NATIONAL QUALITY FORUM
Driving measurable health improvements together

Geriatrics and Palliative Care Fall 2021 Cycle: Public and Member Comments

Contents

https://www.qualityforum.org
1099 14th Street NW, Suite 500 | Washington, DC 20005 | M 202.783.1300 F 202.783.3434

Post-Evaluation Measure-Specific Comments on Geriatrics and Palliative Care Fall 2021 Submissions

NQF #3645 Hospice Visits in the Last Days of Life (Recommended)

Anna Kim, American Geriatrics Society; Submitted by Anna Kim
Comment ID#: 7959 (Submitted: 04/25/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

The AGS supports Measure # 3645: Hospice Visits in the Last Days of Life. While hospice visits in the last days of life are not necessary for all patients, the overwhelming majority and their families need support in the last days, which are the hardest for both patients and families. We believe that hospice visits are critically important during 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

Katie Wehri, Katie Wehri
Comment ID#: 7989 (Submitted: 04/29/2022)

Council / Public: Public

Level of Support: N/A

Comment

Since 1982, the National Association for Home Care & Hospice (NAHC) has been the leading association representing the interests of hospice, home health, and home care providers across the nation. Members are providers of all sizes and types -- from small rural agencies to large national companies -- and including government-based providers, nonprofit voluntary hospices, privately-owned companies and public corporations. As such, we welcome the opportunity to comment on NQF# 3645 - Hospice Visits in the Last Days of Life (HVLDL). The Geriatrics and Palliative Care Standing Committee evaluated and voted on this measure in its March 2022 meeting. NAHC recommends to the Committee that it reconsider this vote for the following reasons: • NAHC understands the reason for the focus on hospice visits delivered during the final days of life; however, we strongly urge the Committee to consider visit data in the context of an individualized plan of care reflective of patient and family wishes. The Meaningful Measure area for the Hospice Visits in Last Days of Life measure is “Person Centered Care” and the Healthcare Priority is

"Strengthen Person & Family Engagement as Partners in their Care" We believe that, as currently constructed, the proposed measure does not fit within this area or priority and NQF should not endorse it. Utilizing data that includes whether the patient/family desired a visit from the disciplines that are part of a measure or exclusion criteria that removes patients/caregivers who refuse visits offered by these various disciplines in the last days of life from the measure denominator would better reflect quality of care and better serve the Meaningful Measure and Healthcare Priority. • Care of the imminently dying patient is an important domain of palliative and hospice care. NAHC further appreciates the individualized plan of care based on patient and family needs and desires and the value of visits and services delivered by all members of the hospice interdisciplinary group (IDG) consistent with such plan of care. The core services of hospice care include a full complement of disciplines – physician, registered nurse, and medical social worker, patoral or other counselor. These disciplines are recognized as the core of hospice care because they address pain and symptoms that occur at the physical, emotional and spiritual level. This is the essence of hospice care. Therefore, the services provided by all core members of the IDG should be included in any visit measure. Also of note is the possibility that the majority of patients/families distinguish hospice staff visits by type, i.e. social worker or nurse, chaplain or aide, but do not distinguish further. Specifically, CMS should consider the possibility that patients/families do not distinguish between an RN and LPN but, rather, simply recognize that a "nurse" is making or made a visit. Of course, credentials of the individual making the visit are present on a nametag, but this is often not scrutinized by patients/families once they know the individual and, after time, the LPN versus RN license is forgotten. At a minimum, CMS should consider inclusion of all nursing visits, RN and LPN, in the HVLDL measure. • The 2020 Abt Associates report, Hospice Quality Reporting Program Quality Measures and Assessment Instruments Development, Modification and Maintenance, and Quality Reporting Oversight Support, provides background information on the development of the HVLDL. Unfortunately, the NQF Geriatrics and Palliative Care Committee was not presented with this report or the details of some of the data from the report prior to voting. Specifically, the data shows positive correlation between RN and MSW visits and CAHPS Hospice Survey outcomes although it is a low correlation. The correlation with chaplain and aide visits is also low. Granted, RN and MSW visit correlations are higher, but these visits are in the same correlation category as chaplain and aide visits per the data. None of the visit types had a moderate (0.5-0.7) or strong (0.7-1) correlation with the CAHPS Hospice Survey outcomes. It was mentioned by Committee members during the March 2022 review that the data does not align with the experience of some Committee members in that visits from other disciplines, specifically chaplain, are valued by patients and families. In fact, in the initial review of the measure the Committee suggested that the measure could be further strengthened by expanding the care disciplines covered, conducting a more holistic review of patient and caregiver end of life desires, and including postmortem visits and pediatric palliative care hospice patients. Hospice visit data and its correlation to the CAHPS hospice survey results should be further analyzed, including analysis that incorporates visits in the context of the individualized plan of care and the patient's wishes regarding visits. If warranted based on this data, CMS should expand any visit data utilized in the HQRP to include all core disciplines. • Claims-based visit data, from which the HVLDL is calculated, do not provide a true picture of hospice services delivered to Medicare hospice beneficiaries. This is because they do not include telehealth visits. Some patients and families prefer this type of visit especially in the final days of life. NAHC strongly recommends that telehealth visits be included in the HVLDL as these visits are indicators of care and services provided by the hospice and are

related to and ordered on the plan of care. An identifier for telehealth visits on hospice claims would allow CMS to capture this data, and NAHC and other hospice stakeholders have urged CMS to create a code or other identifier for this purpose. • The Hospice Outcome & Patient Evaluation (HOPE) instrument, currently in the beta testing phase, will capture data as hospice care is being delivered to patients, a gap in the HQRP that CMS sought to close in recent years. The amount of data and information available not only to consumers but also to the measure steward, CMS, and hospice providers from the HQRP is relatively small. The HOPE will bring significantly more data and information to the HQRP that will allow for more robust quality measures. It is anticipated that the HOPE will be in use soon by hospices. NAHC strongly recommends the impact of the measures anticipated from the HOPE on visits in last days of life be considered, to eliminate any possible future duplication. We appreciate the opportunity to submit these comments and strongly urge the Committee to reconsider its vote on NQF #4635 – Hospice Visits in the Last Days of Life. Sincerely, Katie Wehri Director of Home Health & Hospice Regulatory Affairs

Developer Response

Thank you for your comments regarding Hospice Visits in the Last Days of Life (HVLDL). we appreciate your thoughtful and input, and we have prepared response addressing the important issues you raised. We are grateful that the intent of the measure is understood. We were also happy that the measure's performance met all NQF criteria for variability, validity, and reliability, and was recommended for endorsement. We welcome the opportunity to address the issues raised. Visits by professional hospice staff - registered nurses and social workers - have been cited in focus groups as being particular helpful in the last days of life by bereaved family. Such attestations led CMS to incentivize visits by these staff, only (and not the full IDG team) in the Service Intensity Add-On policy implemented in 2016. Subsequently in development of HVLDL, CMS conducted a per-discipline analysis comparing the receipt of visits with the hospices' CAHPS outcome scores. Visits by registered nurses and social workers were the only two disciplines which yielded a meaningful positive correlation. A previously developed measure, Hospice Visits When Death is Imminent (HVWDII), encompassed a broader array of the disciplines of the IDG. This measure, encompassing the full IDG, failed to meet NQF testing standards, directly resulting from poor validity evidence (i.e., no relation to CAHPS scores), as detailed in a report CMS has published on its website since 2020 (https://www.cms.gov/files/document/hqrphospice-visits-when-death-imminent-testing-re-specification-reportoctober-2020.pdf). Based on our data analysis, we believe another measure broadly encompassing the full IDG team would similarly fail as was the case with HVWDII. CMS re-specified HVWDII as HVLDL, which meets testing criteria, and is moreover calculated using claims data, important information already collected by providers; CMS would be negligent to not publicly report this information, which we have shown to provide value to the Hospice Quality Reporting Program. It should be noted the evidence for chaplain visits was mixed - that is, the additional inclusion of chaplain visits may meet NQF testing standards and bring demonstrated value to the HQRP. However, at present chaplain visits are not captured by claims data. CMS believes HVLDL which focuses on RN/SW visits, only, brings meaningful value to the HQRP., and the lack of chaplain visits should not prevent the public receive otherwise useful data. We appreciate the commenter's note to consider the HOPE data as a source of chaplain visits in the future. The commenter notes that end-of-life visits may not occur due to refusals. CMS had implicitly allowed for refusals during measure design, by specifying the measure to counts visits in two of the last three days or life, instead of visits on each of last three days. Also, CMS believes there is value to a broad, population-based measures. CMS certainly expects that caregiver refusals

of visits will occur - and indeed family wishes of privacy near death a of paramount to be respected - and scores are not expected to ever be 100%. But basic analyses demonstrate there is important variation across hospices, more so than could plausibly be explained by differences in patient refusals across hospices. CMS believes this variation reveals meaningful differences in care delivery that could be useful to patients and their families when making a choice about the type of provider from whom they wish to receive care. The commenter raised the issue of telehealth. While we appreciate the comment, the steward at this time intends to keep the measure as specified, with in-person visits being the focus. CMS is proud of the new HOPE instrument currently in development, which will collect more information on hospice quality of care and will greatly enhance what is currently reported in the Hospice Quality Reporting Program. However, HOPE has not yet been nationally implemented, and no data has been collected, so it will be some time before measures from national HOPE data can be publicly reported. CMS has claims data on hand right now and would be remiss to not report this useful information. Patients and families making a difficult decision during an emotional time need assistance now and HVLDL will assist to help healthcare consumers make an informed choice.

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee notes that there is value in monitoring the quality of care provided by registered nurses and social workers during the last-days-of-life. While the Standing Committee recognizes the concern that certain disciplines are excluded from the measure, the Standing Committee maintains that the measure meets NQF criteria as specified and stands by the decision to recommend the measure for endorsement. However, the Standing Committee encourages the developer to monitor data and billing codes, as they become available, to support the inclusion of other interdisciplinary groups (e.g., chaplains, licensed practical nurses) within future iterations of the measure. The Standing Committee also recommended that the developer consider returning for early NQF maintenance review, prior to the designated three years, if including additional disciplines becomes more feasible.

