said it would depend on the number of measures they are submitting and the turn-around time allowed.)

Similar to that of the upfront technical review team's perception of developer response to review, the CDP activities team—whose members guided the measure concepts through the remainder of the evaluation process subsequent to the upfront technical review—perceived developer response to the review process in Stage 1 as marginal. While members of this team thought that measure submissions did improve (at least somewhat) after the upfront technical review process, staff on this team ranked their overall experience of the review process less positively than did the review team staff, in part because of the difficulties in the transition or "hand-off" from the upfront technical review process to the remainder of the Stage 1 activities.<sup>12</sup>

It is not possible to ascertain whether members of the project Steering Committee believed that the upfront technical review resulted in improved concept submissions because they were given only the final (revised) submissions. However, only two-thirds of those who responded to the Steering Committee survey believed that the final submissions were both complete and responsive.

#### **Were the tools developed to facilitate the upfront technical review beneficial?**

Three tools were developed specifically to aid in the upfront technical review process (and, by extension, to help improve the quality of concept/measure submissions). These tools included the following:

- **Concept submission form.** Because the measure submission form used for measures submitted under the current CDP is designed for fully specified and tested measures, a form was needed to elicit only the information needed to evaluate measure concepts. This form is an abbreviated version of the current measure submission, modified as needed to provide guidance to developers to describe their measure concepts. As with the current measure submission form,

---

<sup>12</sup> These difficulties were articulated as part of the feedback on this hand-off process by the CDP Activities team and are included in the section on project tools (as part of the discussion around the measure tracking log) and in the section on scalability of the upfront technical review process.

data were entered into the concept submission form electronically through the usual NQF measure submission interface.

- **Evidence attachment form.** Past experience with the current measure submission form indicates that developers often are unsure how to respond to the items needed for evaluation of the evidence subcriterion and Steering Committee members often have trouble assimilating the information. Staff designed the evidence attachment to elicit this needed information from developers in a way that is more logical and less repetitive compared to how it is elicited in the current measure submission form, with the hope that the new layout would assist developers in addressing the questions and facilitate Steering Committee member understanding of the information provided. For the pilot, the items on this form were not implemented as part of the electronic submission form but were instead distributed as an "attachment" to the form.
- **Measure tracking log.** This tool was created to track and communicate the progress of a concept/measure through the entire evaluation process, from initial submission, through upfront technical review, hand-off to the CDP activities team, and evaluation of the concept/measure by the Steering Committee. Staff used this log to document the feedback and interaction with developers that was provided as part of the upfront technical review process, to track what changes developers made based on this feedback, to record the time needed for the review process, and to flag which concept submission items were the most difficult for developers to answer.
- **Measure testing attachment form.** Similar to the evidence attachment, staff developed a measure testing attachment for use in Stage 2 of the pilot. Although the same information for evaluating scientific acceptability was elicited in this form as compared to the current measure submission form, the questions were more explicitly worded (e.g., what levels of analysis were tested; what is your interpretation of the results).

Feedback on the tools used in the upfront technical review process for Stage 1 was primarily gathered from the CDP activities team and the project Steering Committee. CDP activities team members indicated their belief that, for the most part, the information needed to evaluate measure concepts in Stage 1 of the pilot was contained in the concept submission form and the evidence attachment. However, each member noted that the items on the form itself could have been better organized. In contrast, the Steering Committee noted several additional items that they believed may have helped inform their evaluation of the measure concept (e.g., a brief statement of measure intent, the value of the measure from a patient perspective). They also noted that some of the information that was asked was not necessarily answered optimally (e.g., explicit statements of how the concept links to desired outcomes; information regarding intended use of the measure).

The CDP team also provided some positive feedback about the evidence attachment (e.g., one member described it as "better" than what is used in the current measure submission form). However, team members still noted "confusion" among the Committee with regard to the information contained in the evidence attachment, and suggested that additional "re-structuring" of the form is needed. The team also noted that, even with the current restructuring of the questions regarding evidence on the evidence attachment, some developers were still unable to provide the necessary information.

