NATIONAL QUALITY FORUM
Driving measurable health improvements together

Primary Care and Chronic Illness Fall 2021 Cycle: Public and Member Comments

Contents

Post-Evaluation Measure-Specific Comments on Primary Care and Chronic Illness Fall 2021 Submissions

NQF #3332 Psychosocial Screening Using the Pediatric Symptom Checklist-Tool (PSC-Tool) (Recommended)

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood
Comment ID#: 7971 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in opposition to measures 3661 as outlined below. UnityPoint Health is one of the nation's most integrated health care systems. Through more than 32,000 employees and our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. Within our organization, a handful of physicians utilize PSC17 but don't find it as helpful as other screenings. Our physicians find it too long to be useful as a screening form in a well child visit and deliver results that are too vague to be useful in a focused mental health encounter. The pediatric symptom checklist is an even longer version (35 questions) plagued by the same problems. Typically, basic history taking indicates whether a more specific and sensitive screen or diagnostic tool is indicated or in the setting of a well visit whether a follow up visit dedicated specifically to mental health is needed. A standard set of psychosocial screening questions is certainly helpful but if the end point of those questions is to just indicate the need for additional assessment, as is the case with the PSC, then those first questions need to be very brief, and that additional assessment deserves dedicated time outside of a well visit. Furthermore, identifying more kids with problems without any infrastructure to support them will be stressful for providers and not helpful to our patients. At this time UnityPoint Health feels this measure, as proposed, is too cumbersome for use in daily practice and would not recommend the measure move forward.

Developer Response

Writing on behalf of UnityPoint Health, Stephanie Collingwood has commented that, in their opinion, NQF measure #3332, "Psychosocial screening with the Pediatric Symptom Checklist Tool (PSC-Tool") should not move forward for endorsement because it is "too cumbersome for use in daily practice". Our reply is that although this may be UnityPoint's opinion, they do not support it with anything but the claim that within their organization, "a handful of physicians utilize PSC17 but don't find it as helpful as other screenings". Beyond this assertion based on very small number of unspecified cases, the commentator does not reference any studies that would support this opinion. Contrast this with a recent paper (Murphy et al, 2020) that showed that in the first year of a best practice commitment to using the PSC-17 in a network of 18 suburban outpatient pediatric practices whose patients were covered predominantly with commercial insurance, 89.3% of the

35,237 well child visits in their organization had been screened with the PSC and that even in the second year of the program, the rate was 77.9%. Other evidence of the feasibility (lack of cumbersomeness) of using the PSC in actual practice are reports from the statewide Massachusetts Children's Behavioral Health Initiative for children with Medicaid (CBHI; Kuhlthau, 2011; Murphy et al, 2020) now in its fifteenth year. The PSC is the primary screening measure for 4–17-year-old children and, according to data on a state website, the rate of psychosocial screening has averaged about 67% over the entire time and has not dropped below 60% (MassHealth Quarterly Screening Report, 2021 [use https://www.mass.gov/info-details/childrens-behavioral-health-initiative-cbhi-data-reports]) since it started, again providing strong evidence for feasibility/lack of cumbersomeness. The UnityPoint Health commentator goes on to note that: Our physicians find it too long to be useful as a screening form in a well child visit and deliver results that are too vague to be useful in a focused mental health encounter. The pediatric symptom checklist is an even longer version (35 questions) plagued by the same problems Again the commentator provides no empirical support for the opinion expressed on this point and, in contrast, the CBHI program website and findings from the empirical papers cited above (and more than 200 other papers that used the PSC) document the feasibility of using the PSC in the real world. With regard to the comment about the length of the PSC-35, it may be important to note that even after fifteen years, the state of Massachusetts continues to approve the use of the PSC-35 as well as the PSC-17 (MassHealth - Learn about the Approved Screening Tools [use https://www.mass.gov/info-details/learn-about-the-approved-masshealth-screening-tools]).To our knowledge, there are no published studies of the UnityPoint approach. The commentator next provides a snapshot of the approach that UnityPoint uses: Typically, basic history taking indicates whether a more specific and sensitive screen or diagnostic tool is indicated or in the setting of a well visit whether a follow up visit dedicated specifically to mental health is needed. A standard set of psychosocial screening questions is certainly helpful but if the end point of those questions is to just indicate the need for additional assessment, as is the case with the PSC, then those first questions need to be very brief, and that additional assessment deserves dedicated time outside of a well visit. Although the brief outline of the UnityPoint system described in the comment may sound like a less cumbersome approach, the comment does not describe in any detail its alternative approach (which would permit a reader to assess whether it does indeed seem less cumbersome). Again, there no evidence of the feasibility and effectiveness of an alternative system, and no data comparing the approach outlined by UnityPoint to a PSC-based system. As for the comment's implication that the PSC takes too long and is not brief enough, it is important to note that all versions of the PSC are filled out prior to the well child visit, so there is no burden on pediatricians at all and even the burden on the parent or youth who completes the PSC is very light. The PSC-35 takes only about 5 minutes to complete, and the PSC-17 takes only two minutes. It is hard to see how, even if the implied briefer UnityPoint screen took only one minute, this time saving would be experienced as significantly less cumbersome. Although the commentator does not give enough details for the reader to be sure, it sounds like their approach may actually be to do away with first stage screening entirely, skipping right to diagnosis-specific measures if the pediatrician becomes aware of a specific problem (e.g. child seems anxious during the well child visit). This kind of approach flies in the face of dozens of studies over several decades showing that, in the absence of a policy that endorses routine psychosocial screening, pediatricians often fail to detect problems during the well child visit and therefore would lose the chance to administer diagnosis-specific measures. All of this is not to imply that current PSC screening systems are perfect or that other approaches which might prove