Marian Grant, C-TAC; Submitted by Dr. Marian Grant, DNP, RN

Comment ID#: 7979 (Submitted: 04/27/2022)

Council / Public: CON

Level of Support: N/A

Comment

C-TAC supports this measure

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

Mr. George Handzo, Rev., BCC, CSSBB, HealthCare Chaplaincy Network
Comment ID#: 7986 (Submitted: 04/29/2022)

Council / Public: Public

Level of Support: N/A

Comment

The HealthCare Chaplaincy Network appreciates CMS' goal of developing a measure that seeks to improve the quality of care currently being provided to beneficiaries and their families at the end-of-life. We welcome the opportunity to comment on this measure. There is clear evidence that the last days of life for a hospice patient can be times of high symptom burden. Both patients and family caregivers can have high needs in all domains of care. If the patient and family desire a visit from hospice staff, it is certainly within the hospice's responsibility to deliver it. The evidence suggests that many hospices may not be meeting this goal reliably and there appears justification for a quality measure to help improve this situation. However, the measure itself should be of high quality and patient-centered. The Meaningful Measure area for the Hospice Visits in Last Days of Life measure is "Person Centered Care" and the Healthcare Priority is "Strengthen Person & Family Engagement as Partners in their Care". We believe that, as currently constructed and without modification, the proposed measure does not fit within this priority and NQF should not endorse it. Arguably, the central tenet of the patient and family-centered care that CMS promises to deliver to all beneficiaries is to enable the care that the patient and family want when they want it and not impose care on patients that they do not desire. This measure as currently constructed incentivizes hospices to impose RN and social worker visits on patients and families who may not want them and where a visit from a different member of the hospice interdisciplinary group may be more appropriate. Furthermore, the HVLDL measures does not allow the virtual telehealth visits that some families prefer and discourages hospices from meeting the end-of-life spiritual and religious needs of patients and families by focusing heavily on only the medical portion of the interdisciplinary team. This measure does not serve the Meaningful Measure or Healthcare Priority as intended. It is our strong belief that analysis utilizing data that includes whether the patient/family desired a visit from the disciplines that are part of a measure or exclusion criteria that removes patients/caregivers who refuse visits offered by these various disciplines in the last days of life from the measure denominator would better reflect quality of care and better serve the Meaningful Measure and Healthcare Priority. Collecting and monitoring data of visits in the last days of life is understandable, and the Coalition strongly urges CMS to consider visit data in the context of an individualized plan of care reflective of patient and family wishes. Furthermore, we suggest a clear claims-based indicator that outlines spiritual care visits. This provision will accurately serve the purpose of reducing administrative and reporting burden that this measure seems to bring forward. The literature is clear that spiritual needs are prevalent at the end of life and rise in importance as death approaches. There are indications that spiritual well-being buffers the potential impact of end-of-life despair. Most religious traditions also have important end-of-life rituals. The National Consensus Project Clinical Guidelines for Quality Palliative Care propose that

maximizing the benefits of supporting spiritual and religious needs requires a trained professional health care chaplain-a spiritual care specialist. The current measure is effectively a barrier to the delivery of that spiritual/religious care by not permitting chaplain visits to be included in this measure, even if this is what the family chooses and desires at the end of life. There seems to be a major disconnect between the stated goal of the measure and the research used to describe it. The developer states the goal of the measure as follows: “Collecting information about hospice staff visits for measuring quality of care will encourage hospices to visit patients and caregivers and provide services that will address their care needs and improve quality of life during the patients' last days of life.” It would seem that the goal is to have as many patients as possible and caregivers who want a visit to be visited so that their needs can be addressed. We completely agree with this goal. However, there is no discussion here of the quality of the visits-only the fact that those visits were made. The developers justify the measure based on patient satisfaction scores on the CAHPS. In the process they limit the hospice’s incentive to deliver visits by limiting the disciplines (RN and social workers only) whose visits count (excluding physicians, LPNs, aides and chaplains) and excluding virtual visits which many patients might prefer. It seems that the developers have conflated two different goals-increasing the number of patients visited and meeting patient needs on one hand and measuring family satisfaction with the visits that are made on the other. We suggest they focus on the former and remove all restrictions on the members of the hospice staff whose visits count for the purposes of this measure. The 2020 Abt Associates report, Hospice Quality Reporting Program Quality Measures and Assessment Instruments Development, Modification and Maintenance, and Quality Reporting Oversight Support, provides background information on the development of the HVLDL. Unfortunately, the NQF Geriatrics and Palliative Care Committee was not presented with this report or the details of some of the data from the report prior to voting. Specifically, the data shows positive correlation between RN and MSW visits and CAHPS Hospice Survey outcomes although it is a low correlation. The correlation with chaplain and aide visits is also low. Granted, RN and MSW visit correlations are higher, but these visits are in the same correlation category as chaplain and aide visits per the data. None of the visit types had a moderate (0.5-0.7) or strong (0.7-1) correlation with the CAHPS Hospice Survey outcomes. It was mentioned by Committee members during the March 2022 review that the data does not align with the experience of some Committee members in that visits from other disciplines, specifically chaplain, are valued by patients and families. In fact, in the initial review of the measure the Committee suggested that the measure could be further strengthened by expanding the care disciplines covered, conducting a more holistic review of patient and caregiver end of life desires, and including postmortem visits and pediatric palliative care hospice patients. The Hospice Outcome & Patient Evaluation (HOPE) instrument, currently in the beta testing phase, will capture data as hospice care is being delivered to patients, a gap in the HQRP that CMS sought to close in recent years. The amount of data and information available not only to consumers but also to CMS and hospice providers from the HQRP is relatively small. The HOPE will bring significantly more data and information to the HQRP that will allow for more robust quality measures. It is anticipated that the HOPE will be in use soon by hospices. The Coalition urges CMS to consider the impact of the measures anticipated from the HOPE on visits in last days of life, to eliminate any possible future duplication. In sum, this measure is much needed however, it must be of the highest quality itself. The measure can reach that status with three simple changes: 1) Removing the restrictions on the disciplines of the staff whose visits count, 2) allowing virtual visits, and 3) inserting an exception to the denominator for patients and families who are documented to not want a visit of any kind at

end of life (last 3 days). Therefore, until such time as appropriate changes can be made, HCCN recommends the Committee reconsider its vote to endorse the HVLDL measure. HealthCare Chaplaincy Network Since its founding in 1961, HealthCare Chaplaincy Network (HCCN) has led the way in the integration of spiritual care in health care through clinical practice, education, research, and advocacy. The organization has grown from a small program providing hospital chaplaincy in the New York metropolitan area into an internationally recognized model for multi-faith spiritual care, education, and research. The parent company of the Spiritual Care Association (SCA) and the SCA University of Theology and Spirituality (UTS), HCCN has catalyzed spiritual care research through a grant from the John Templeton Foundation, which has resulted in ground-breaking studies that provide an evidence base for the effectiveness of spiritual care in health care. Through the publication of several key white papers, and the annual Caring for the Human Spirit Conference, HCCN's outreach and advocacy is now felt throughout the field of chaplaincy, nationally and internationally. Balboni TA, Paulk ME, Balboni MJ, Phelps AC, Loggers ET, Wright AA, Block SD, Lewis EF, Peteet JR, Prigerson HG. Provision of spiritual care to patients with advanced cancer: associations with medical care and quality of life near death. J Clin Oncol. 2010 Jan 20;28(3):445-52. doi: 10.1200/JCO.2009.24.8005. McClain, C. S., Rosenfeld, B., & Breitbart, W. (2003). Effect of spiritual well-being on end-of-life despair in terminally-ill cancer patients. The lancet, 361(9369), 1603-1607. Chen, J., Lin, Y., Yan, J., Wu, Y., & Hu, R. (2018). The effects of spiritual care on quality of life and spiritual well-being among patients with terminal illness: a systematic review. Palliative medicine, 32(7), 1167-1179. Murray, S. A., Kendall, M., Grant, E., Boyd, K., Barclay, S., & Sheikh, A. (2007). Patterns of social, psychological, and spiritual decline toward the end of life in lung cancer and heart failure. Journal of pain and symptom management, 34(4), 393-402.

Developer Response

Thank you for your comments regarding Hospice Visits in the Last Days of Life (HVLDL). we appreciate your thoughtful and input, and we have prepared response addressing the important issues you raised. We are grateful that the intent of the measure is understood. We were also happy that the measure's performance met all NQF criteria for variability, validity, and reliability, and was recommended for endorsement. We welcome the opportunity to address the issues raised. Visits by professional hospice staff - registered nurses and social workers - have been cited in focus groups as being particular helpful in the last days of life by bereaved family. Such attestations led CMS to incentivize visits by these staff, only (and not the full IDG team) in the Service Intensity Add-On policy implemented in 2016. Subsequently in development of HVLDL, CMS conducted a per-discipline analysis comparing the receipt of visits with the hospices' CAHPS outcome scores. Visits by registered nurses and social workers were the only two disciplines which yielded a meaningful positive correlation. A previously developed measure, Hospice Visits When Death is Imminent (HVWDII), encompassed a broader array of the disciplines of the IDG. This measure, encompassing the full IDG, failed to meet NQF testing standards, directly resulting from poor validity evidence (i.e., no relation to CAHPS scores), as detailed in a report CMS has published on its website since 2020 (https://www.cms.gov/files/document/hqrphospice-visits-when-death-imminent-testing-re-specification-reportoctober-2020.pdf). Based on our data analysis, we believe another measure broadly encompassing the full IDG team would similarly fail as was the case with HVWDII. CMS re-specified HVWDII as HVLDL, which meets testing criteria, and is moreover calculated using claims data, important information already collected by providers; CMS would be negligent to not publicly report this information, which we have shown to provide value to the

Hospice Quality Reporting Program. It should be noted the evidence for chaplain visits was mixed - that is, the additional inclusion of chaplain visits may meet NQF testing standards and bring demonstrated value to the HQRP. However, at present chaplain visits are not captured by claims data. CMS believes HVLDL which focuses on RN/SW visits, only, brings meaningful value to the HQRP., and the lack of chaplain visits should not prevent the public receive otherwise useful data. We appreciate the commenter's note to consider the HOPE data as a source of chaplain visits in the future. The commenter notes that end-of-life visits may not occur due to refusals. CMS had implicitly allowed for refusals during measure design, by specifying the measure to counts visits in two of the last three days or life, instead of visits on each of last three days. Also, CMS believes there is value to a broad, population-based measures. CMS certainly expects that caregiver refusals of visits will occur - and indeed family wishes of privacy near death a of paramount to be respected - and scores are not expected to ever be 100%. But basic analyses demonstrate there is important variation across hospices, more so than could plausibly be explained by differences in patient refusals across hospices. CMS believes this variation reveals meaningful differences in care delivery that could be useful to patients and their families when making a choice about the type of provider from whom they wish to receive care. The commenter raised the issue of telehealth. While we appreciate the comment, the steward at this time intends to keep the measure as specified, with in-person visits being the focus. CMS is proud of the new HOPE instrument currently in development, which will collect more information on hospice quality of care and will greatly enhance what is currently reported in the Hospice Quality Reporting Program. However, HOPE has not yet been nationally implemented, and no data has been collected, so it will be some time before measures from national HOPE data can be publicly reported. CMS has claims data on hand right now and would be remiss to not report this useful information. Patients and families making a difficult decision during an emotional time need assistance now and HVLDL will assist to help healthcare consumers make an informed choice.