Although the Steering Committee was asked for specific feedback regarding the evidence attachment itself (e.g., around format, items included, and suggestions for improvement) during the structured discussion, many of their responses revealed questions or concerns about how to evaluate the evidence for measure concepts (e.g., when the only evidence provided is consensus based guidelines/expert opinion; when an exception to the evidence subcriterion should be invoked). Further, in their

responses to the survey, only 78 percent of members agreed or strongly agreed that they understood the criteria for concept evaluation; this result resonates with the feedback of CDP activities team members, who described Committee members as both engaged and eager, but inexperienced and sometimes uncomfortable in using the evaluation criteria. Nonetheless, the Committee did make a few specific suggestions for improving the attachment. For example, they noted some duplication in the attachment and also suggested that a tool to rate the evidence would be useful. They also made suggestions regarding the actual discussion of the evidence, including having a more standardized approach for the discussion and rating of the quantity, quality, and consistency of the evidence and beginning the discussions with the stronger concepts/measures. Finally, Committee members also suggested that, in some cases, there seemed to be misunderstandings by the developers regarding the evidence that is required for NQF approval/endorsement of a concept/measure.

As part of their feedback around the success of the "hand-off" after the upfront technical review process, the CDP team noted that the measure tracking log was useful, yet wasn't completely adequate for communicating whether, and in what ways, developers implemented recommended changes to their submissions that were suggested during the upfront technical review process. Thus, the CDP team reported that additional "re-work" was necessary so as to understand what changes were (or were not) made.

No specific feedback on the concept submission form and evidence attachment elicited from developers who used the forms in Stage 1 of the pilot. However, feedback regarding the measure testing attachment was obtained from developers who are participating in Stage 2 of the pilot. In general, the developers seemed to appreciate the flexibility of the attachment (particularly the ability to include nicely formatted tables and figures). However, this flexibility did cause some problems (e.g., questions inadvertently deleted, inconsistencies between versions of Word). Also, developers viewed the attachment as too long, and several brought up questions about the wording of the questions in the form.

Note that a by-product of improved measure submissions should be increased consistency in Steering Committees evaluations (between measures in a project as well as between measures across projects). However, assessment of this objective was not included as part of the evaluation of the pilot.

### **Were delays in the project avoided?**

Although the pilot actually began several weeks later than originally planned, Stage 1 of the pilot concluded only 10 days after the initially-scheduled completion date and was completed within the 17 weeks originally envisioned. Thus, for Stage 1 of the pilot project, any potential delays in the review and/or endorsement timeline were avoided. However, it is unknown if, or to what extent, the provision of upfront technical review could potentially contribute to the project teams' ability to adhere to the timelines. Notably, staff from both the upfront technical review team and the CDP activities team underscored the inability to implement in the pilot some of the components of the proposed two-stage process that might serve to prevent project delays (i.e., rejecting inadequate submissions and then accepting revised submissions in the next submission cycle).

### **Scalability**

Scalability of the proposed two-stage process—the extent to which the process can be expanded across multiple topic areas given the resources available—is critical to the evaluation of the proposed redesign effort as a whole. While a full evaluation of the scalability of the proposed two-stage CDP is not possible

(given that not all of the components and assumptions of the proposed two-stage CDP could be tested directly through the pilot project), the potential scalability of the upfront technical review process and of the tools developed to facilitate concept and measure reviews was evaluated, particularly as pertaining to the objective of increasing the efficiency of NQF resources.

### Is the upfront technical review process scalable?

For the upfront technical review process to be scalable across multiple concept/measure evaluation, projects staff involved in the review must have the appropriate skills to accomplish the reviews; further, staff must have adequate time for staff to complete the reviews and offer the appropriate feedback.

According to the Technical Review team, the skills needed for adequate upfront technical review include an in-depth understanding of NQF's measure evaluation criteria; however, team members also noted that clinical knowledge and/or experience is necessary. Typically, staff at either the Senior Project Manager level or the Senior Director level would have this skill set.

During Stage 1 of the pilot, 14 of the 18 submitted measures concepts received an upfront technical review and two of these concepts required feedback more than once. Review team members reported a difficulty in balancing the time required for the upfront technical reviews with other work commitments, noting that they considered the reviews to be a priority given the public commitment to provide thoughtful review of the concept submissions.

Although they did not comment specifically on the time/staffing required for the upfront technical review process, the CDP activities team offered feedback regarding the transition from the upfront technical review process to the remainder of the Stage 1 activities. Specifically, they described the hand-off that occurred in the pilot as "fair to good", but noted that quite a bit of "re-work" was required to clarify whether developers had, in fact, implemented the recommendations from the upfront technical review. It is currently unknown to what extent that additional tools/processes would improve this transition.

Some additional quantitative data from Stage 1 support the findings gleaned from interviews of the project staff. In particular, initial data on actual hours versus budgeted hours for Stage 1 of the project indicate that both the upfront technical review process and the remaining CDP activities required more staff hours than were initially expected. It is unknown to what extent staff familiarity with the tools and the process itself might ameliorate this problem.