to be more effective might not exist now or in the future. Quite the contrary, the literature shows that there are other ways to screen, although systematic comparisons usually favor the PSC (Pourat et al, 2017). For example, the second most frequently used brief psychosocial screen is the Strengths and Difficulties Questionnaire, which, as its name suggests, includes an assessment of strengths as well as problems. The SDQ's authors believed that adding some questions about strengths would make the SDQ superior to a questionnaire like the PSC that only screened for difficulties…but there are, to our knowledge, no studies that investigate this hypothesis. The SDQ was originally endorsed by Massachusetts CBHI but was removed after several years due to its infrequent usage (MassHealth - Learn about the Approved Screening Tools [use https://www.mass.gov/info-details/learn-about-the-approved-masshealth-screening-tools]). Other researchers have explored whether greater screening accuracy can be obtained by using the PSC in tandem with a second brief screen (like the PHQ-9; Jellinek et al, 2021), with longer and shorter screens like the CBCL and SDQ (Young and Takala, 2018), or when both parent and youth complete a PSC (Montano, 2011). Although each of these screening alternatives undoubtedly has some plusses, they undoubtedly also have some minuses and there has been no empirical research that demonstrates a more effective way to screen in the real world than by using a single PSC measure. The UnityPoint comment concludes by stating that "…identifying more kids with problems without any infrastructure to support them will be stressful for providers and not helpful to our patients." While, again, this assertion might be true in general, it confounds a number of issues. Although identifying children and adolescents with psychosocial problems without providing ways to respond to them is unethical as well as pointless, the published research provides evidence that this is not what happens in the real world. For example, numerous studies by Hacker and her associates (2014a, 2014b, and 2016) have documented that after the implementation of the CBHI screening program, thousands of additional referrals were made and there were substantial increases in the number of children and adolescents who actually received outpatient mental health services. Screening mandates do not in and of themselves guarantee a greater access to mental health services, but they do appear to create pressures within healthcare systems to find ways to help the children and adolescents who are newly identified. In conclusion, although the UnityPoint commentator has not provided evidence that should lead NQF to withdraw its endorsement of the PSC, the UnityPoint comment outlines its own alternative approach which might, over time, be able to provide evidence of its feasibility and effectiveness so that it could be compared to similar evidence from the PSC. Until that time, NQF's continued endorsement of psychosocial screening with the PSC-tool will keep encouraging providers to use a screen with proven feasibility and effectiveness and, thereby, to facilitate research that can sharpen a more complete understanding of the most important aspects of routine screening in pediatrics. References: Hacker KA, Penfold R, Arsenault L, Zhang F, Murphy M, Wissow LS. Screening for behavioral health issues in children enrolled in Massachusetts Medicaid. Pediatrics. 2014;133(1):46-54. Hacker KA, Penfold RB, Arsenault LN, Zhang F, Murphy M, Wissow LS. Behavioral health services following implementation of screening in Massachusetts Medicaid children. Pediatrics. 2014;134(4):737-746. Hacker KA, Penfold R, Arsenault LN, Zhang F, Soumerai SB, Wissow LS. The Impact of the Massachusetts Behavioral Health Child Screening Policy on Service Utilization. Psychiatric Services. 2016;68(1):25-32. Jellinek M, Bergmann P, Holcomb JM, et al. Recognizing Adolescent Depression with Parent-and Youth-Report Screens in Pediatric Primary Care. The Journal of Pediatrics. 2021;233:220-226. Kuhlthau K, Jellinek M, White G, VanCleave J, Simons J, Murphy M. Increases in behavioral health screening in pediatric care for Massachusetts Medicaid patients. Archives of Pediatrics &

Adolescent Medicine. 2011;165(7):660-664. MassHealth. Quarterly Behavioral Health Screening Report: Behavioral health screening January 2012-2020. Hingham, MA: Executive Office of Health and Human Services. 2021. https://www.mass.gov/info-details/childrens-behavioral-health-initiative-cbhi-data-reports MassHealth. Learn about the approved MassHealth screening tools. Hingham, MA. 2022. https://www.mass.gov/info-details/learn-about-the-approved-masshealth-screening-tools. Montano Z, Mahrer NE, Nager AL, Claudius I, Gold JI. Assessing psychosocial impairment in the pediatric emergency department: Child/caregiver concordance. Journal of Child and Family Studies. 2011;20(4):473-477. Murphy JM, Stepanian S, Riobueno-Naylor A, et al. Implementation of an electronic approach to psychosocial screening in a network of pediatric practices. Academic Pediatrics. 2021;21(4):702-709. Murphy JM, Riobueno-Naylor A, Haile H, et al. Behavioural Health Screening and Service Use in a Statewide Sample of Medicaid-Eligible Pediatric Outpatients. Science Repository. 2020;doi:10.31487/j.PDR.2020.03.04 Pourat N, Zima B, Marti A, Lee C. California child mental health performance outcomes system: Recommendation report. UCLA Center for Health Policy Research. 2017. Young ND, & Takala CR. Sequential screening to improve behavioral health needs detection in primary care. Journal of the American Academy of Child & Adolescent Psychiatry. 2018;57(8):603-609.

NQF Response

N/A

NQF Committee Response

N/A

NQF #3661 Mismatch Repair (MMR) or Microsatellite Instability (MSI) Biomarker Testing Status in Colorectal Carcinoma, Endometrial, Gastroesophageal, or Small Bowel Carcinoma (Recommended)

Anna Kim, American Geriatrics Society; Submitted by Anna Kim

Comment ID#: 7958 (Submitted: 04/25/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

The AGS does not support Measure #3661 : Mismatch Repair (MMR) or Microsatellite Instability (MSI) Biomarker Testing Status in Colorectal Carcinoma, Endometrial, Gastroesophageal, or Small Bowel Carcinoma. The guidelines recommend testing for patients with concern of familial cancers and we were not able to find clinical data that suggests outcomes will improve if the recommendation is broadened to all patients. Furthermore, the analysis was only on an individual basis without support from a group-level analysis.