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee notes that there is value in monitoring the quality of care provided by registered nurses and social workers during the last-days-of-life. While the Standing Committee recognizes the concern that certain disciplines are excluded from the measure, the Standing Committee maintains that the measure meets NQF criteria as specified and stands by the decision to recommend the measure for endorsement. However, the Standing Committee encourages the developer to monitor data and billing codes, as they become available, to support the inclusion of other interdisciplinary groups (e.g., chaplains, licensed practical nurses) within future iterations of the measure. The Standing Committee also recommended that the developer consider returning for early NQF maintenance review, prior to the designated three years, if including additional disciplines becomes more feasible.

Ms. Amy Melnick, MPA, National Coalition for Hospice and Palliative Care

Comment ID#: 7980 (Submitted: 04/29/2022)

Council / Public: QMRI

Level of Support: Member Does NOT Support

Comment

The National Coalition for Hospice and Palliative Care appreciates CMS' goal of developing a measure that seeks to improve the quality of care currently being provided to beneficiaries and their families at the end-of-life. We welcome the opportunity to comment on #3645 Hospice Visits in the Last Days of Life. The Coalition is comprised of 13 national organizations working together so that all patients, families and caregivers will have equitable access to quality hospice and palliative care. There is clear evidence that the last days of life for a hospice patient can be times of high symptom burden. Both patients and family caregivers can have high needs in all domains of care. If the patient and family desire a visit from hospice staff, it is the hospice's responsibility to deliver it. The evidence suggests that many hospices may not be meeting this goal reliably and there appears justification for a quality measure to help improve this situation. However, the measure itself should be of high quality and patient centered. The Meaningful Measure area for the Hospice Visits in Last Days of Life measure is "Person Centered Care" and the Healthcare Priority is "Strengthen Person & Family Engagement as Partners in their Care". We believe that, as currently constructed and without modification, the proposed measure does not fit within this priority and NQF should not endorse it. The central tenet of the patient and family-centered care that CMS promises to deliver to all beneficiaries is to enable the care that the patient and family want when they want it and not impose care on patients that they do not desire. This measure, as currently constructed incentivizes hospices to impose RN and social worker visits on patients and families who may not want them and where a visit from a different member of the hospice interdisciplinary group may be more appropriate. Furthermore, the HVLDL measures does not allow the virtual telehealth visits that some families prefer and discourages hospices from meeting the end-of-life spiritual and religious needs of patients and families by focusing heavily on only the medical portion of the interdisciplinary team. This measure does not serve the Meaningful Measure or Healthcare Priority as intended. It is our strong belief that analysis utilizing data that includes whether the patient/family desired a visit from the disciplines that are part of a measure or exclusion criteria that removes patients/caregivers who refuse visits offered by these various disciplines in the last days of life from the measure denominator would better reflect quality of care and better serve the Meaningful Measure and Healthcare Priority. Collecting and monitoring data of visits in the last days of life is highly desirable and needed, however, the Coalition strongly urges CMS to consider visit data in the context of an individualized plan of care reflective of patient and family wishes. The literature is clear that spiritual needs are prevalent at the end of life and rise in importance as death approaches. There are indications that spiritual well-being buffers the potential impact of end-of-life despair. Most religious traditions also have important end-of-life rituals. The National Consensus Project Clinical Guidelines for Quality Palliative Care propose that maximizing the benefits of supporting spiritual and religious needs requires a trained professional health care chaplain-a spiritual care specialist. The current measure is effectively a barrier to the delivery of that spiritual/religious care by not permitting chaplain visits to be included in this measure, even if this is what the family chooses and desires at the end of life. There seems to be a major disconnect between the stated goal of the measure and the research used to describe it. The developer states the goal of the measure as follows: "Collecting information about hospice staff visits for measuring quality of care will encourage hospices to visit patients and caregivers and provide services that will address their care needs and improve quality of life during the patients' last days of life." The Coalition supports the overarching goal of members of the interdisciplinary hospice team visiting patients and families during the last three days of life (unless of course the patient and family

declines a visit). However, there is no discussion here of the quality of the visits-only the fact that those visits were made. The developers justify the measure based on patient satisfaction scores on the CAHPS. In the process they limit the hospice's incentive to deliver visits by limiting the disciplines (RN and social workers only) whose visits count (excluding physicians, LPNs, aides and chaplains) and excluding virtual visits which many patients might prefer. It seems that the developers have conflated two different goals-increasing the number of patients visited and meeting patient needs on one hand and measuring family satisfaction with the visits that are made on the other. We strongly recommend restrictions on the members of the hospice staff whose visits count for the purposes of this measure. The 2020 Abt Associates report, Hospice Quality Reporting Program Quality Measures and Assessment Instruments Development, Modification and Maintenance, and Quality Reporting Oversight Support, provides background information on the development of the HVLDL. Unfortunately, the NQF Geriatrics and Palliative Care Committee was not presented with this report or the details of some of the data from the report prior to voting. Specifically, the data shows positive correlation between RN and MSW visits and CAHPS Hospice Survey outcomes although it is a low correlation. The correlation with chaplain and aide visits is also low. Granted, RN and MSW visit correlations are higher, but these visits are in the same correlation category as chaplain and aide visits per the data. None of the visit types had a moderate (0.5-0.7) or strong (0.7-1) correlation with the CAHPS Hospice Survey outcomes. In the initial review of the measure, the Committee suggested that the measure could be further strengthened by expanding the care disciplines covered, conducting a more holistic review of patient and caregiver end of life desires, and including postmortem visits and pediatric palliative care hospice patients. The Hospice Outcome & Patient Evaluation (HOPE) instrument, currently in the beta testing phase, will capture data as hospice care is being delivered to patients, a gap in the HQRP that CMS sought to close in recent years. The amount of data and information available not only to consumers, but also to CMS and hospice providers from the HQRP is relatively small. The HOPE will bring significantly more data and information to the HQRP that will allow for more robust quality measures. It is anticipated that the HOPE will be in use soon by hospices. The Coalition urges CMS to consider the impact of the measures anticipated from the HOPE on visits in last days of life, to eliminate any possible future duplication. In sum, this measure is much needed however, it must be of the highest quality itself. The measure can reach that status with three simple changes: 1) Removing the restrictions on the disciplines of the staff whose visits count, 2) allowing virtual visits, and 3) inserting an exception to the denominator for patients and families who are documented to not want a visit of any kind at end of life (last 3 days). Therefore, until such time as appropriate changes can be made, the Coalition recommends the Committee reconsider its vote to endorse the HVLDL measure. Balboni T, Balboni M, Paulk ME, Phelps A, Wright A, Peteet J, Block S, Lathan C, Vanderweele T, Prigerson H. Support of cancer patients' spiritual needs and associations with medical care costs at the end of life. Cancer. 2011 Dec 1;117(23):5383-91. doi: 10.1002/cncr.26221. Balboni TA, Paulk ME, Balboni MJ, Phelps AC, Loggers ET, Wright AA, Block SD, Lewis EF, Peteet JR, Prigerson HG. Provision of spiritual care to patients with advanced cancer: associations with medical care and quality of life near death. J Clin Oncol. 2010 Jan 20;28(3):445-52. doi: 10.1200/JCO.2009.24.8005. McClain, C. S., Rosenfeld, B., & Breitbart, W. (2003). Effect of spiritual well-being on end-of-life despair in terminally ill cancer patients. The lancet, 361(9369), 1603-1607. Chen, J., Lin, Y., Yan, J., Wu, Y., & Hu, R. (2018). The effects of spiritual care on quality of life and spiritual well-being among patients with terminal illness: a systematic review. Palliative medicine, 32(7), 1167-1179. Murray, S. A., Kendall, M., Grant, E., Boyd, K., Barclay, S., & Sheikh, A. (2007).

Patterns of social, psychological, and spiritual decline toward the end of life in lung cancer and heart failure. Journal of pain and symptom management, 34(4), 393-402.

Developer Response

Thank you for your comments regarding Hospice Visits in the Last Days of Life (HVLDL). We appreciate your thoughtful and input, and we have prepared response addressing the important issues you raised. We are grateful that the intent of the measure is understood. We were also happy that the measure's performance met all NQF criteria for variability, validity, and reliability, and was recommended for endorsement. We welcome the opportunity to address the issues raised. Visits by professional hospice staff - registered nurses and social workers - have been cited in focus groups as being particular helpful in the last days of life by bereaved family. Such attestations led CMS to incentivize visits by these staff, only (and not the full IDG team) in the Service Intensity Add-On policy implemented in 2016. Subsequently in development of HVLDL, CMS conducted a per-discipline analysis comparing the receipt of visits with the hospices' CAHPS outcome scores. Visits by registered nurses and social workers were the only two disciplines which yielded a meaningful positive correlation. A previously developed measure, Hospice Visits When Death is Imminent (HVWDII), encompassed a broader array of the disciplines of the IDG. This measure, encompassing the full IDG, failed to meet NQF testing standards, directly resulting from poor validity evidence (i.e., no relation to CAHPS scores), as detailed in a report CMS has published on its website since 2020 (https://www.cms.gov/files/document/hqrphospice-visits-when-death-imminent-testing-re-specification-reportoctober-2020.pdf). Based on our data analysis, we believe another measure broadly encompassing the full IDG team would similarly fail as was the case with HVWDII. CMS re-specified HVWDII as HVLDL, which meets testing criteria, and is moreover calculated using claims data, important information already collected by providers; CMS would be negligent to not publicly report this information, which we have shown to provide value to the Hospice Quality Reporting Program. It should be noted the evidence for chaplain visits was mixed - that is, the additional inclusion of chaplain visits may meet NQF testing standards and bring demonstrated value to the HQRP. However, at present chaplain visits are not captured by claims data. CMS believes HVLDL which focuses on RN/SW visits, only, brings meaningful value to the HQRP., and the lack of chaplain visits should not prevent the public receive otherwise useful data. We appreciate the commenter's note to consider the HOPE data as a source of chaplain visits in the future. The commenter notes that end-of-life visits may not occur due to refusals. CMS had implicitly allowed for refusals during measure design, by specifying the measure to counts visits in two of the last three days or life, instead of visits on each of last three days. Also, CMS believes there is value to a broad, population-based measures. CMS certainly expects that caregiver refusals of visits will occur - and indeed family wishes of privacy near death a of paramount to be respected - and scores are not expected to ever be 100%. But basic analyses demonstrate there is important variation across hospices, more so than could plausibly be explained by differences in patient refusals across hospices. CMS believes this variation reveals meaningful differences in care delivery that could be useful to patients and their families when making a choice about the type of provider from whom they wish to receive care. The commenter raised the issue of telehealth. While we appreciate the comment, the steward at this time intends to keep the measure as specified, with in-person visits being the focus. CMS is proud of the new HOPE instrument currently in development, which will collect more information on hospice quality of care and will greatly enhance what is currently reported in the Hospice Quality Reporting Program. However, HOPE has not yet been nationally implemented, and no data has been collected, so it will be some time

before measures from national HOPE data can be publicly reported. CMS has claims data on hand right now and would be remiss to not report this useful information. Patients and families making a difficult decision during an emotional time need assistance now and HVLDL will assist to help healthcare consumers make an informed choice.