### Can the tools developed for the GI/GU pilot be used in other projects?

Six tools were developed for the use in the two-stage pilot, four of which have already been described. The two remaining tools include:

- **A Developer Guidebook.** The purpose of this guidebook was to orient concept/measure developers to the proposed two-stage process, to explain the roles and expectations for measure steward/developers during the evaluation process, and to provide in-depth guidance on submitting concepts and measures for evaluation
- **A Steering Committee Guidebook.** The purpose of this guidebook was to orient Steering Committee members to NQF and to the measure evaluation process, to document NQF criteria and guidance for measure evaluation criteria, and to introduce the proposed two-stage CDP.

A notable feature included in both of these guidebooks is an Appendix with examples of "What Good Looks Like" for the importance criteria<sup>13</sup>. The examples included in this section of the guidebooks are meant to illustrate the type of information that would be considered both complete and responsive to the submission items, and thus, ready for Steering Committee evaluation. Some of the examples were adapted from actual measure submissions and some were developed by NQF staff.

Steering Committee feedback on the Steering Committee Guidebook was overwhelmingly positive: 100% of respondents either agreed or strongly agreed that the guidebook was useful in providing information about NQF and the endorsement process. In general, the Committee members also were positive about the section on "What Good Looks Like" (only one respondent disagreed that the examples and key points were useful for evaluating the measure concepts in Stage 1 of the pilot).

Developer feedback on the Developer Guidebook was also quite positive: 30% of individual respondents rated the guidebook as valuable, 60% rated it somewhat valuable, and only 10% (one respondent) rated it as not valuable. However, developers' perceptions of the section on "What Good Looks Like" was somewhat mixed: 50% rated the examples as valuable, 20% rated them as somewhat valuable, and 30% rated them as not valuable. Interestingly, individuals within two of the developer groups had opposing perceptions regarding the value of this section of the guidebook (responses were "valuable" vs. "not valuable" and "somewhat valuable" vs. "not valuable"). Unfortunately, those developers who did not find value in these resources did not provide any comment about why these resources were not useful to them, nor did they provide any suggestions on how to improve the resources.

### Earlier and more balanced opportunities for feedback

As currently designed and implemented in Stage 1 of the pilot, stakeholders have two opportunities to offer feedback on measure concepts: a 15-day commenting period for NQF members prior to the in-person Committee meeting and a 30-day public and member commenting period prior to CSAC review.

A total of 15 comments from 2 member organizations (from the health plan council and the health professionals council) were received during the pre-review comment period. A total of 90 comments from 18 organizations were received during the post-review comment period.

Because a pre-review commenting period is not a feature of the current CDP, the relative response rate for this commenting opportunity during Stage 1 of the pilot cannot be evaluated. However, the response to the post-review commenting period was similar to that of other CDP projects (in the past two years, comments were received from as few as three organizations and as many as 72 organizations, depending on the project; on average, 20 organizations submitted comments during this period).

### More frequent endorsement cycles and more predictable schedules

Because the pilot includes only one full endorsement cycle, and because use of standing steering committees has not yet been implemented, an assessment of these objectives could not be tested directly through the pilot project.

---

<sup>13</sup> An additional resource that includes examples of "What Good Looks Like" for the Scientific Acceptability criteria is still under development and will be provided as a resource to the project Steering Committee for Stage 2 of the pilot.



## Earlier identification and resolution of issues related to competing measures and measure harmonization.

Staff members were able to provide earlier feedback to developers regarding the identification of issues related to competing and related measures (in the current CDP, identification of related or competing measures typically occurs just prior to the in-person Steering Committee meeting). A total of 10 of the concepts submitted to Stage 1 of the pilot had at least one related or competing measure; while developers correctly identified at least some of these related/competing measures for eight of the concepts, NQF staff made the identification for the remaining two concepts.

However, due to the design of the pilot—specifically, the launch of Stage 2 directly after the conclusion of Stage 1—developers did not have the 12-18 month "gap period" allowed in the proposed two-stage design to resolve any issues related to competing and/or related measures.<sup>14</sup> Therefore, an assessment regarding earlier resolution of harmonization/competing measures cannot be accomplished in the pilot project.

## Conclusions Based on Evaluation Results

Several conclusions regarding the success of the pilot itself, as well as the success of the two-stage process as implemented in the pilot, can be made based on the evaluation of the pilot to date.

First, the goals for the GI/GU pilot were, for the most part, met. Many of the process steps of the proposed two-stage process were implemented and evaluated through the pilot project. Also, several new tools needed for implementation were developed and evaluated as part of the pilot. Similarly, even though the pilot project was delayed, and the timelines were somewhat compressed, NQF staff were successful in adhering to the timeline that was developed for completion of Stage 1 of the pilot. Finally, the pilot allowed for testing of staffing requirements for staffing and infrastructure supports, which allows for some assessment of the scalability of the proposed two-stage process.