Developer Response

Thank you for your comments. To clarify, we are not suggesting that every patient is a candidate for MMR or MSI testing. However, recent guidelines broaden recommendations beyond familial cancers to include patients being considered for checkpoint inhibitor therapy (see https://www.cap.org/protocols-and-guidelines/cap-guidelines/current-cap-guidelines/mismatch-

repair-and-microsatellite-instability-testing-for-immune-checkpoint-inhibitor-therapy). This is the reason for the Exception category "patients not a candidate for checkpoint inhibitor therapy". With the FDA's approval of pembrolizumab for any advanced tumor that is microsatellite instable or mismatch repair deficient, it is increasingly important to consider not only familial occurrences of these genetic changes such as those found in Lynch syndrome but spontaneous as well. We also appreciate the concern regarding individual vs group level analysis. As noted by NQF staff, this was addressed to the satisfaction of the reviewers. However, we continue to collect data on this measure (which was in use in 2021 and is in use in 2022) at the clinician and group level so that further testing can be performed to ensure complete reliability.

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee accepted the updated guideline which was submitted by the developer to support the broadening of the measure population to assist in therapeutic decision making and thus to include all patients being considered for checkpoint inhibitor therapy. Further, the Standing Committee evaluated the measure as specified by the College of American Pathologists, with the level of analysis at the group/practice level and individual level. At the meeting, the developer stated that the analysis results at the individual level demonstrated sufficient reliability and that aggregating at the group level would only improve the reliability. The Standing Committee accepted this rationale and found reliability testing sufficient for both the individual and group levels.

Leslie Narramore, American Gastroenterological Association; Submitted by Leslie Narramore

Comment ID#: 7952 (Submitted: 04/07/2022)

Council / Public: Public

Level of Support: N/A

Comment

AGA supports NQF Measure 3661. Mismatch repair (MMR) and microsatellite instability (MSI) are key biomarkers in colorectal cancer (CRC) and other GI tumors, with crucial diagnostic, prognostic, and predictive implications. Gastroenterologists and other clinicians order testing for MMR/MSI during screening for Lynch syndrome and/or prognostic stratification for patients with CRC or with a personal history of colon and rectal cancer. Gastroenterologists and other ordering clinicians depend on pathologists' interpretations of and any recommendations for tests in order to provide quality patient care. If the status of genetic testing is not indicated in each pathology report, important tests may be missed, or unnecessary repeat testing may be performed leading to inappropriate treatment and/or increasing cost. Having a quality measure would provide a strict framework for management with the multi-specialty team managing the patient's oncology care. This is a measure that is applicable to several specialties and fits the larger paradigm of cross-cutting measure, which are particularly relevant. Measure 3661 represents a crucial step in the care process by promoting effective communication of critical information for the purpose of care coordination and efficient use of resources.

Developer Response
N/A

NQF Response
Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response
N/A

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7970 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment
UnityPoint Health respectfully offers comments in support of measures 3661. UnityPoint Health is one of the nation's most integrated health care systems. Through more than 32,000 employees and our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. It was noted that the tumor type for the numerator should specify "adenocarcinoma" or include an additional exclusion for "neuroendocrine carcinoma." Though uncommon, MMR testing is not currently indicated for neuroendocrine carcinomas. From an operational perspective, concern was raised regarding the ability to capture results of MMR testing performed on the biopsy when a resection is received. Sometimes the biopsy is read at a different institution and there may not be an efficient mechanism to determine whether MMR testing was performed and what the results were. However, the metric is written broadly enough that it would be satisfied by mentioning the need to correlate with biopsy MMR testing or recommending that MMR testing be requested if not performed on the biopsy.

Developer Response
Thank you for your comments. We appreciate the careful consideration of the details of the measure specification and will consider whether an exclusion is needed for neuroendocrine carcinoma in the future. At the moment, scientific evidence does not definitely rule out MSI testing on poorly differentiated neuroendocrine colorectal carcinoma, so we did not exclude this from the measure completely (see 2019 ESMO recommendations found here: https://www.annalsofoncology.org/article/S0923-7534(19)31269-4/fulltext). However, we will continue to engage with stakeholders and monitor scientific consensus to determine whether additional clarification is needed. We also appreciate the difficulty in determining whether MMR testing was previously performed on a biopsy. As noted, we wrote the measure broadly to account for such circumstances and to discourage repeat testing by allowing "recommended" or "previously performed" as Met conditions.

NQF Response
N/A

NQF Committee Response
N/A

NQF #3667 Days at Home for Patients with Complex, Chronic Conditions (Not Recommended)

Anna Kim, American Geriatrics Society; Submitted by Anna Kim
Comment ID#: 7957 (Submitted: 04/25/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

The AGS supports this measure and believes it is important to managing patients at home, particularly as it is patient-centered and one that patients care about deeply. It is critical to capture patients with substantial disease(s) and have specific measures for complex, chronic conditions to do so effectively. The measure is also increasingly being used in scientific literature as a valid composite outcome measure. While the AGS agrees that the issues of risk adjustment and incorporating social determinants of health (SDOH) are crucial, these challenges are not unique to this specific measure. We encourage efforts to improve measures by improving risk adjustment of SDOH.