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee notes that there is value in monitoring the quality of care provided by registered nurses and social workers during the last-days-of-life. While the Standing Committee recognizes the concern that certain disciplines are excluded from the measure, the Standing Committee maintains that the measure meets NQF criteria as specified and stands by the decision to recommend the measure for endorsement. However, the Standing Committee encourages the developer to monitor data and billing codes, as they become available, to support the inclusion of other interdisciplinary groups (e.g., chaplains, licensed practical nurses) within future iterations of the measure. The Standing Committee also recommended that the developer consider returning for early NQF maintenance review, prior to the designated three years, if including additional disciplines becomes more feasible.

Stephanie Collingwood, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 7964 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

UnityPoint Health respectfully offers comments in support of measure 3645. 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. While this measure may result in additional operational burden, its important and meaningful for patients to have direct care in the last few days prior to their death. This allows providers to deliver additional support to family/loved ones on medication and treatment plans to reduce pain and suffering. Direct care providers are trained to identify the signs of an impending death, which family/loved ones are not. Those instructions and preparation for the patient's passing allow for a more peaceful dying process. As a health care organization, we understand family members may not always have the knowledge to recognize clinical rationale for direct care provided within the few days prior to their loved one's death, and therefore, it would not be captured on the Hospice CAHPS survey. UnityPoint Health is supportive of this measure.

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

NQF #3665 Ambulatory Palliative Care Patients' Experience of Feeling Heard and Understood (Recommended)

Anna Kim, American Geriatrics Society; Submitted by Anna Kim

Comment ID#: 7960 (Submitted: 04/25/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

The AGS supports patients' experience of feeling heard and understood as a key goal and benefit of palliative care. Patients want to be treated as an individual and have their symptoms and goals of care managed effectively, which may be challenging at times given provider time constraints.

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

Marian Grant, C-TAC; Submitted by Dr. Marian Grant, DNP, RN

Comment ID#: 7977 (Submitted: 04/27/2022)

Council / Public: CON

Level of Support: N/A

Comment

C-TAC strongly supports NQF endorsement of this measure. It is unique in capturing the patient's experience of communication with a health care provider and team and is thus an important way to incorporate the patient's voice. Although it was tested in palliative care programs, we strongly feel it should be considered for use in all quality programs as a core measure. C-TAC has been supportive of its development and validation and will now advocate for its use with our health system and payer members and in our policy advocacy with the Centers for Medicare and Medicaid and the Innovation Center there as well.

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. Amy Melnick, MPA, National Coalition for Hospice and Palliative Care

Comment ID#: 7981 (Submitted: 04/29/2022)

Council / Public: QMRI

Level of Support: Member Does Support

Comment

The National Coalition for Hospice and Palliative Care strongly supports NQF endorsement of this measure, #3665. The Coalition praises NQF for recognizing the utility and suitability of Feeling Heard and Understood measure to help measure quality care from the patient's perspective. This measure is a patient reported measures and was developed with patients and caregivers at the table from the very beginning including in each phase of measure development. The Coalition notes and appreciates that the Standing Committee and draft report recognize that this measure is demonstratively meaningful directly to patients. Although the issue of survey fatigue was raised in the report, we would like to remind NQF that 87% of patient and caregiver respondents during the robust public comment period reported that they would be likely to complete a survey of their experience with their health care provider. We look forward to working with the measure steward, AAHPM at broadening this measure to other patient populations and settings.

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

Polly Friend

Comment ID#: 7973 (Submitted: 04/26/2022)

Council / Public: Public

Level of Support: N/A

Comment

WiserCare offers support for measure 3665. As a company, WiserCare respects the voice of the patient and the need for comprehensive discussions with a focus on understanding the patient's goals and preferences for care. We support this measure which helps to identify the extent the patient feels heard and understood, as well as to determine what is important in their life.

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#: 7963 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment
UnityPoint Health respectfully offers comments in support of measure 3665. 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. The Center to Advance Palliative Care (CAPC) standards for quality palliative care support the importance and value of comprehensive provider and interdisciplinary team discussions focused on understanding, advocating, and incorporating the patient's goals into their care plan. Therefore, a measurement capturing the patient's experience with feeling understood, their best interests advocated, and goals reflected in their care as a unique person is meaningful data to support an evidenced based intervention. From an operational perspective, new information capture mechanisms would have to be introduced through patient experience survey. Today, this is supported in a variety of methods with the majority of home care surveys captured through paper mail. Overall, UnityPoint Health is supportive of this measure.

Developer Response
Thank you for your support of the measure. We agree that understanding the patient's experience of feeling heard and understood is meaningful information to drive improvements in care.

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

NQF Committee Response
N/A

NQF #3666 Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain (Recommended)

Anna Kim, American Geriatrics Society; Submitted by Anna Kim
Comment ID#: 7961 (Submitted: 04/25/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

The AGS supports Measure #3666 : Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain, and believes patients' experience of receiving desired help for pain is also a key goal and benefit of palliative care. The interdisciplinary team structure of palliative care offers patients a more holistic mechanism of addressing their pain and with appropriate follow-up. Multiple modalities may also help with improved pain management.

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

Dr. Marian Grant, DNP, RN, Coalition to Transform Advanced Care (C-TAC)
Comment ID#: 7978 (Submitted: 04/27/2022)

Council / Public: CON

Level of Support: N/A

Comment

C-TAC supports this measure and appreciate that it incorporates the patient's perspective regarding what their goal for pain management is. As with NQF# 3665 Patients' Experience of Feeling Heard and Understood, once this is endorsed C-TAC will advocate for its use with our health system and payer members and in our policy advocacy with the Centers for Medicare and Medicaid and the Innovation Center there as well.

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. Amy Melnick, MPA, National Coalition for Hospice and Palliative Care
Comment ID#: 7985 (Submitted: 04/29/2022)

Council / Public: QMRI

Level of Support: Member Does Support

Comment

The National Coalition for Hospice and Palliative Care strongly supports NQF endorsement of this measure, #3666. The Coalition praises NQF for recognizing the utility and suitability of Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain. This measure is from the patient's perspective of getting help they desired for pain, and pain is broadly defined to include spiritual and other non-physical sources of pain. As a patient reported measure, it was developed with patients and caregivers at the table from the very beginning including in each phase of measure development. Importantly, the Receiving Desired Help for Pain measure assesses whether patients are getting the kind of care that they want. This is very different from surveying whether standardized clinical outcomes have been met. Assessment of pain-related clinical outcomes (e.g., asking how bad pain is, on a scale of 1 to 10) is already possible through existing performance measures, yet this is a one-size-fits-all approach that does not incorporate the patient's goals of care. Asking patients to report on their experience of the care they received, and whether they feel their problem was addressed as they wished, is the only way to reflect the patient's perspective. We look forward to working with the measure steward, AAHPM, at broadening this measure to other patient populations and settings.

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
Comment ID#: 7962 (Submitted: 04/25/2022)

Council / Public: PRO

Level of Support: N/A

Comment

3666- Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain UnityPoint Health respectfully offers comments in opposition of measure 3666 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. The Center to Advance Palliative Care (CAPC) Serious Illness Model priorities include symptom management, however, is not limited to pain management. Serious illness management is aligned with a broader definition of symptom management. The unique and limited focus on only pain is an archaic lens on mature palliative care practice. If this measure were to broaden the focus to include serious illness symptom management versus the limited lens of pain, it would align with best known practice. As reflected in measure 3665, the palliative care practice symptom management priorities should align with the patient's unique needs assessment and aligned with the patient's goals of care. Example: "The percentage of patients aged 18 years and older who had an ambulatory palliative care visit and report getting the help they wanted for their [serious illness symptoms] from their palliative care provider and team within 6 months of the ambulatory palliative care visit". From an operational perspective, new information capture mechanisms would have to be introduced through patient experience survey. Today, this is supported in a variety of methods with the majority of home care surveys captured through paper mail. UnityPoint Health opposes this measure as currently drafted and would recommend changing the language from "pain" to "serious illness symptoms".

Developer Response

Thank you for your comment. We agree that palliative care practice prioritizes serious illness symptom management broadly and not limited to pain. We limited the current measure development effort to pain management because it is a symptom commonly encountered in serious illness and was rated as a high priority for patients during our information gathering phase. Our measure was developed with input from a 30-member TECUPP which included patients and caregivers. The TECUPP discussed and ultimately decided against adding additional symptoms to the measure, in part due to concerns about measurement issues and difficulty comparing providers, since the measure was created for use in MIPS. Future work should expand on this to include other symptoms that may have different lookback periods, require additional cognitive testing to ensure appropriate wording and item structure, and as noted, require different information capture mechanisms.

NQF Response

N/A

NQF Committee Response

Thank you for your comment. The Standing Committee found that this measure meets NQF criteria as specified and maintained its decision to recommend the measure for endorsement.

Public Comments on Geriatrics and Palliative Care Fall 2021 Draft Report

N/A

Pre-Evaluation Measure-Specific Comments on Geriatrics and Palliative Care Fall 2021 Submissions

NQF #3665 Ambulatory Palliative Care Patients' Experience of Feeling Heard and Understood (Recommended)

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW
Comment ID#: 7864 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

These comments are in response to SMP review.

- **Issue 7: Cognitive Testing**

  R6:Was **cognitive debriefing** done with patients before the measure was tested? I have a few issues with the survey items. Specifically, **item #2 is double barreled** and research indicates that this leads to measurement error.