However, as noted earlier, two key components of the proposed redesign were not implemented in the pilot project. First, standing steering committees were not operational at the time of the pilot. The lack of a standing steering committee structure has hampered the establishment of more predictable endorsement schedules, a key objective of the two-stage CDP. Also, the pilot was designed to allow for only one full measurement cycle (i.e., only one round each of Stage 1 and Stage 2 is included in the pilot). Consequently, measure stewards did not have the 12-18 months allotted in the two-stage design to fully specify and test their measures and then bring them back for the Stage 2 review. Further, the back-to-back design for implementation of Stage 1 and Stage 2 of the pilot impacted the ability of NQF staff to reject incomplete or non-responsive submissions from the pilot. Accordingly, any potential savings of developer resources could not be assessed through the pilot,<sup>15</sup> nor was it possible to assess

---

<sup>14</sup> Instead, for Stage 2 of the pilot, the Committee will provide feedback/recommendations regarding combining or harmonizing measures, and developers will be asked, as necessary, for a joint response regarding plans to address these issues.

<sup>15</sup> Although based on anecdote rather than on formally gathered feedback, NQF staff members involved in the pilot project believe that most of the developers who submitted concepts to the project had already made significant investments in developing their concepts and were too far along in the process to easily make substantial changes. Thus, depending on each individual developer's processes, it is possible that concept review may result in savings of resources that would have been used for measure testing, but not necessarily in savings of resources used for measure specification.

whether the proposed two-stage process would enable more frequent measurement cycles and reduce project delays.

Second, results from the evaluation of the pilot suggest that the proposed two-stage process could enable increased efficiencies for NQF staff and Steering Committees. In Stage 1 of the pilot, one-third of the measure concepts were judged not to meet the Importance criterion, and thus, further review and evaluation against the remaining criteria will not be needed in Stage 2. Typically, Committee members spend the majority of their time on the Importance and Scientific Acceptability criteria; conservatively, therefore, the rejection of one-third of the measure concepts in Stage 1 could translate to a 16 percent decrease in the time needed for complete review of measures by staff and Steering Committees over the life of the project.

However, results from the evaluation of the pilot suggest that the proposed two-stage process may not enable increased efficiencies for the CSAC and Board—and in fact may increase the workload for these groups. Specifically, as designed and implemented in the pilot, there will potentially be a need for a second round of CSAC and Board reviews for two-thirds of the measure concepts that were reviewed in Stage 1 of the pilot (assuming the associated measures meet the remaining evaluation criteria). Although different issues would be considered in this second round, at least some overlap in the two rounds would be expected.

Third, the evaluation results suggest that upfront technical review can help to improve the quality of submissions. However, this finding is tempered by the reality that upfront technical review and associated feedback (for either stage) did not seem to be equally embraced or acted upon by all the developers involved in the pilot.

Fourth, the tools developed for the upfront technical review were found to be beneficial, even though the evaluation uncovered the need for some reorganization of the concept submission form, and potential additions and/or modifications to the evidence and measure testing attachments. The evaluation also uncovered the need for alterations to the measure tracking tool to improve the "hand-off" process between the technical review team and the CDP activities team and minimize the amount of "re-work" associated with the tracking process.

Fifth, there were no substantial delays in the project timeline once the pilot began. However, this result cannot be attributed to the provision of upfront technical review, particularly given that staff members were not able to implement the option to reject consideration of submissions that were deemed incomplete or non-responsive.

Sixth, experience from the pilot project suggests that the upfront technical review process is scalable, at least in terms of having an adequate number of staff members with the skills required to conduct the upfront technical review. However, there is still some question as to whether the process is scalable in terms of the time requirements needed for the process. The pilot project included a relatively small number of measure concepts (a total of 18, rather than the 25-30 measures that are more typically evaluated in a CDP project). Further, the staff time needed to accomplish the reviews was more than was originally anticipated. Yet these findings should be considered in light of the fact that a review process for measure submissions already occurs—albeit on a much more informal basis—in current CDP projects. Furthermore, the time required for upfront technical review could possibly decrease as staff members become more familiar with the process itself and with the tools used in the process—

particularly if the hand-off process from the upfront technical review staff to the remainder of project activities staff is improved.

Seventh, results of the evaluation suggest that the guidebooks developed for the pilot—particularly the sections that give examples of “What Good Looks Like”—are a valuable addition to the educational materials provided to developers and steering committees going forward.