Developer Response
N/A

NQF Response

Thank you for your comment. This measure was submitted as outcome measure not a composite outcome measure. The Standing Committee reviewed the measure as it was submitted. Your comment has been shared with the Standing Committee and the measure developer

NQF Committee Response
N/A

Dr. Clarke Ross, DPA, American Association on Health and Disability
Comment ID#: 7955 (Submitted: 04/22/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

The American Association on Health and Disability, Altarum, and the Lakeshore Foundation appreciate the opportunity to provide comments. We write to support the measure recommended by the NQF committee. A core tenet of the disability rights movement, enshrined in the Americans with Disabilities Act (ADA) and L.C. v. Olmstead, is that people with disabilities of all ages have a

right to receive services and supports in the most integrated setting, regardless of the source of payment for services or the intensity of their service needs. Most people far prefer to age in their homes, and research has shown that individuals who receive needed services in their communities – including individuals with the most complex intellectual disabilities who require the most substantial supports -- experience improved quality of life. The Consortium for Citizens with Disabilities (CCD) and the Disability and Aging Collaborative (DAC) address the services and supports that enable older adults and individuals with disabilities of all ages to live in their homes and communities. We are CCD and DAC members, and Altarum is also a DAC member. In particular, these coalitions focus on the Medicaid Home and Community-Based Services (HCBS) program, recognizing that HCBS is the key to community integration, full participation, independent living, and economic self-sufficiency for many people with disabilities and older adults. These critical services make it possible for people with disabilities and older adults to fully exercise their civil and human rights. NQF Measure #3667 is an innovative provider group-level measure of days at home or in a community setting. It is stewarded by CMS and is a Yale Center for Outcomes Research and Evaluation (CORE) measure. The proposed measure is focused on Medicare fee-for-service (FFS) beneficiaries with complex chronic conditions, and the level of analysis is the Accountable Care Organization. This measure is coming forward at a key moment, as the U.S. health care system moves further toward provision of multiple services in home and community-based settings (other than for specialty care in hospitals and medical centers). NQF Measure #3667 therefore has outsize public policy significance, given that it is poised to set an important precedent for analyzing provider performance in the context of person-centered care. The disability and aging communities have been promoting HCBS services and programs for decades as desired alternatives to institutional settings and as strategies for "rebalancing" Medicaid and other public program financing away from institutions. To fully realize these approaches requires dissemination and use of a meaningful quality metric that measures how providers fare in keeping their patients and clients out of medical institutions. Measure #3667 now being considered for use through Medicare and ACOs, is a clear recognition that the larger health care system is moving to meaningfully promote the objectives of home and community living. We therefore urge NQF to approve the measure, and to move work forward that will measure what matters to millions of people who need and want their medical care to help them return home as soon as possible, and to remain there for as long as possible, with appropriate support services if required. In closing, the disability and aging advocacy and research communities are committed to supporting quality measurement experts whose work aims to expand funding and programming for care in HCBS settings. We are heartened to see Congressional proposals being considered that would further incentivize HCBS services and supports, and we believe that if approved, NQF Measure #3667 would strengthen and reinforce these trends over time.

Developer Response

N/A

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee concluded that the developer's approach to risk adjustment was not sufficient. Therefore, the Standing Committee did not pass the measure on validity; a must pass criterion.

Dr. Clarke Ross, DPA, American Association on Health and Disability

Comment ID#: 7956 (Submitted: 04/22/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

Comments - part 2 - info on 3 submitting organizations: The American Association on Health and Disability, Altarum, and the Lakeshore Foundation appreciate the opportunity to provide comments. We write to support the measure recommended by the NQF committee. The American Association on Health and Disability (AAHD) (www.aahd.us) is a national non-profit organization of public health professionals, both practitioners and academics, with a primary concern for persons with disabilities. The AAHD mission is to advance health promotion and wellness initiatives for persons with disabilities. AAHD is specifically dedicated to integrating public health and disability into the overall public health agenda. The Lakeshore Foundation (www.lakeshore.org) mission is to enable people with physical disability and chronic health conditions to lead healthy, active, and independent lifestyles through physical activity, sport, recreation and research. Lakeshore is a U.S. Olympic and Paralympic Training Site; the UAB/Lakeshore Research Collaborative is a world-class research program in physical activity, health promotion and disability linking Lakeshore's programs with the University of Alabama, Birmingham's research expertise. Altarum is a nonprofit health services research organization (www.altarum.org) that helps federal and state health agencies and foundations improve health equity and outcomes through better systems of care, primarily for disenfranchised populations. Altarum strives to produce solutions that go beyond being road maps for improvement; rather they serve to catalyze, accelerate and implement innovations.

Developer Response

N/A

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A

Kyle Bagshaw, Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (CORE)

Comment ID#: 7972 (Submitted: 04/25/2022)