- **Developer Response 7:**

- We conducted 25 one-hour telephone cognitive interviews using a convenience sample of outpatient palliative care patients and caregivers to cognitively test survey items, with positive results. Participants generally understood the intended meaning of the question content. Some changes were made to improve the clarity of specific items. (See published manuscript: Rollison et al, Incorporating the Patient and Caregiver Voice in Palliative Care Quality Measure Development, Journal of Pain and Symptom Management, 2021)

- In particular, the "feeling heard and understood" concept was generally well-understood in its intended meaning as validation and acknowledgement from one's provider. It was determined necessary the two words – "heard" and "understood" together, because when asked separately, interviewees mistakenly understood the terms to refer to hearing (auditory ability) and comprehension (cognitive ability). This confusion also arose in early work to develop the single-item for use in inpatient settings (see: Gramling R et al, Feeling Heard and Understood: A Patient-Reported Quality Measure for the Inpatient Palliative Care Setting, Journal of Pain and Symptom Management, 2016), reinforcing our decision to use both words together to represent the single construct of feeling seen, respected, acknowledged.

- **Issue 8: Communicating scores to providers**

R9:Will there be any effort to communicate to the Provider the Top Box score on each of the four items so that the Provider can take a targeted intervention? The average score, while informative, does not provide the opportunity to make a targeted intervention.

- **Developer Response 8:** AAHPM will consider this option and if it's feasible, we will strive to provide the opportunity for targeted intervention.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7869 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 3: Risk Adjustment**

R3, R4:The risk model seems overly simplified, there are many factors that should have been looked into and potentially included, for example, administrative home type, disease status and others; Considered only a small number of patient level risk factors; lack of risk adjustment for patient level factors. Although I understand that this is because of lack of patient-level data on risk factors, this is not an "excuse" for the lack of risk adjustment.

- **Developer Response 3:**
- Using the data available to us (which was limited in terms of what programs were able to provide to us, and how much we could reliably capture via survey-based self-report), we did explore some potential program- and patient-level risk adjustment factors.
- Only survey mode was significant in its relationship with the HU performance measure ($p = 0.013$) and with programs ($p = 0.001$) after adjustment for multiple comparisons.
- At the patient-level, a single data element ("I felt this provider and team understood what is important to me out of life") of the four Feeling Heard and Understood data elements was significantly associated with diagnosis group ($p < 0.01$), and the raw measure score was significantly associated with diagnosis group. These results held after multiple comparison adjustments. Because of challenges with data quality, we were unable to conduct further analyses within the scope of this effort, but these findings provide preliminary indication that diagnosis might affect responses to the performance measure data elements and overall measure performance. **We acknowledge the importance of further research in this area before the measure is used for high-stakes decisions.**

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW
Comment ID#: 7875 (Submitted: 01/18/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

Reliability

- **Issue 3: Measure Score Calculation**

R3: Additionally, the developer should clearly describe how the measure score is calculated. Specifically, how the results from the hierarchical logistic regression model are rolled up to a measure score. The model specified seems to be a 2-level model, what's unit of analysis? Survey item level response (which is binary) or patient level score (which is not binary)? Measure score reliability was assessed via hierarchical logistic regression model, although it is not clear how it was done, at survey item response level or patient score level?

- **Developer Response 3:**

- The Feeling Heard and Understood measure score is calculated using a hierarchical (two-level) risk-adjusted binomial model (see mathematical details below). The scores are rolled up using a set of baseline characteristics (in this case proxy assistance status and survey mode) such that each provider has the same set of characteristics. The patient is the unit of analysis. We use a binomial model because each respondent contributes two pieces of information: 1) the number of responses provided across the four Feeling Heard and Understood items; and 2) the number of top-box responses. The individual's score on this measure is the proportion of top-box responses on these four items, i.e., a set of n = 4 trials with probability p of success. The average score is the estimated p (that as the reviewer notes, is not binary).

- More mathematically, our measure assumes that within provider i for each individual j, the k = 1,2,3,4 questions that they respond to are from the following parametric distribution, Y(subscript ij) ~ Binomial(n[subscript ij], p[subscript ij]) where n(subscript ij) = Sigma(subscript k)R(subscript ijk) <= 4 where R(subscript ijk) is one if a question k is responded to and zero otherwise. Thus, the unit of analysis is the patient-level, n(subscript ij) is the number of questions that an individual responded to, and p(subscript ij) represents an individual's average number of top-box responses on the four items. Explicitly, the individual's score arises as a non-continuous value because we have up to four binary outcomes that are contributing to the likelihood function.

- Let P(subscript i) represent an indicator that individual j received care from provider i, X(subscript ij) represents the patient's characteristics, then the risk-adjusted model for a provider score assumes the following generalized linear model logit(E[Y(subscript ij)|X(subscript ij), a, standardized beta]) = logit(p[subscript ij]) = (standardized beta[subscript 0] + b[subscript i]P[subscript i]) + X(superscript T, subscript ij)a with an

assumption that b(subscript i ~ N(0, lowercase omega[superscript 2, subscript b]). In this model, standardized beta(subscript 0) represents the average score across providers (i.e., grand mean), b(subscript i) is the difference between the average program score across providers (higher values represent better than average care) and a specific provider i's score, and a are risk adjusted coefficients.

- To calculate a specific providers score, let X* be a set of "baseline" characteristics to standardize an individual provider's score against, in our example, the characteristics were a fixed survey mode and no proxy assistance. The score for provider i is estimated using the following p-hat(subscript i) = logit(superscript -1)((standardized beta[subscript 0] + b[subscript i]) + X*[superscript T]a)
- In our specific submission X* was set to zero (indicating the baseline survey mode and no proxy assistance) and therefore the adjusted score is: p-hat(subscript i) = logit(superscript -1)(standardized beta[subscript 0] + b[subscript i])
- Hopefully this clarifies both our model and the estimation of the provider risk-score.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7862 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 6: Sampling**

R6: Also, on page 35 it is indicated that data should be collected from "eligible palliative care patients that are representative of the palliative care provider program." This indicates to me that some sampling technique is used but up to this point in the application I thought the practice would send data on all of the patients who met the criteria - not sample. This is an easy fix and just needs a clarification.

- **Developer Response 6:** Depending on the volume of patients and to support feasibility for programs, palliative care practices may survey all eligible patients or a *random* sample of eligible patients. The target population for sampling includes patients aged 18 years or older who received ambulatory palliative care services from a MIPS-eligible provider within the three months prior to the start of survey fielding. Findings from the alpha pilot test and beta field test support the feasibility of identifying eligible patients using administrative data and using a survey vendor to support survey administration and data collection. The provider or program will provide a vendor with an extract file of all patients who received care during the measurement period. To prevent gaming and to minimize administration

and social desirability bias, the vendor will apply the eligibility criteria to identify the patient sample and field the survey to eligible patients.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7861 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 5: Data Element Reliability**

R3:The developer ascertained both internal consistency and test-retest reliability for data elements. Each survey item has 5 response categories, however, for the measure, top box scoring is used. Therefore, the developer needs to clarify if the testing was consistent with the top box scoring approach; Data element reliability testing needs to be consistent with the top box scoring approach.

- **Developer Response 5:** To clarify, reliability of the four-data element scale using all 5 categorical options was high (Cronbach's alpha = 0.90) and similarly high for the dichotomous top-box option (Cronbach's alpha = 0.84).

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7857 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to the an SMP member's concerns.

- **Issue 2: Proxy Response**

R6:Also, it is stated throughout the application that responses completed by a proxy with our assistance from the patient will be excluded. I'm assuming (perhaps wrongly) that question 10 of the survey (option 3 - Answered the questions for me) will be used to determine this. If that is the case, I have an issue with this as I would not understand that response to indicate no patient involvement. Thus, I feel like this question needs to be re-worked. Also, it is indicated through the application that **surveys that were completely**

filled out by a proxy are excluded. However, it is unclear to me how this would be identified. I'm assuming (perhaps wrongly) that survey question 11 is used for this purpose and that option "answered the questions for me" is used to signify that the patient was not involved. However, I find this option unclear and I would not have understood it to indicate that the patient was not involved. Thus, I think this item needed to be re-worked to increase clarity before use.

- **Developer Response 2:**
- We excluded from the denominator patients for whom a proxy completed the entire survey on their behalf for any reason i.e., with no patient involvement, (proxy-only responses), but retained proxy-assistance responses, adjusting slightly upward for the latter in our measure scoring procedure, as indicated by our risk adjustment analysis.
- We defined “proxy-only” as the response option “answered the questions for me” to the question “How did that person help you complete the survey?”. This was the only response that indicated that the proxy actually provided the answers to the questions. Based on cognitive interviews and TECUPP input, we felt comfortable that this response option was indicative of **no** patient involvement. In contrast, we defined “proxy-assistance” as any or all of these responses: “read the questions to me”, “wrote down the answer I gave”, “translated the questions into my language; “helped in some other way”. Further work could reinforce these distinctions and identify slight revisions to increase clarity; the work done to date provides general support for the language currently used.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7859 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 3: Measure Score Calculation**

R3: Additionally, the developer should clearly describe how the measure score is calculated. Specifically, how the results from the hierarchical logistic regression model are rolled up to a measure score. The model specified seems to be a 2-level model, what's unit of analysis? Survey item level response (which is binary) or patient level score (which is not binary)? Measure score reliability was assessed via hierarchical logistic regression model, although it is not clear how it was done, at survey item response level or patient score level?

- **Developer Response 3:**

- The Feeling Heard and Understood measure score is calculated using a hierarchical (two-level) risk-adjusted binomial model (see mathematical details below). The scores are rolled up using a set of baseline characteristics (in this case proxy assistance status and survey mode) such that each provider has the same set of characteristics. The patient is the unit of analysis. We use a binomial model because each respondent contributes two pieces of information: 1) the number of responses provided across the four Feeling Heard and Understood items; and 2) the number of top-box responses. The individual's score on this measure is the proportion of top-box responses on these four items, i.e., a set of n = 4 trials with probability p of success. The average score is the estimated p (that as the reviewer notes, is not binary).
- Please see mathematical calculations and equations provided separately to NQF staff due to inability to copy them in this form.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7860 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 4: Measure Score Calculation**

R1:(Re: reliability testing) I was unclear if the hierarchical model accounted for nesting within patient and facility. It looks like items (not patient scores) were the level of analysis, but without accounting for the nesting within patient? If the patient score was the level of analysis, then I don't understand the model form (logistic); Entity-level testing revealed good signal to noise reliability, but I'm a little unclear how the beta binomial distribution was used when the patient score is a proportion (not binomial).