Eighth, the provision of earlier opportunities for feedback was shown to be somewhat successful during Stage 1 of the pilot project. Although only two organizations took advantage of this opportunity, the response indicates that at least some NQF member organizations would be interested in providing earlier feedback. Further, some of the feedback offered related specifically to measure specifications and to concerns around feasibility of data collection—feedback that might be particularly useful to developers who are putting forward true measure concepts and have not yet fully specified the associated measures.

Ninth, the pilot also demonstrated that identification of competing and/or related concepts/measures can be done earlier in the CDP process than is currently the case. However, because the pilot is being implemented for only one full measurement cycle with back-to-back stages, the utility of earlier identification of competing measures and/or harmonization issues—for example, earlier collaboration between developers, effects on subsequent measure development, and potential avoidance of project delays—was not evaluated.

Finally, and perhaps most importantly, the results of the CSAC deliberations at the end of Stage 1 of the pilot provided insight on an unforeseen weakness inherent in separating the evaluation process into stages. Per the design of the two-stage process, the CSAC had the option of disapproving measure concepts, even if approved by the project Steering Committee. In the pilot, two concepts that were approved by the Committee were later overturned by the CSAC, due to concerns with the lack of evidence linking the concepts to outcomes and the lack of importance of these concepts to the overall NQF measure portfolio. Comparatively, this reversal rate was quite high, given that the CSAC has typically overturned Committee decisions only 2-3 times per year.

For each of these two concepts, the project Steering Committee agreed that insufficient information was submitted to rate the quantity, quality, and consistency of body of evidence. However, for one concept, the Committee invoked the evidence exception; for the other concept, the Committee agreed that—although not submitted by the developer—the quantity, quality, and consistency of the body of evidence would meet NQF criteria for evidence. It is possible that the decisions regarding these two concepts were due to the Committee’s unfamiliarity with, or misunderstanding of, the evaluation criteria and process. However, such shortcomings are not unique to the pilot project committee but are present to various degrees in all steering committees. For example, project staff has suggested that some Committee members may have judged those measure concepts more favorably than was warranted by the evidence due to their perceptions of the importance of the topic area. Also, steering committees historically have had difficulties in disentangling the evidence and validity criteria. As currently implemented, the concept review process does not allow for consideration of validity; specifically, there are no criteria under Importance that would allow the Committee to reject a concept when the specifications are not consistent with the evidence presented.<sup>16</sup> Finally, it is possible that

---

<sup>16</sup> The assessment of whether specifications are consistent with the evidence is included as part of the evaluation of the Scientific Acceptability of Measure Properties, which will occur in Stage 2 of the pilot.

some Committee members may have been willing to give those concepts the “benefit of the doubt” at the concept review stage either because they provided recommendations for improving the concepts that must be addressed prior to submission to Stage 2 of the pilot and/or because of the additional opportunity to evaluate the full measure in Stage 2.

## Implications for the Two-Stage Process and the CDP redesign

Several implications regarding the proposed two-stage process and the overall CDP redesign effort can be made based on the results of the evaluation of the pilot project, as follows.

First, while there does seem to be value in allowing for review of measure concepts, the complete redesign of the CDP as a two-stage process to facilitate concept review is not -advised. Many of the measures that will be evaluated in future CDP projects have been previously endorsed (and thus are already fully specified and tested). For stewards of these measures, the two-stage process will actually increase the workload (i.e., two measure submission processes, two in-person meetings, two rounds of responding to comments), but without any of the attendant savings that could come from not having to specify and/or test measures that do not pass the Importance criterion. Similarly, while there would be savings in time and resources for NQF staff, steering committees, the CSAC, and Board when subsequent review against the remaining evaluation criteria in Stage 2 is not necessary, those savings would likely be offset due to the duplicative nature of the processes in the two stages, as well as to the required staff resources needed for the upfront technical review process.

Second, the benefits derived from the formal upfront technical review process do not appear to be a good use of staff time and resources. Although successful to some extent—particularly for developers new to the NQF endorsement process—the desired outcome of higher quality concept submissions did not necessarily materialize as a result of the upfront technical review process. Instead, the success of the process seemed related more to the acceptance of the given feedback rather than to the existence or quality of that feedback.

Third, the information required for the concept review—and the criteria used to evaluate that information—were not sufficient to provide adequate feedback to measure stewards. Specifically, there are no criteria available under Importance to Measure and Report to justify not recommending the concept when the preliminary specifications are not consistent with the evidence. Also, there were no criteria in place that would allow the Committee to offer consistent and binding feedback regarding measure specifications. Finally, , although information regarding the numerator, denominator, and exclusions for the measure concept appeared sufficient to flag potential issues around competing or related measures, it is unknown to what extent these issues can be resolved prior to submission to Stage 2.