Council / Public: Public

Level of Support: N/A

Comment

As developer of NQF#3667 "Days at Home for Patients with Complex, Chronic Conditions" (hereafter "Days at Home"), the Yale New Haven Health Services Corporation Center for Outcomes Research & Evaluation (CORE) was disappointed by the Standing Committee's decision not to support endorsement of the measure. We are submitting this comment to bring attention to and provide corrections in response to important instances of mischaracterization of the measures and testing provided both in the "Primary Care and Chronic Illness Standing Committee Measure Evaluation Web Meeting Summary" published for public comment and the information provided to the Standing Committee before their review meeting. First, we were concerned to note mischaracterizations of the measure and the measure testing that were presented in the draft "Primary Care and Chronic Illness Standing Committee Measure Evaluation Web Meeting Summary" document, published for the current Public Comment period. We respectfully request that NQF correct the following items in the final version of the report: 1. The summary states that "the construct validity testing found that NQF #3667 did not correlate well with the other measures. The developer emphasized that the lack of correlation may be due to the other measures having smaller sample sizes and not being risk-adjusted." o This statement is incorrect. CORE tested Days at Home against six conceptually related measures and found modest to high correlation with four of the six as anticipated; the hypotheses for lack of correlation were only applicable to two measures which were narrowly defined and not risk adjusted. Overall, this testing demonstrated that the Days at Home did correlate well with other measures, supporting construct validity. o We also note that these results showing construct validity, together with the face-validity established by unanimous support from our large and diverse Technical Expert Panel, demonstrate that the Days at Home measure meets NQF criteria for validity testing of new measures. 2. The summary states: "The Standing Committee expressed concerns about social determinants of health (SDOH) factors not being included in the risk adjustment model. The developer noted that there is no national, standardized approach to address SDOH factors, and the small sample size hindered the developer's ability to account for SDOH factors. Thus, the developer decided to utilize dual eligibility as an alternative to SDOH in the risk adjustment model." o The measure is adjusted for beneficiaries' dual-eligible status. Dual-eligible status is included not as an alternative to SDOH, it is a conceptually valid indicator of low income and lack of wealth and was the most significant SDOH factor identified in measure testing. o Dual-eligible status is the preferred SDOH factor by the Office of the Assistant Secretary for Planning and Evaluation (ASPE) as detailed in their December 2016 Report to Congress. While imperfect, this indicator represents lack of income and wealth in a way that correlates highly with other issues. o It is also not correct that "small sample size hindered the developer's ability to account for SDOH factors." The approach to SDOH testing was constrained by the lack of relevant and reliable data available at the patient level. Nonetheless, the measure was tested using data that are currently available at a geographic level to consider other factors for inclusion in the measure; these factors proved to have low predictive value in measure testing and were ultimately excluded for that reason. Thank you for considering these corrections for the final Fall 2021 Cycle report for the PCCI Standing Committee. In addition, we are concerned that some errors in the information packet provided to the Standing Committee prior to their meeting may have contributed to some misunderstanding of the measure. For the accuracy of future information about the measure, we would like to provide the following clarifications in particular: • The measure does not count "days after death occurs" as days in care. • The measure does not

exclude long-term nursing home residents; current residents are considered to be "at home" and eligible for subsequent days in care. • The decision not to exclude care in some settings (such as emergency department visits) and count these settings toward "days in care" was made in order to reflect the priorities and preferences of patients. While there may be individual cases in which a "day in care" is preferable to a "day at home," the Technical Expert Panel unanimously supported this broad conception of "days in care," noting that a measure called "days at home" would lack face validity if any care in an inpatient setting was defined to be "at home" and agreed that in aggregate counting these settings would be inappropriate. Finally, CORE is very appreciative of the thoughtful consideration of the Standing Committee. We note, however, that at times the Committee's discussion did not provide clear indications of how their concerns could be addressed. For example, some Committee members noted they would have liked to see comparisons to other measures of care coordination; however, currently no such measures exist, and it is unclear how the measure developer could meet such a request. Similarly, some Committee members would have liked to see testing for different social risk factors but acknowledged the lack of availability of data elements. Given the substantial time and resources that go into measure development and testing, we request that in the post-comment discussion the Committee clarify more concretely what modifications or feasible future testing would address concerns about the measure's validity so we can plan for future phases of measure testing and evaluation.

Developer Response

N/A

NQF Response

Thank you for your comment. NQF will make the appropriate adjustments to the draft report.

NQF Committee Response

Thank you for your comment. The Standing Committee would like to provide the following recommendations for the measure developer to consider.

1. Introducing a survey instrument or a patient-reported outcome measure that would assess factors, which may affect the quality of care and feasibility of care being provided at home.
2. Focus assessment of the measure on the continuum of care versus location of care (i.e., home).
3. Dual-eligibility risk identifier is not an accurate capture of SDOH factors. Not all patients who are able to receive care at home are duel-eligible and this could penalize the provider. Additionally, there are significant policy variations in Medicaid from state-to-state, which impact entity-level SDOH factors.

The Standing Committee maintains its decision to not recommend the measure for endorsement, based on the measure failing to pass the validity criterion.

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 7953 (Submitted: 04/20/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

Comment

The American Medical Association agrees with the concerns raised by the Standing Committee on this measure, particularly around the validity of the measure. We support the Committee's recommendation to not endorse the measure at this time.

Developer Response

N/A

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A

Ms. Tilithia McBride

Comment ID#: 8007 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

The Federation of American Hospitals (FAH) agrees with the concerns raised by the Standing Committee on this measure, particularly around its validity. We support the Committee's recommendation to not endorse the measure at this time.

Developer Response

N/A

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF Committee Response

N/A

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7969 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in support of measures 3667. UnityPoint Health is one of the nation's most integrated health care systems. Through more than 32,000 employees and

our relationships with more than 480 physician clinics, 40 hospitals in urban and rural communities and 14 home health agencies throughout our 9 regions, UnityPoint Health provides care throughout Iowa, central Illinois, and southern Wisconsin. On an annual basis, UnityPoint Health hospitals, clinics and home health provide a full range of coordinated care to patients and families through more than 8.4 million patient visits. UnityPoint Health is supportive of this measure; however, challenges exist today in operationalizing. Market variations exist regionally with a metric like this. For patient outcomes, it's all about access. If there are no care at home options, then patients may have no choice but to go to the emergency department/hospital. For example, if a patient needs an IV diuresis for heart failure, while optimal to offer within the home, this isn't always an option, particularly with RN staffing shortages. Another concern for providers is that it's frustrating to have metrics where there's a perceived or actual lack of control. For example, we have numerous patients in our hospitals that are there for weeks or even months while they wait for an open bed at a mental health facility or long-term care facility. Lack of bed access is entirely out of a provider's control. Claim base measures bring their own challenges as data is delayed, in some cases up to six months, making reactive action less effective. This is where population health management needs to become stronger and align with a global value base strategy.