- **Developer Response 4:**
- **We believe we understand where the confusion arises. The model is estimated at the patient-level, where a patient's score is a summary of the four binary items. We use a binomial model where the number of trials (i.e., n) are the items and the outcome is the proportion (i.e., p) of top-box (binary) responses to these items and represents an individual's average top-box response. We are considering an individual's score as a simple average and not explicitly modeling an individual effect for items (i.e., no nesting). Where there might be confusion is that under a binomial model, the form of the model is very similar to a standard logistic regression, but the number of trials (i.e., n) is included in the actual estimation (see the probability distribution**

here https://en.wikipedia.org/wiki/Binomial_distribution **and when n = 1 we get back what is normally the logistic regression model for binary outcomes). There are more details in the previous response that explicitly mathematically describes the model that was estimated.**

- **We also should make clear that the model in the previous response is not exactly the same as the traditional beta-binomial model (**https://en.wikipedia.org/wiki/Beta-binomial distribution**) because we do not place beta priors on each individual's probability of success. Beta-binomial models have historical relevance in Bayesian estimation because of computational tractability, but recent software (e.g. the Stan programming language for Bayesian models** https://mc-stan.org/**) have made alternative models possible. Our model is a hierarchical generalized linear model where we assume a linear form on the probability of success to perform risk-adjustment and differs somewhat in structure from the beta-binomial but achieves the same effect.**

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7865 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

These comments are in response to SMP review.

Validity

- **Issue 1: Non-Response**

R1, R3:am concerned about survey non-response. Although not very large, there is variation in non-response between programs and demographic differences between responders and non-responders. I'm curious is the former is related to the latter. Are there better methods to account for survey non-response than just ignoring it? Non-response bias needs to be addressed with known differences between respondents and non-respondents.

- **Developer Response 1:**
- Of the 7,595 surveys we fielded, 2,804 were included as cases for analysis. Another 1,435 were deemed to be ineligible for the measure (eg: patient had died or disavowed the reference program or provider) and are thus not considered non-responders.
- Of the remaining 3,356 non-responders (i.e., surveys sent to presumably eligible patients but not returned to us), the majority (80%) were not reachable: 63% were not reachable after the maximum 8 phone call attempts and 17% had non-working phone numbers). Of note, another 14% were reachable but refused to complete the survey.

- As prior survey research has established, it is likely that people who do not return or respond to surveys are systematically different than those who do. This is particularly likely among respondents who explicitly decline or refuse to answer the survey. Our data suggest that survey respondents were slightly older than nonrespondents (mean age 63.4 versus 60.9; $p < 0.01$). The proportion of women was also higher among respondents as compared with nonrespondents (56.2 percent versus 54.5 percent), but the difference was not statistically significant ($p = 0.21$). Although information on patient race was self-reported via the survey instrument, a subset of 12 participating palliative care programs provided patient race for at least 90 percent of their patients in their submitted data files. Among this subset, there was a greater proportion of White patients (88.1 percent versus 80.2 percent) and a lower proportion of Black patients (8.8 percent versus 11.9 percent) in the respondent group compared with the nonrespondent group. The results of a chi-squared test indicate that this difference is statistically significant ($p < 0.01$).

- Because the non-responders did not return a survey, we were unable to compare differences in measure scores between them and responders. Although outside the scope of this initial testing effort, future work could attempt to explore other differences between these two groups, for example, to qualitatively understand whether their care experiences differed, in order to shed light on potential response bias.

- **Issue 2: Telehealth**

R6: I think Telehealth visits should be considered for inclusion in the future. R6, others: Concern about the exclusion of telehealth visits, should be included in the future

- **Developer Response 2:** We strongly agree that telehealth visits should be considered for inclusion in the future. Although we explored the inclusion of telephone and video visits as eligible visits at the outset of our alpha test, we decided not to include those visits because of their low frequency and difficulty identifying these visits. Thus, our initial performance measure eligibility criteria relied on coding in-person office visits. However, because of the COVID-19 pandemic, we were faced with an unexpected situation when participating palliative care programs shifted rapidly to providing telehealth services for their patients. With the input of our TECUPP and project advisory group, as well as input from participating programs, we decided to continue to disallow telehealth visits as eligible for the performance measure when we restarted data collection from September 2020 to February 2021. This ensured consistency in our results (i.e., we were measuring patient experiences with only in-person visits throughout the national beta field test) and avoided any potential confounding effects of the pandemic and telehealth use. However, it is likely that telehealth visits will continue in greater frequency than before the pandemic and should be included in measurement programs in the future. In interviews we conducted with palliative care programs during our testing phase, though most programs had little to no experience with telehealth prior to the pandemic, all programs converted to telehealth after March 2020 and continue to sustain telehealth services in some form. Closer attention to the development and testing of these and other patient experience measures within a telehealth context is warranted prior to widespread use in accountability programs.

- **Issue 3: Risk Adjustment**

R3, R4: The risk model seems overly simplified, there are many factors that should have been looked into and potentially included, for example, administrative home type, disease status and others; Considered only a small number of patient level risk factors; lack of risk adjustment for patient level factors. Although I understand that this is because of lack of patient-level data on risk factors, this is not an "excuse" for the lack of risk adjustment.

- **Developer Response 3:**
- Using the data available to us (which was limited in terms of what programs were able to provide to us, and how much we could reliably capture via survey-based self-report), we did explore some potential program- and patient-level risk adjustment factors.
- Only survey mode was significant in its relationship with the HU performance measure ($p = 0.013$) and with programs ($p = 0.001$) after adjustment for multiple comparisons.
- At the patient-level, a single data element ("I felt this provider and team understood what is important to me out of life") of the four Feeling Heard and Understood data elements was significantly associated with diagnosis group ($p < 0.01$), and the raw measure score was significantly associated with diagnosis group. These results held after multiple comparison adjustments. Because of challenges with data quality, we were unable to conduct further analyses within the scope of this effort, but these findings provide preliminary indication that diagnosis might affect responses to the performance measure data elements and overall measure performance. **We acknowledge the importance of further research in this area before the measure is used for high-stakes decisions.**

NQF Response

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

Katherine Ast, American Academy of Hospice and Palliative Medicine; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7855 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to a concern raised by one of the Scientific Methods Panel reviewer. Reliability

- **Issue 1: Attribution**

R3: My main concern is with the potential misalignment of provider attribution and patient-reported outcome attribution. Provider was identified based on a three-month period, MIPS-eligible provider who the patient saw most often during the three-month period. However, the attached survey form refers to "the last 6 months". Given that provider who the patient saw most often in the 3-month period may not be the same one in the 6-month period, and it is quite

likely that patient might have seen multiple providers during the 6-month period. Therefore, this may potentially cause provider misattribution. To further complicate things, the survey form does not identify the eligible ambulatory palliative care visit, so there is no explicit anchor visit for the patient to refer to even though the developer referred to the eligible ambulatory palliative visit repeatedly in this application, for example, the developer mentioned that patients who had transitioned to hospice could still answer the survey by reflecting on their experience with the visits.

- **Developer Response 1:** Our eligibility and sampling procedures, informed by input from our TECUPP, was designed to reduce the potential for misattribution as much as possible, while enhancing patient recall and their evaluation of the care they received from the palliative care provider and team.

- From the data files outpatient palliative care programs sent us, we first filtered to include only visits with Merit-based Incentive Payment System (MIPS)-eligible provider types that occurred in the three months prior to the anticipated start date of survey fielding (i.e., the planned date for mailing the prenotification letter to patients). We limited to 2019 MIPS-eligible providers so that these measures could be used for MIPS reporting). We limited eligible visits to a three-month period to ensure the recency of the visit patients should consider when responding about their experience. Setting this time frame also allowed each program's "clock" to start at the same time.

- We then identified a reference provider to be named on the survey instrument for each patient by selecting the MIPS-eligible provider whom the patient saw most often within the three-month period, with ties in numbers of visits broken by provider type, giving preference to providers holding primary responsibility for patient care outcomes (e.g., physician or physician-designee over nurse or therapist). If patients had multiple visits, we selected the most recent visit for each patient with the reference provider.

- The survey instrument included additional protections against misattribution. In both the survey cover letter as well as the instrument itself, we name the provider and team (eg: "Dr. Jones and team"). We included mention of the "team" because palliative care is an interdisciplinary team effort, and we anticipated that many patients would have seen the primary provider as well as other palliative care team members across and within visits that they had in the 3-month period. By naming the specific palliative care provider seen most often during the 3-month period, we hoped to avoid confusion with other providers outside palliative care that the patient might have seen.

- The survey instrument refers to a 6-month timeframe rather than the 3-month visit eligibility timeframe to cover potential lags in timing between when the palliative care program sent their data files, and when the survey was fielded and ultimately reached the patient.

- As an example, a program might have submitted a data file to us on September 1st, 2019, covering visits from March 1st through August 31st, 2019. We would sample visits June through August 2019 (the most recent 3 months of data), and field the survey September 25th (once all data files had been cleaned and prepared). The patient might then

receive/open the survey on October 1st, 2019. Referring to a 6-month timeframe (rather than a 3-month timeframe) thus covers the full sampling timeframe of June-August 2019.

- Guided by input from our TECUPP, we did not anchor the survey instrument to a specific single visit. Rather, we intentionally wanted patients to reflect on their experience of palliative care as a whole, rather than segmented into what happens in just a single visit, because palliative care as a discipline is intended to be holistic and comprehensive, with a longer-term care relationship. As such, the proposed measures reflect the experience of care over time and cannot be justifiably assessed after a single visit. For example, ensuring that a patient receives the help that they desire for their pain necessarily takes place over time rather than in a single visit.

NQF Response

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

Ms. Katherine Ast, MSW, LCSW, American Academy of Hospice and Palliative Medicine
Comment ID#: 7867 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

These comments are in response to SMP review.

Validity

- **Issue 1: Non-Response**

R1, R3:am concerned about survey non-response. Although not very large, there is variation in non-response between programs and demographic differences between responders and non-responders. I'm curious is the former is related to the latter. Are there better methods to account for survey non-response than just ignoring it? Non response bias needs to be addressed with known differences between respondents and non respondents.

- **Developer Response 1:**

- Of the 7,595 surveys we fielded, 2,804 were included as cases for analysis. Another 1,435 were deemed to be ineligible for the measure (eg: patient had died or disavowed the reference program or provider) and are thus not considered non-responders.

- Of the remaining 3,356 non-responders (i.e., surveys sent to presumably eligible patients but not returned to us), the majority (80%) were not reachable: 63% were not reachable after the maximum 8 phone call attempts and 17% had non-working phone numbers). Of note, another 14% were reachable but refused to complete the survey.