**Given the results of the evaluation of the pilot project, NQF will not utilize the two-stage CDP going forward.**

## Considerations for Future CDP Improvement

Based on the findings of the evaluation of the pilot to date, the following suggestions for CDP improvement are presented:

- **Incorporate what worked into the CDP.** Many of the design components from the proposed two-stage CDP were successful and can be incorporated into the existing CDP. These include the following:
  - Earlier opportunity for multi-stakeholder input prior to evaluation of measures by the Steering Committee
  - Use of the Steering Committee and Developer Guidebooks
  - Continued improvement of the measure submission form, incorporating many of the ideas from the Evidence and Measure Testing Attachments
  - More consistent and earlier—but still informal—upfront technical review, using tools and lessons learned from the formalized processes implemented in the pilot
- **Allow for concept review.**
  - Look for ways to allow for concept review at any point in time—both within and outside of scheduled CDPs
  - Do not insist that concept reviews be conducted as part of a formal process
  - Do not require approval for concepts or passing of criteria
  - Do not require concept reviews for fully-specified and tested measures
  - Provide technical assistance and recommendations as part of concept review
- **Continue to look for ways to allow for more timely and flexible review of fully specified and tested measures.**
- **Continue to explore implementation of standing Steering Committees.**
- **Continue to elicit feedback from all stakeholders regarding the CDP.**

## Next Steps

Stage 2 of the GI/GU pilot project is currently underway and will be completed as planned. Based on the evaluation to date, additional information from the second stage of the pilot likely will not change the decision to not move forward with the two-stage process.

A more global redesign of the CDP is under consideration by the NQF Consensus Task Force as part of their charge to review and recommend enhancements for defining and achieving consensus within the CDP. Information on this effort can be found on the Consensus Task Force project page.<sup>17</sup>

---

<sup>17</sup> [http://www.qualityforum.org/Measuring\\_Performance/Improving\\_NQF\\_Process/Consensus\\_Task\\_Force.aspx](http://www.qualityforum.org/Measuring_Performance/Improving_NQF_Process/Consensus_Task_Force.aspx)

## Appendix A: Concepts submitted to the GI/GU pilot

Number	Title	Developer	Type
0030	Urinary Incontinence Management in Older Adults - a. Discussing urinary incontinence, b. Receiving urinary incontinence treatment	NCQA	Maintenance measure
0098	Urinary Incontinence: Assessment, Characterization, and Plan of Care for Urinary Incontinence in Women Aged 65 Years and Older – an administrative measure	NCQA	Maintenance measure
0622	GERD - Upper Gastrointestinal Study in Adults with Alarm Symptoms	ActiveHealth	Maintenance measure
0635	Chronic liver disease - Hepatitis A vaccination	ActiveHealth	Maintenance measure
0658	Endoscopy/polyp surveillance: Appropriate follow-up interval for normal colonoscopy	AMA-PCPI	Maintenance measure
0659	Endoscopy/Poly Surveillance: Colonoscopy Interval for Patients with a History of Adenomatous Polyps- Avoidance of Inappropriate Use	AMA-PCPI	Maintenance measure
C2037	Objective characterization of pelvic organ prolapse, including the bladder and the urethra, prior to surgery	AUGS	Concept
C2038	Performing vaginal apical suspension (uterosacral, iliococcygeus, sacrospinous or sacral colpopexy) at the time of hysterectomy to address uterovaginal prolapse	AUGS	Concept
C2049	Complete Workup for Assessment of Stress Urinary Incontinence Prior to Surgery	AUA	Concept
C2050	Patient counseling on treatment options, including behavioral and surgical treatments prior to SUI surgery	AUA	Concept
C2051	Patients Counseled About Risks Associated with the Use of Mesh in Sling Surgery Prior to Surgery	AUA	Concept
C2052	Reduction of Complications through the use of Cystoscopy during Surgery for Stress Urinary Incontinence	AUA	Concept
C2054	Assessment of treatment within one year of SUI surgery	AUA	Concept
C2056	Colonoscopy Quality Index	Quality Quest for Health of Illinois	New measure
C2059	IBD preventive care: corticosteroid sparing therapy	AGA	Concept
C2062	IBD preventive care: corticosteroid related iatrogenic injury – bone loss assessment	AGA	Concept
C2063	Appropriate use of cystoscopy in pelvic prolapse repair	AUGS	Concept
C2065	GI Hemorrhage Mortality Rate (IQI #18)	AHRQ	New measure