Developer Response

- UnityPoint Health Comment: UnityPoint Health is supportive of this measure; however, challenges exist today in operationalizing. Market variations exist regionally with a metric like this. For patient outcomes, it's all about access. If there are no care at home options, then patients may have no choice but to go to the emergency department/hospital. For example, if a patient needs an IV diuresis for heart failure, while optimal to offer within the home, this isn't always an option, particularly with RN staffing shortages. o Developer Response: We appreciate your support of this concept and your thoughtful consideration of the measure. We discussed the issue of regional differences in patient access to services extensively with our Technical Expert Panel and acknowledge this as a concern for some providers. However, we have not found that any providers are systematically disadvantaged in performance on the measure as a result. During testing for potential risk factors, we found that urban residence and local density (per 100,000 population) of hospital beds were not significantly associated with patients' days in care. Greater local density of primary care physicians and specialists was associated with fewer days in care, but the practical magnitude of this effect was quite small compared to that of clinical risk factors and dual-eligibility. Conversely, greater local density of nursing home beds was associated with more days in care, but the practical magnitude of this effect was also quite small. o Furthermore, the population-based focus and broad outcome of this measure is intended in part to allow flexibility and promote innovation to meet the goal of reducing the use of acute inpatient care utilization across their patients, in recognition that there is no one-size-fits-all approach for every provider group's situation. • UnityPoint Health Comment: Another concern for providers is that it's frustrating to have metrics where there's a perceived or actual lack of control. For example, we have numerous patients in our hospitals that are there for weeks or even months while they wait for an open bed at a mental health facility or long-term care facility. Lack of bed access is entirely out of a provider's control. o Developer Response: We acknowledge that some factors contributing

to days at home are outside of providers' ability to control. Accordingly, the goal for the measure is not to eliminate "days in care" entirely but to encourage providers to explore other options when feasible, as one piece of a larger quality strategy. Furthermore, the measure is intended for organizations like ACOs that provide comprehensive services to patients across the continuum of care and so have more opportunities to engage with patients both to mitigate the risk of health deterioration leading to hospitalization and to organize care to provide for needed outpatient services. • UnityPoint Comment: Claim base measures bring their own challenges as data is delayed, in some cases up to six months, making reactive action less effective. o Developer Response: Unfortunately, it is true that claims-based measures will have some delay in providing feedback. The reporting delay associated with Days at Home is comparable to that of many other claims-based measures in current use. • UnityPoint Comment: This is where population health management needs to become stronger and align with a global value base strategy. o Developer Response: We agree that promoting good population health management is a key strategy. We hope that introducing this measure will put a spotlight on this issue and highlight further opportunities to improve care, outcomes and experiences of patients.

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee concluded that the developer's approach to risk adjustment was not sufficient. Therefore, the Standing Committee did not pass the measure on validity; a must pass criterion.

Public Comments on Primary Care and Chronic Illness Fall 2021 Draft Report

N/A

Pre-Evaluation Measure-Specific Comments on Primary Care and Chronic Illness Fall 2021 Submissions

NQF #3661 Mismatch Repair (MMR) or Microsatellite Instability (MSI) Biomarker Testing Status in Colorectal Carcinoma, Endometrial, Gastroesophageal, or Small Bowel Carcinoma (Recommended)

Colleen Skau, College of American Pathologists ; Submitted by Ms. Colleen Skau, PhD

Comment ID#: 7844 (Submitted: 01/12/2022)

Council / Public: HPR

Level of Support: N/A

Comment

For reliability testing, the CAP only performed testing at the individual level. This was for two reasons: first, since the testing we did (signal to noise with a beta-binomial model) is dependent on the number of measured entities, we started with the testing that would yield the lower reliability value, which is testing at the individual level. Given that our individual-level reliability was very high, we did not proceed to group level testing. Second, for purposes of MIPS reporting (which is the only program this measure is for), the group score is simply the sum of the individual scores, there is no separate method of calculating a group score. So calculating "group" reliability doesn't have an independent meaning.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

NQF #3667 Days at Home for Patients with Complex, Chronic Conditions (Not Recommended)

Jake Miller

Comment ID#: 7845 (Submitted: 01/12/2022)

Council / Public: Public

Level of Support: N/A

Comment

Yale/CORE clarifications to the methods panel evaluation summary (1 of 2):

Specifications:

- *In their preliminary analyses, a few SMP members found the specifications confusing and occasionally arbitrary. Some members expressed concerns about the potential misalignment of concept presentations within the submission and noted the denominator statement appeared to lack an explanation of the target population, conditions, settings, and other pertinent measure constructs information. They were also concerned that several concepts included in the submission were not documented as exclusions in the*

specifications, which both threatens the measure's validity and may incentivize under-treatment of conditions potentially outside the locus of control of the accountable entity, including very low outliers that can never reach the expected performance gains, permanent nursing home residents **NQF Clarification:** Please note that the issues noted here were raised by some but not all SMP members and that the summary should clearly reflect these as individual opinions, not the consensus of the entire SMP. Over 60% of subgroup members voted to support this measure on both reliability and validity in the preliminary analysis, indicating they were able to follow the information we provided in the submission. In the final vote after the SMP discussion of these issues, 4 of 10 SMP members still supported the measure validity and voted to pass the measure. It is important not to base this summary solely on the views of a few individual SMP members. **Clarification:** The Days at Home measure is population-based and intended to capture performance broadly across eligible beneficiaries. The target population is patients with complex, chronic conditions (who have higher risk for needing complex care) as defined by the inclusion criteria. **This is clearly documented in the submission and should not be noted as lacking**. There are intentionally no denominator exclusions – all beneficiaries meeting the inclusion criteria are included in the denominator because conceptually all are at risk for days in care, and any further exclusions would lack face validity. Some members of the committee may have been confusing the cohort (included beneficiaries) and outcome (days in care that count in the model). We clarify the outcome below. However, it is not accurate to present the measure as "not documenting exclusions." **Clarification**: The description of the SMP evaluation seems to reference comments related both to the cohort of included patients (as addressed above) and in the outcome definition of days in care (as clarified here). The measure uses a broad definition of "days in care" consistent with feedback from the Technical Expert Panel (TEP) and aligned with previous work by the Medicare Payment Advisory Commission (MedPAC), reflecting that patients tend to view any time in settings such as inpatient hospitals and facilities as disruptive to their daily life. The consensus recommendation of the TEP was to maintain a broad conception of days in care, so that no types of hospital admission were counted as "days at home." Such a broad definition is not intended to suggest every admission is avoidable, but instead to represent a patient-centered outcome definition which allows for flexibility in improvement strategies. The goal is not to achieve zero days in care, but to reduce the total days in care compared to expectation for a given case mix. **Clarification**: It is not accurate to say that "very low outliers" or "permanent nursing home residents" are categorically "outside the locus of control of the accountable entity." Clinical groups and ACOs do have capacity to impact days in acute care for these populations (for example, through more proactive preventive care and improved care coordination to avoid preventable admissions) as confirmed by the TEP. • *The SMP also questioned whether the consideration of exclusions included (i.e., patients treated in emergency departments, admitted to acute care settings, and days after a death occurs), indicated low-quality care. Another SMP member expressed concerns with adjusting for transitions to the nursing home, which purports that moving from home to a nursing home, is always negative. Other concerning date elements included permanent nursing home admissions requiring skilled nursing care, which may include personal and community resources that are not be modifiable by the accountable entity.* **Clarification:** As noted above, the Days at Home measure does not conceptually assume that *all* days in an