- As prior survey research has established, it is likely that people who do not return or respond to surveys are systematically different than those who do. This is particularly likely among respondents who explicitly decline or refuse to answer the survey. Our data suggest that survey respondents were slightly older than nonrespondents (mean age 63.4 versus

60.9; p < 0.01). The proportion of women was also higher among respondents as compared with nonrespondents (56.2 percent versus 54.5 percent), but the difference was not statistically significant (p = 0.21). Although information on patient race was self-reported via the survey instrument, a subset of 12 participating palliative care programs provided patient race for at least 90 percent of their patients in their submitted data files. Among this subset, there was a greater proportion of White patients (88.1 percent versus 80.2 percent) and a lower proportion of Black patients (8.8 percent versus 11.9 percent) in the respondent group compared with the nonrespondent group. The results of a chi-squared test indicate that this difference is statistically significant (p < 0.01).

- Because the non-responders did not return a survey, we were unable to compare differences in measure scores between them and responders. Although outside the scope of this initial testing effort, future work could attempt to explore other differences between these two groups, for example, to qualitatively understand whether their care experiences differed, in order to shed light on potential response bias.

- **Issue 2: Telehealth**

R6: I think Telehealth visits should be considered for inclusion in the future. R6, others: Concern about the exclusion of telehealth visits, should be included in the future

- **Developer Response 2:** We strongly agree that telehealth visits should be considered for inclusion in the future. Although we explored the inclusion of telephone and video visits as eligible visits at the outset of our alpha test, we decided not to include those visits because of their low frequency and difficulty identifying these visits. Thus, our initial performance measure eligibility criteria relied on coding in-person office visits. However, because of the COVID-19 pandemic, we were faced with an unexpected situation when participating palliative care programs shifted rapidly to providing telehealth services for their patients. With the input of our TECUPP and project advisory group, as well as input from participating programs, we decided to continue to disallow telehealth visits as eligible for the performance measure when we restarted data collection from September 2020 to February 2021. This ensured consistency in our results (i.e., we were measuring patient experiences with only in-person visits throughout the national beta field test) and avoided any potential confounding effects of the pandemic and telehealth use. However, it is likely that telehealth visits will continue in greater frequency than before the pandemic and should be included in measurement programs in the future. In interviews we conducted with palliative care programs during our testing phase, though most programs had little to no experience with telehealth prior to the pandemic, all programs converted to telehealth after March 2020 and continue to sustain telehealth services in some form. Closer attention to the development and testing of these and other patient experience measures within a telehealth context is warranted prior to widespread use in accountability programs.

NQF Response

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

NQF #3666 Ambulatory Palliative Care Patients' Experience of Receiving Desired Help for Pain (Recommended)

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW
Comment ID#: 7866 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

These comments are in response to SMP review.

Validity

- **Issue 1: Non-Response**

R1, R3:am concerned about survey non-response. Although not very large, there is variation in non-response between programs and demographic differences between responders and non-responders. I'm curious is the former is related to the latter. Are there better methods to account for survey non-response than just ignoring it? Non response bias needs to be addressed with known differences between respondents and non respondents.

- **Developer Response 1:**
- Of the 7,595 surveys we fielded, 2,804 were included as cases for analysis. Another 1,435 were deemed to be ineligible for the measure (eg: patient had died or disavowed the reference program or provider) and are thus not considered non-responders.
- Of the remaining 3,356 non-responders (i.e., surveys sent to presumably eligible patients but not returned to us), the majority (80%) were not reachable: 63% were not reachable after the maximum 8 phone call attempts and 17% had non-working phone numbers). Of note, another 14% were reachable but refused to complete the survey.
- As prior survey research has established, it is likely that people who do not return or respond to surveys are systematically different than those who do. This is particularly likely among respondents who explicitly decline or refuse to answer the survey. Our data suggest that survey respondents were slightly older than nonrespondents (mean age 63.4 versus 60.9; $p < 0.01$). The proportion of women was also higher among respondents as compared with nonrespondents (56.2 percent versus 54.5 percent), but the difference was not statistically significant ($p = 0.21$). Although information on patient race was self-reported via the survey instrument, a subset of 12 participating palliative care programs provided patient race for at least 90 percent of their patients in their submitted data files. Among this subset, there was a greater proportion of White patients (88.1 percent versus 80.2 percent) and a lower proportion of Black patients (8.8 percent versus 11.9 percent) in the respondent group compared with the nonrespondent group. The results of a chi-squared test indicate that this difference is statistically significant ($p < 0.01$).
- Because the non-responders did not return a survey, we were unable to compare differences in measure scores between them and responders. Although outside the scope

of this initial testing effort, future work could attempt to explore other differences between these two groups, for example, to qualitatively understand whether their care experiences differed, in order to shed light on potential response bias.

- **Issue 2: Telehealth**

R6: I think Telehealth visits should be considered for inclusion in the future. R6, others: Concern about the exclusion of telehealth visits, should be included in the future

- **Developer Response 2:** We strongly agree that telehealth visits should be considered for inclusion in the future. Although we explored the inclusion of telephone and video visits as eligible visits at the outset of our alpha test, we decided not to include those visits because of their low frequency and difficulty identifying these visits. Thus, our initial performance measure eligibility criteria relied on coding in-person office visits. However, because of the COVID-19 pandemic, we were faced with an unexpected situation when participating palliative care programs shifted rapidly to providing telehealth services for their patients. With the input of our TECUPP and project advisory group, as well as input from participating programs, we decided to continue to disallow telehealth visits as eligible for the performance measure when we restarted data collection from September 2020 to February 2021. This ensured consistency in our results (i.e., we were measuring patient experiences with only in-person visits throughout the national beta field test) and avoided any potential confounding effects of the pandemic and telehealth use. However, it is likely that telehealth visits will continue in greater frequency than before the pandemic and should be included in measurement programs in the future. In interviews we conducted with palliative care programs during our testing phase, though most programs had little to no experience with telehealth prior to the pandemic, all programs converted to telehealth after March 2020 and continue to sustain telehealth services in some form. Closer attention to the development and testing of these and other patient experience measures within a telehealth context is warranted prior to widespread use in accountability programs.

- **Issue 3: Risk Adjustment**

R3, R4:The risk model seems overly simplified, there are many factors that should have been looked into and potentially included, for example, administrative home type, disease status and others; Considered only a small number of patient level risk factors; lack of risk adjustment for patient level factors. Although I understand that this is because of lack of patient-level data on risk factors, this is not an "excuse" for the lack of risk adjustment.

- **Developer Response 3:**

- Using the data available to us (which was limited in terms of what programs were able to provide to us, and how much we could reliably capture via survey-based self-report), we did explore some potential program- and patient-level risk adjustment factors.

- Only survey mode was significant in its relationship with the HU performance measure ($p = 0.013$) and with programs ($p = 0.001$) after adjustment for multiple comparisons.

- At the patient-level, a single data element ("I felt this provider and team understood what is important to me out of life") of the four Feeling Heard and Understood data elements was significantly associated with diagnosis group ($p < 0.01$), and the raw measure score was significantly associated with diagnosis group. These results held after multiple comparison adjustments. Because of challenges with data quality, we were unable to conduct further analyses within the scope of this effort, but these findings provide preliminary indication that diagnosis might affect responses to the performance measure data elements and overall measure performance. **We acknowledge the importance of further research in this area before the measure is used for high-stakes decisions.**

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7863 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 6: Sampling**

R6: Also, on page 35 it is indicated that data should be collected from "eligible palliative care patients that are representative of the palliative care provider program." This indicates to me that some sampling technique is used but up to this point in the application I thought the practice would send data on all of the patients who met the criteria - not sample. This is an easy fix and just needs a clarification.

- **Developer Response 6:** Depending on the volume of patients and to support feasibility for programs, palliative care practices may survey all eligible patients or a *random* sample of eligible patients. The target population for sampling includes patients aged 18 years or older who received ambulatory palliative care services from a MIPS-eligible provider within the three months prior to the start of survey fielding. Findings from the alpha pilot test and beta field test support the feasibility of identifying eligible patients using administrative data and using a survey vendor to support survey administration and data collection. The provider or program will provide a vendor with an extract file of all patients who received care during the measurement period. To prevent gaming and to minimize administration and social desirability bias, the vendor will apply the eligibility criteria to identify the patient sample and field the survey to eligible patients.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW
Comment ID#: 7858 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to an SMP member's concerns.

- **Issue 2: Proxy Response**
  R6:Also, it is stated throughout the application that responses completed by a proxy with our assistance from the patient will be excluded. I'm assuming (perhaps wrongly) that question 10 of the survey (option 3 - Answered the questions for me) will be used to determine this. If that is the case, I have an issue with this as I would not understand that response to indicate no patient involvement. Thus, I feel like this question needs to be re-worked. Also, it is indicated through the application **that surveys that were completely filled out by a proxy are excluded**. However, it is unclear to me how this would be identified. I'm assuming (perhaps wrongly) that survey question 11 is used for this purpose and that option "answered the questions for me" is used to signify that the patient was not involved. However, I find this option unclear, and I would not have understood it to indicate that the patient was not involved. Thus, I think this item needed to be re-worked to increase clarity before use.

- **Developer Response 2:**

- We excluded from the denominator patients for whom a proxy completed the entire survey on their behalf for any reason i.e., with no patient involvement, (proxy-only responses), but retained proxy-assistance responses, adjusting slightly upward for the latter in our measure scoring procedure, as indicated by our risk adjustment analysis.

- We defined “proxy-only” as the response option “answered the questions for me” to the question “How did that person help you complete the survey?”. This was the only response that indicated that the proxy actually provided the answers to the questions. Based on cognitive interviews and TECUPP input, we felt comfortable that this response option was indicative of **no** patient involvement. In contrast, we defined “proxy-assistance” as any or all of these responses: “read the questions to me”, “wrote down the answer I gave”, “translated the questions into my language; “helped in some other way”. Further work could reinforce these distinctions and identify slight revisions to increase clarity; the work done to date provides general support for the language currently used.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW
Comment ID#: 7870 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

These comments are in response to SMP review.

Reliability

- **Issue 3: Measure Score Reliability**

R6: Measure score - The adjusted ICC (0.079 with CI 0.02-0.175) is extremely low and is concerning. However, the individual program reliability (especially when taking into account the programs that met the minimum number of respondents is 0.735 which is good. R6: I rated low based solely on the ICC results. R9: This is a benefit of the doubt rating, measure score reliability was low.

- **Developer Response 3:**

- Since reliability is a function of both sample size and ICC, we believe the adjusted ICC on its own is not concerning. Various patient experience surveys have very low ICC's for item responses. For example, from "Psychometric Properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician and Group Adult Visit Survey" it was reported that the item ICCs for Access to Care were an average of 0.08, ranging from 0.07 to 0.11, all above the set 0.05 criterion (see section on "Multilevel Analyses"). Similarly, for CAHPS Hospice, ICCs for both their composite and single item measures range from 0.010 to 0.021 (see "Development of Valid and Reliable Measures of Patient and Family Experiences of Hospice Care for Public Reporting"). Considering ***both*** sample size and ICC, our measure test suggests that to achieve a reliability around 0.7, providers must have at least 33 respondents. We acknowledge that only 30% of programs ***in our test*** met this threshold (in implementation, this number could be higher, as we describe in our response to the comment below).