## Appendix B: Data Collection Instruments

### CDP ACTIVITIES TEAM STRUCTURED DISCUSSION QUESTIONS, STAGE 1

*This qualitative review with members of the CDP team is intended to elicit feedback on the overall experience, perceptions of the value of the process, quality of submitted measures, and assessment of resources. Data will be collected through 30-minute individual interviews with each team member and one 45-minute team discussion.*

**As part of the pilot for the redesigned CDP, how would you describe your overall experience on the GI/GU CDP Team?**

- Excellent
- Very good
- Good
- Fair
- Poor

**COMMENTS:**

**What was your involvement during the technical review process? How would you describe the hand-off between the technical review team and your team?**

- Excellent
- Very good
- Good
- Fair
- Poor

**Why? COMMENTS:**

**Do you feel the technical review process resulted in the availability of better/more complete information for the Steering Committee's consideration?**

- Yes
- No

**Why? COMMENTS:**

**From your perspective, was the information collected to describe the concept (i.e., numerator (and details), denominator (and details), exclusions, usability info, taxonomy sufficient for the Steering Committee to evaluate the concept?**

- Yes
- No

**Why? COMMENTS:**

**Is there additional information that should be collected to better help the Committee evaluate importance of the concept?**

- Yes
- No

**What information? COMMENTS:**

**From your perspective, was the evidence attachment clear and easy for the Steering Committee to understand?**

- Yes
- No

**COMMENTS:**

**From your perspective, did you have sufficient resources (staffing and/or time) to adequately perform your responsibilities related to this phase of the project?**

- Yes
- No

**If not, what additional resources would you have needed? COMMENTS:**

**ADDITIONAL COMMENTS:**

**Thank you! We appreciate your feedback.**



**UPFRONT TECHNICAL REVIEW TEAM  
STRUCTURED DISCUSSION QUESTIONS, STAGE 1**

*This qualitative review with members of the technical review team is intended to elicit feedback on the overall experience, perceptions of the value of the process, quality of submitted measures, and assessment of resources. Data will be collected through 30-minute individual interviews with each team member and one 45-minute team discussion.*

**Overall, how would you describe your experience on the Technical Review team?**

- Excellent
- Very good
- Good
- Fair
- Poor

**COMMENTS:**

**From your perspective, did you have sufficient resources (staffing and/or time) to adequately perform your reviews?**

- Yes
- No

If not, what additional resources would you have needed?

**ADDITIONAL COMMENTS:**

**What skill set do you think is necessary to perform a thorough technical review? Could this function be performed by any member of the Performance Measures team (i.e. project analyst, project manager, senior project manager, senior director)?**

**For Level 1 review?**

**For Level 2 review?**

**Generally, were measure developers receptive to the feedback provided by the Technical Review team?**

- Very receptive
- Receptive
- Somewhat receptive
- Not receptive at all

**COMMENTS:**

**Do you think the pre-submission conference calls were valuable resources for developers?**

- Yes
- No

**COMMENTS:**

**Overall, how would you rate the value of the entire NQF Technical Review process?**

- Very valuable
- Valuable
- Somewhat valuable
- Not valuable at all

**COMMENTS:**

**OTHER COMMENTS ON THE TECHNICAL REVIEW PROCESS:**

**Thank you! We appreciate your feedback.**

## **MEASURE DEVELOPERS SURVEY QUESTIONS, STAGE 1**

The purpose of the technical review period is to provide measure developers with in-depth feedback on the completeness and responsiveness of their measure submissions to the GI/GU project prior to submission. In order to gauge the effectiveness of the technical review period, NQF will examine the reaction to the technical assistance itself from all those involved both internally and externally. Those measure developers receiving technical assistance will be asked to give this assistance an overall rating and will be asked more specifics such as whether they found this assistance to be useful.

**Please rate the value of the following components and or resources of the Technical Review period:**

Rating Scale

- Valuable
- Somewhat valuable
- Not valuable
- Did not use technical assistance

1. NQF Technical Review period pre-submission conference call.
2. NQF Developer Guidebook.
3. Examples of “what good looks like” in the NQF Developer Guidebook.
4. NQF Technical review period feedback on concept submission.
5. Overall value of entire technical review process.
  