included setting indicate low-quality care, and the goal is not to achieve zero days in these settings. Rather, the goal is to encourage providers to explore home-based options or other feasible means so that their patients can spend *fewer* days in these settings. Moreover, days after a death occurs are not counted as either days at home or days of acute care, but rather as unmeasured days. **Clarification:** The goal of adjusting for nursing home transitions is to encourage providers to explore care options, such as providing home-and community-based care, preventive care services, or improved care coordination, that relieve some of the burden on their patients (and family/caregivers) while allowing patients to remain in their home and community longer. While in some cases a transition to nursing home is the best outcome for a patient, the TEP and CMS agreed this outcome is more often less desirable than remaining in the community setting and that the measure should not have the unintended consequence of rewarding providers who are quicker to transition patients to nursing homes. The adjustment is designed to have a modest effect on measure scores in those cases where there are much higher rates of transition than expected given the case-mix of patients. The current approach was developed as a compromise between counting days in a nursing home as "acute care days" and counting them as "days at home," both options that include notable drawbacks as discussed by the TEP. **Clarification:** While most long-term nursing home residence days are considered "days at home," days in which **skilled** nursing care is utilized *do* count as "days in care." • *SMP members also noted that the unit of analysis reported in the measure vacillates between accountable care organizations (ACOs) and provider group.* **Clarification**: The measure is intended for use in different settings in which accountable entities comprise groups of individual providers, including provider groups and ACOs; the specifications have used the general term "provider group" to capture these different organizations. The term "ACO" is used only in documentation pertinent to the *testing* of the measure, which used a dataset of 2017-2018 Shared Savings Programs ACOs and aligned beneficiaries. • *One SMP member questioned whether this measure, which combines multiple risk models calculations into a single overall score, should be considered a cost composite measure.* **Clarification:** Days at Home is not a composite measure; it measures a single outcome. The mortality and nursing home transition component models are not standalone measures, nor are they intended to capture different outcomes. These component models are included as a means of safeguarding against potential adverse consequences for the measure that were identified in conversation with CMS, the TEP, and other experts. The only outcome is days at home, which is adjusted for multiple risk factors, as well as for unexpectedly high mortality or nursing home transition rates. This is demonstrated empirically in test results as noted in the additional comments in the final measure submission; the quality signal of the measure is dominated by the Days in Care component and the additional adjustments result in modest changes for a small number of ACOs. **Validity** • *The developer conducted construct validity with Pearson correlations to six other ACO-level measures hypothesizing that quality conceptually relates to excess days in care (EDIC) for patients with complex chronic diseases.*

- *Pearson's correlations did not correlate well, ranging between -0.549 and +0.048 resulting in a high inverse correlation for unplanned admissions (expected), moderate correlation with other measures, no correlation with fall risk, and an unexpected inverse correlation with patient experience.*

- *The developer explained that this is possibly due to endogeny of the hospital admissions and readmissions measures. The developer also reported the poor correlations may result from testing against measures using smaller sample sizes and which were not risk adjusted for clinical variables.* ***Clarification:*** *This summary does not accurately reflect the developer's explanation. We documented the expected modest correlations in a direction that was pre-specified. The measures with significant correlation in the expected direction have key and notable differences in cohort (the patients included and the time period for measurement) and outcome (the settings included and the outcome metric) from Days at Home, despite some overlap. These measures were intended to assess construct validity because they measure similar aspects of quality in distinctly different way. These results do not undermine the validity of the measure as we would expect similar results across providers between similar measures.*

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

Jake Miller

Comment ID#: 7846 (Submitted: 01/12/2022)

Council / Public: Public

Level of Support: N/A

Comment

Yale/CORE clarifications to the methods panel evaluation summary (2 of 2): Risk-Adjustment