  Dyer, Naomi, Joann S. Sorra, Scott A. Smith, Paul Cleary, and Ron Hays. "Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) clinician and group adult visit survey." *Medical care* 50, no. Suppl (2012): S28.

  Anhang Price, Rebecca, Brian Stucky, Layla Parast, Marc N. Elliott, Ann Haas, Melissa Bradley, and Joan M. Teno. "Development of valid and reliable measures of patient and family experiences of hospice care for public reporting." Journal of palliative medicine 21, no. 7 (2018): 924-932. · **Issue 4: Minimum Patient Volume** R9: The average reliability for all group/programs for the measure score was 0.482 with a wide range of values. However, when the requirement of n=33 was imposed, reliability jumped to 0.735 with a narrow range of values. However, this reduced the reportability of these results to only 30% of the beta (field) test sample groups/programs. Will reportability be an issue when the measure is scaled to a national roll-out? R3: Average reliability was around 0.48. After imposing 33 volume restriction, average reliability was around 0.73 but it would remove many programs. **Developer Response 4: As noted by the reviewer, only 13 of 43 programs (30%) had sufficient patient volume to meet the minimum required respondents for a reliability measure score. Although our sample of outpatient palliative care programs did not include all programs in the United States who might have been able to participate, this drop-off in the number of programs does raise concerns about reportability and participation upon**

national implementation. It is possible that more programs would participate if the measures are implemented. It is also possible that the data submitted by participating programs to us for the test was limited (eg: by lack of dedicated resources to prepare data files, by the onset of the pandemic) and that once implemented, more of these programs would meet the minimum numbers of respondents. Further work will be important to address this and other issues related to implementation, that can only be accomplished once these measures are rolled out more widely.

NQF Response

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

Katherine Ast, AAHPM; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7871 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to SMP review.

- **Issue 3: Risk Adjustment**

R3: Risk adjustment approach seem incomplete. While there are data availability issues, important factors such as disease status could have been captured and included. R4: Lack of meaningful risk adjustment.

- **Developer Response 3:**
- Using the data available to us (which was limited in terms of what programs were able to provide to us, and how much we could reliably capture via survey-based self-report), we did explore some potential program- and patient-level risk adjustment factors.
- None of the potential risk adjustment variables were significant in their relationship with the pain measure after adjustment for multiple comparisons. However, our TECUPP emphasized the importance of considering inclusion of some variables, such as survey mode and proxy assistance, to increase the face validity of our modeling.
- At the patient level, the *Receiving Desired Help for Pain* data element was significantly associated with diagnosis group ($p<0.01$). The quality measure score was also significantly associated with diagnosis group. These results held after multiple comparison adjustments. Because of challenges with data quality, we were unable to conduct further analyses within the scope of this effort, but these findings provide preliminary indication that diagnosis might affect responses to the performance measure data elements and overall measure performance. **We acknowledge the importance of further research in this area before the measure is used for high-stakes decisions.**

NQF Response

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

Katherine Ast, American Academy of Hospice and Palliative Medicine; Submitted by Ms. Katherine Ast, MSW, LCSW

Comment ID#: 7856 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

This comment is in response to a concern raised by one of the Scientific Methods Panel reviewer. Reliability

- **Issue 1: Attribution**

R3: My main concern is with the potential misalignment of provider attribution and patient-reported outcome attribution. Provider was identified based on a three-month period, MIPS-eligible provider who the patient saw most often during the three-month period. However, the attached survey form refers to "the last 6 months". Given that provider who the patient saw most often in the 3-month period may not be the same one in the 6-month period, and it is quite likely that patient might have seen multiple providers during the 6-month period. Therefore, this may potentially cause provider misattribution. To further complicate things, the survey form does not identify the eligible ambulatory palliative care visit, so there is no explicit anchor visit for the patient to refer to even though the developer referred to the eligible ambulatory palliative visit repeatedly in this application, for example, the developer mentioned that patients who had transitioned to hospice could still answer the survey by reflecting on their experience with the visits.

- **Developer Response 1:** Our eligibility and sampling procedures, informed by input from our TECUPP, was designed to reduce the potential for misattribution as much as possible, while enhancing patient recall and their evaluation of the care they received from the palliative care provider and team.

- From the data files outpatient palliative care programs sent us, we first filtered to include only visits with Merit-based Incentive Payment System (MIPS)-eligible provider types that occurred in the three months prior to the anticipated start date of survey fielding (i.e., the planned date for mailing the prenotification letter to patients). We limited to 2019 MIPS-eligible providers so that these measures could be used for MIPS reporting). We limited eligible visits to a three-month period to ensure the recency of the visit patients should consider when responding about their experience. Setting this time frame also allowed each program's "clock" to start at the same time.

- We then identified a reference provider to be named on the survey instrument for each patient by selecting the MIPS-eligible provider whom the patient saw most often within the three-month period, with ties in numbers of visits broken by provider type, giving preference to providers holding primary responsibility for patient care outcomes (e.g.,

physician or physician-designee over nurse or therapist). If patients had multiple visits, we selected the most recent visit for each patient with the reference provider.

- The survey instrument included additional protections against misattribution. In both the survey cover letter as well as the instrument itself, we name the provider and team (eg: "Dr. Jones and team"). We included mention of the "team" because palliative care is an interdisciplinary team effort, and we anticipated that many patients would have seen the primary provider as well as other palliative care team members across and within visits that they had in the 3-month period. By naming the specific palliative care provider seen most often during the 3-month period, we hoped to avoid confusion with other providers outside palliative care that the patient might have seen.

- The survey instrument refers to a 6-month timeframe rather than the 3-month visit eligibility timeframe to cover potential lags in timing between when the palliative care program sent their data files, and when the survey was fielded and ultimately reached the patient.

- As an example, a program might have submitted a data file to us on September 1st, 2019, covering visits from March 1st through August 31st, 2019. We would sample visits June through August 2019 (the most recent 3 months of data), and field the survey September 25th (once all data files had been cleaned and prepared). The patient might then receive/open the survey on October 1st, 2019. Referring to a 6-month timeframe (rather than a 3-month timeframe) thus covers the full sampling timeframe of June-August 2019.

- Guided by input from our TECUPP, we did not anchor the survey instrument to a specific single visit. Rather, we intentionally wanted patients to reflect on their experience of palliative care as a whole, rather than segmented into what happens in just a single visit, because palliative care as a discipline is intended to be holistic and comprehensive, with a longer-term care relationship. As such, the proposed measures reflect the experience of care over time and cannot be justifiably assessed after a single visit. For example, ensuring that a patient receives the help that they desire for their pain necessarily takes place over time rather than in a single visit.

NQF Response

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

Ms. Katherine Ast, MSW, LCSW, American Academy of Hospice and Palliative Medicine
Comment ID#: 7868 (Submitted: 01/17/2022)

Council / Public: HPR

Level of Support: Member Does Support

Comment

These comments are in response to SMP review.

Validity

- **Issue 1: Non-Response**

R1, R3:am concerned about survey non-response. Although not very large, there is variation in non-response between programs and demographic differences between responders and non-responders. I'm curious is the former is related to the latter. Are there better methods to account for survey non-response than just ignoring it? Non response bias needs to be addressed with known differences between respondents and non-respondents.

- **Developer Response 1:**
- Of the 7,595 surveys we fielded, 2,804 were included as cases for analysis. Another 1,435 were deemed to be ineligible for the measure (eg: patient had died or disavowed the reference program or provider) and are thus not considered non-responders.
- Of the remaining 3,356 non-responders (i.e., surveys sent to presumably eligible patients but not returned to us), the majority (80%) were not reachable: 63% were not reachable after the maximum 8 phone call attempts and 17% had non-working phone numbers). Of note, another 14% were reachable but refused to complete the survey.
- As prior survey research has established, it is likely that people who do not return or respond to surveys are systematically different than those who do. This is particularly likely among respondents who explicitly decline or refuse to answer the survey. Our data suggest that survey respondents were slightly older than nonrespondents (mean age 63.4 versus 60.9; $p < 0.01$). The proportion of women was also higher among respondents as compared with nonrespondents (56.2 percent versus 54.5 percent), but the difference was not statistically significant ($p = 0.21$). Although information on patient race was self-reported via the survey instrument, a subset of 12 participating palliative care programs provided patient race for at least 90 percent of their patients in their submitted data files. Among this subset, there was a greater proportion of White patients (88.1 percent versus 80.2 percent) and a lower proportion of Black patients (8.8 percent versus 11.9 percent) in the respondent group compared with the nonrespondent group. The results of a chi-squared test indicate that this difference is statistically significant ($p < 0.01$).
- Because the non-responders did not return a survey, we were unable to compare differences in measure scores between them and responders. Although outside the scope of this initial testing effort, future work could attempt to explore other differences between these two groups, for example, to qualitatively understand whether their care experiences differed, in order to shed light on potential response bias.
- **Issue 2: Telehealth**

R6: I think Telehealth visits should be considered for inclusion in the future. R6, others: Concern about the exclusion of telehealth visits, should be included in the future

- **Developer Response 2:** We strongly agree that telehealth visits should be considered for inclusion in the future. Although we explored the inclusion of telephone and video visits as eligible visits at the outset of our alpha test, we decided not to include those visits because of their low frequency and difficulty identifying these visits. Thus, our initial performance measure eligibility criteria relied on coding in-person office visits. However, because of the COVID-19 pandemic, we were faced with an unexpected situation when participating

palliative care programs shifted rapidly to providing telehealth services for their patients. With the input of our TECUPP and project advisory group, as well as input from participating programs, we decided to continue to disallow telehealth visits as eligible for the performance measure when we restarted data collection from September 2020 to February 2021. This ensured consistency in our results (i.e., we were measuring patient experiences with only in-person visits throughout the national beta field test) and avoided any potential confounding effects of the pandemic and telehealth use. However, it is likely that telehealth visits will continue in greater frequency than before the pandemic and should be included in measurement programs in the future. In interviews we conducted with palliative care programs during our testing phase, though most programs had little to no experience with telehealth prior to the pandemic, all programs converted to telehealth after March 2020 and continue to sustain telehealth services in some form. Closer attention to the development and testing of these and other patient experience measures within a telehealth context is warranted prior to widespread use in accountability programs.

NQF Response

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