6. Did you revise your concept submission based on the suggested feedback received during the technical review period? (Y/N)
7. Would you take advantage of technical review if offered (not required) on a future project? (Yes/Maybe/No) (Comment box)
8. Please provide comments and suggestions for enhancements to the technical review period. (Comment box)

**Thank you! We appreciate your feedback.**

**STEERING COMMITTEE  
SURVEY QUESTIONS, STAGE 1**

**Have you served on an NQF Steering Committee and performed measure evaluations before?**

- Yes
- No

**Please indicate to what degree you agree with the following statements:**

**When I arrived at the meeting I understood how to evaluate the concepts against the “Importance to Measure and Report” criteria.**

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Unsure

**COMMENTS:**

**The Steering Committee Guidebook was useful in providing information about NQF and the endorsement process.**

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Unsure
- I did not look at the Steering Committee Guidebook

**COMMENTS:**

**The examples and key points in the Steering Committee Guidebook were useful for evaluating the concepts.**

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Unsure

**COMMENTS:**

**The concept submission forms were complete (no blanks where information was needed) and answered the questions appropriately.**

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Unsure

**COMMENTS:**

**The SharePoint site was easy to use to find needed documents and evaluation materials.**

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Unsure

**COMMENTS:**

**OTHER COMMENTS ON THE EVALUATION PROCESS:**

**Thank you! We appreciate your feedback.**

## MEASURE DEVELOPERS TELEPHONE SURVEY QUESTIONS, STAGE 2

### *Email Script*

Good [Morning/Afternoon],

Thank you for participating in the Gastrointestinal/Genitourinary (GI/GU) Stage 2 technical review period. The purpose of the technical review period is to provide measure developers with in-depth feedback on the completeness and responsiveness to NQF criteria of their measure prior to submission to the [Gastrointestinal/Genitourinary \(GI/GU\) Measure Endorsement project](#). In order to gauge the effectiveness of the Stage 2 technical review period, NQF is requesting feedback about the technical review from those developers who participated. The technical review period occurred during the entire submission of approved concepts phase. Measure developers that received technical review are being asked to participate in a 30-minute conference call to discuss whether they found this assistance to be useful. During the call, a few questions will also focus on the measure testing form. We have attached a copy of the form for your convenience. Please select one of the time slots listed below for the conference call. If none of these slots work for you, please send us your availability. Please send us your selected time slot by COB \_\_\_\_\_.

Thank you for your time,

### *Telephone Script*

Hello, \_\_\_\_\_ [your colleague's name] and I would like to thank you for taking time out of your busy schedule to provide us with your feedback on the GI/GU Stage 2 technical review period. The technical review period occurred during the entire submission of approved concepts phase. The purpose of the technical review period is to provide measure developers with in-depth feedback on the completeness and responsiveness to NQF criteria of their measure prior to measure submission. We will use the feedback you provide to us on today's call to evaluate this component of the pilot.

1. What parts of the technical review period, the written itemized feedback or the post-technical review submission conference call, were most valuable?
2. Did you revise your measure submission based on the suggested feedback received during the technical review period? (Yes/No—why or why not)
3. Would you take advantage of technical review if offered (not required) on a future project? (Yes/No—why or why not)
4. Would you like to provide any additional comments regarding the technical review period? Do you have any suggestions on how to enhance the technical review period?

### *Telephone Script*

Now we would like to ask you a few questions regarding the measure testing form. The measure testing form, a separate document from the measure submission form, was designed to allow measure

developers to submit information on measure testing with the option to include illustrations, such as tables, graphs and diagrams. Modifications were made to some questions to offer a more clear understanding of the type of information NQF is seeking for testing.

5. Were there specific items on the measure testing form that were unclear, hard to understand or difficult to answer based on your testing information?
6. Was the format of the measure testing form conducive to completing the form? (Yes/No—why or why not)
7. Would you like to provide any comments or suggestion for improving the measure testing form?

*Telephone Script*

We would like to thank you for taking time to provide us with your feedback. We appreciate your thoughts and will use your feedback as a part of the evaluation of this element of the process.

## Appendix C: NQF Staff

**Helen Burstin, MD, MPH**

Senior Vice President  
Performance Measures

**Heidi Bossley, MSN, MBA**

Vice President  
Performance Measures

**Helen Imbernino, BSN, MPH**

Director, Quality Management and Compliance  
Operations

**Taroon Amin, MA, MPH**

Senior Director  
Performance Measures

**Karen Johnson, MS**

Senior Director  
Performance Measures

**Karen Pace, PhD, RN**

Senior Director  
Performance Measures

**Elisa Munthali, MPH**

Senior Project Manager  
Performance Measures

**Alexis Morgan, MPH**

Senior Project Manager  
Performance Measures

**Ashlie Wilbon, RN, MPH**

Senior Project Manager  
Performance Measures

**Evan M. Williamson, MPH, MS**

Project Analyst  
Performance Measures