The SMP members had concerns with the model construction, which they agreed lacked vital adjustment and consideration for many variables without theoretical or empirical justifications and used arbitrary measure weighting. The developers acknowledge these were not empirically assessed, but rather are subjective and based solely on TEP recommendation. **Clarification:** The Days in Care statistical count model includes an offset for days alive, so that "mortality days" are not counted in either the numerator nor denominator of the main measure component ("excess days in care" or EDIC). The Days in Care measure does incorporate an adjustment to EDIC for the excess mortality risk of the measured provider groups, as well as the excess risk of transition to nursing home. These adjustments are made by multiplying the EDIC by a standard mortality ratio (SMR) and by 0.5 times a standard nursing home transition ratio (SNHR). The SNHR is scaled to have the same distribution as SMR and then given a relative weight of 0.5, to accommodate feedback received from the TEP that nursing home transition as an outcome is less severe than death but should still be reflected in performance scores. Both the SMR and SNHR adjustments have a minor impact on the overall score, except in the case of extreme differences from the average provider group risk of mortality or nursing home transition. **Clarification:** The "nursing home start date of January 1" refers to the classification of beneficiaries; those already in a nursing home on January 1 are not considered for a nursing home transition during the measurement period. This start date aligns exactly with the specified performance period for the measure of January 1 to December 31 (the calendar year). *o A few SMP members discussed the effect of*

specific chronic conditions on the risk model, such as cancer, dementia, and congestive heart failure that increase EDIC by nature of the disease states. **Clarification:** The measure includes risk adjustment to account for differences in case mix between providers, including for these stated factors. While these conditions may result in more observed (unadjusted) days in care for patients, risk adjustment accounts for this increased risk and these patients will not necessarily have more *excess* days in care. *o The greatest concern for the risk adjustment model expressed from the SMP members was the development approach for days at home, and the mortality and nursing models. The SMP noted that formulas in the approach may include doubling the EDIC estimates for enrolled ACOs and negative impacts to the penalty schematic* **Clarification:** It is unclear what "faulty formulas" are being referenced here, what "doubling" is described, or how the specifications compromise the validity of the measure. The formulas used were endorsed by the TEP, which included members with expertise in measure development who had reviewed the approach and results in great detail. Performance on the measure is driven by the Days in Care model, which is a conventional risk adjustment model. The score is then modified such that only provider groups with both outlying performance in Days in Care and nursing home transitions and/or mortality are noticeably impacted. It is not true that this results in "doubling the estimates" for some providers. It is also not clear what "negative impacts to the penalty schematic" means in this statement or what "fault" in the specifications is proposed to give rise to that. Without more detail, it is difficult to further address the challenges being put forward. **Exclusions** *o The SMP questioned the process-outcome pathway that resulted in increased, rather than decreased, days in care, and the lack of exclusions for long-term nursing home residents prior to a measurement period, who have no chance of "at home" days defined in the specifications.* **Clarification:** This is not an accurate description of the methodology. **Patients who reside in long-term nursing homes are considered "at home" for purposes of the Days in Care model.** For example, a nursing home resident on January 1 with no other care use during the year would be considered "at home" for the full 365 days. Similarly, for patients who transition to a nursing home during the performance year, all subsequent days in the nursing home with no other care use are counted as "days at home." *o SMP members indicated the discrimination and calibration were generally acceptable but had concerns related to the low outliers. The developer described this as an unintended consequence of the measure construct as the measure attempts to balance days at home with other unintended consequences.* **Clarification**: The measure does not have a strict definition of outliers, nor is it proposed to report outliers. In clarification of results the SMP may be referring to, certain ACOs observed in testing with scores much lower than average did not arise as a result of "attempting to balance days at home with other intended consequences." These ACOs in the test dataset already had substantially more days in care than expected, based on the Days in Care model results even before accounting for nursing home transitions and mortality, and their low performance is unrelated to the additional adjustments. The nursing home and mortality adjustments simply have the greatest *potential* impact for provider groups that are already outliers (either high or low) in Excess Days in Care. The measure was designed to ensure that it is extremely difficult for a provider group with near-average Excess Days in Care to become a very high or very low performer due solely to outlying performance in the nursing home or mortality models. **Meaningful Differences** *o A few SMP members questioned the presence of meaningful differences in performance and the use of the measure for quality improvement purposes, and whether the measure could be used to identify differences in patient function or health-related quality of life.* **Clarification:** While scores are reported as "days at home" to align with the conceptual focus of the measure, *differences* in

performance should be considered relative to days *in care* which are the basis of the main Days in Care model. As noted in the measure submission, the interquartile range of 3.0 days at home (329.1 – 332.1) reflects that patients of a provider at the 25th percentile of performance can each expect to spend on average 3.0 days *more* in care than they could expect at a provider at the 75th percentile of performance. As the average patient in the cohort spends 12.8 days in care, 3 days more or fewer represents a meaningful amount of time for each patient who, as noted above, strongly prefer to minimize time in these care settings when possible.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 7847 (Submitted: 01/12/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on #3667, Days at Home for Patients with Complex, Chronic Conditions. We note that while the submission form indicates that the measure is intended to be used at the Accountable Care Organization (ACO) level, the wording, "provider groups", is used frequently throughout the submission. We request clarification on whether the measure is intended to be used for ACO reporting only or if it would also be applied to other levels such as clinician groups. Based on the specifications and testing completed, we do not believe that it would be appropriate to be applied to any other level but the submission is not clear on its intent. In addition, The AMA strongly supports the inclusion of individuals with dual eligibility status in the risk model but remains concerned that CMS continues to test social risk factors **after** the assessment of clinical and demographic risk factors and it is unclear why this multi-step approach is preferable. On review of the Evaluation of the NQF Trial Period for Risk Adjustment for Social Risk Factors report, it is clear that the approaches to testing these data should be revised to strategies such as multi-level models or testing of social factors prior to clinical factors and that as access to new data becomes available, it may elucidate more differences that are unrelated to factors within an entity's control. Additional testing that evaluates clinical and social risk factors at the same time or social prior to clinical variables rather than the current approach with clinical factors prioritized should be completed. References: National Quality Forum. Evaluation of the NQF Trial period for Risk Adjustment for Social Risk Factors. Final report. July 18, 2017. Available at:

http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=85635. Last accessed January 8, 2022.

NQF Response

Thank you for your comment. It has been shared with the Standing Committee and the measure developer.