Hospice Quality Reporting Program Quality Measures and Assessment Instruments Development, Modification and Maintenance, and Qualitry Reporting Oversight Support

Hospice Visits When Death is Imminent: Measure Validity Testing Summary and Re-Specifications

Contract No. 75FCMC18D0014,

Task Order No. 75FCMC19F0001

Prepared for:
Centers for Medicare & Medicaid Services
7500 Security Blvd.
Baltimore, MD 21244-1850
Attn: Cindy Massuda

Prepared by:
Abt Associates
10 Fawcett Street
Cambridge, MA 02138

September 11, 2020

Background

CMS implemented the Hospice Visits when Death is Imminent (HVWDII) measure pair in the Fiscal Year (FY) 2017 Hospice Final rule (*Medicare Program; FY 2017 Hospice Wage Index and Payment Rate Update and Hospice Quality Reporting Requirements*). The information used to calculate HVWDII is sourced from the Hospice Item Set (HIS) V2.00.0. Data collection for the measure pair began April 1, 2017.

The measure pair focuses on visits received by hospice patients in the last week of life. This period is typically associated with a higher burden of physical and emotional symptoms, which necessitates close attention from the interdisciplinary hospice care team. The first measure in the pair, Measure 1, addresses clinical hospice visits, including visits by registered nurses (RN), physicians, nurse practitioners or physician assistants, in the final three days of life. CMS found that this measure met established standards for reliability, validity, and reportability, and has publicly reported Measure 1 since August 2019.

The second measure of the pair, Measure 2, addresses non-clinical hospice visits in the final seven days of life, including visits by medical social workers (MSW), chaplains or spiritual counselors, licensed practical nurses (LPN) and hospice aides. As stated in the FY 2017 Hospice Final rule, focusing Measure 2 on the interdisciplinary team was intended to give hospices the "flexibility to provide individualized care that is in line with the patient, family, and caregiver's preferences and goals for care" (81 FR 25498, 25522).

Unlike Measure 1, Measure 2 did not meet CMS's readiness standards for public reporting. In particular, the measure failed to meet CMS's *validity testing* criteria. Validity testing, as described by the National Quality Forum (NQF), establishes whether "the [quality] measure score correctly reflects the quality of care provided, adequately identifying differences in quality" (National Quality Forum, 2019).[1] Specifically, these analyses seek to establish whether a measure is able to distinguish high- and low-quality hospices.

In the FY 2020 Hospice Wage Index and Payment Rate Update final rule (84 FR 38484, 38527-38529), CMS stated that Measure 2 would not be publicly reported on Hospice Compare website until further testing was conducted to determine if and how the measure specifications might be modified or re-specified so that the measure would meet the standards of scientific acceptability and reportability. As finalized on page 38529 of the FY 2020 rule:

> *Final Decision*: After considering the comments received in response to the proposed rule and for the reasons discussed in the above paragraph, we are finalizing our proposal to continue collection of this data to complete additional testing and to make a determination about the public reporting of Measure 2 of the ''Hospice Visits when Death is Imminent'' measure pair. We expect to complete our analysis by the end of FY 2020, and determine next steps for public reporting based on meeting established standards for reliability, validity, and reportability.

[1] See: Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement. (2019). National Quality Forum. Available via: http://www.qualityforum.org/docs/measure_evaluation_criterias.aspx.

During the time since Measure 1 has been publically reported, CMS has continued to collect the data necessary to calculate Measure 2. CMS has provided calculated Measure 2 scores privately to hospices using these data through the present. The data have also enabled CMS to continue testing of Measure 2 performance.

Results from the initial testing of Measure 2 suggested that CMS could re-specify the measure concept and achieve better testing result than achieved currently by the HVWDII measure pair; moreover, re-specification of Measure 2 could result in lower public burden.

Further analysis led to the development of a re-specified version of the HVWDII measure, calculated as claims data and referred to as the Hospice Visits in the Last Days of Life (HVLDL) measure. Comprehensive measure testing demonstrates that, relative to the HVWDII measure pair, the new HVLDL measure achieves:

- Improved ability to differentiate higher from lower quality hospices
- Quality rankings more consistent with those produced with other quality measures in the HQRP
- Alignment with the Service Intensity Add-On (SIA), CMS's payment policy initiative[2] implemented in 2016 which seeks to incentivize visits by registered nurses and medical social workers when patients are near death
- Reliance solely on existing administrative data for calculation, removing the need for data collection through clinician assessment.

The purpose of this report is to share findings of the testing on HVWDII Measure 2 which led to further re-specification and measure analysis towards the development of the HVLDL measure. The report reviews the methodology, analysis, results, and conclusions that motivated CMS to favor adoption of the HVLDL specification of the measure concept.

Methods

CMS's approach to measure performance testing adheres to the evaluation criteria outlined by the National Quality Forum.[3] The key testing analyses aim to ensure that quality measures exhibit appropriate reportability (whether sufficient providers have enough patients to calculate and report scores publically while ensuring patient privacy), variability (whether a sufficient range in performance exists to differentiate providers), reliability (whether the measure can be implemented consistently and distinguish differences in performance)[4], and validity (as defined above). As mentioned in the background section, Measure 2 failed to meet the standards for public reporting due to a lack of *validity*.

[2] Under the Service Intensity Add-On (SIA), implemented by CMS in 2016 (80 FR 47173), visits by RNs or (in-person) MSWs occurring in the last 7 days of life on Routine Home Care (RHC) days receive an additional payment in addition to the RHC per diem payment. The payment is equal to the Continuous Home Care (CHC) rate billed in 15 minute increments for services up to 4 hours total daily.

[3] See: Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement. (2019). National Quality Forum. Available via: http://www.qualityforum.org/docs/measure_evaluation_criterias.aspx.

[4] For more information about CMS's measure criteria, please see the *Blueprint for the CMS Measures Management System* (Version 15.0; September 2019), available via: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Blueprint.pdf.

The approach to establishing validity was to calculate correlation coefficients between the receipt of visits at the end of life and outcomes calculated from the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Hospice Survey data collected from caregivers of decedent hospice patients. Validity testing demonstrates to extent to which a measure's quality rankings are consistent with other measures' rankings. In the HQRP, the only other quality measure (besides the CAHPS® Hospice Survey outcomes) are calculated from data collected through the HIS and relate to processes of care around the time of admission. The CAHPS® Hospice Survey outcomes, calculated from caregiver responses after the patient's death, are a more suitable benchmark to a measure concerned with receipt of visits near to the patient's death.

The eight CAHPS® Hospice Survey outcomes have collectively been endorsed by NQF as NQF #2651 and are currently publically reported on the Hospice Compare website.[5] The eight outcomes are:

1. Communication with family
2. Getting timely help
3. Treating patient with respect
4. Emotional and spiritual support
5. Help for pain and symptoms
6. Training family to care for patient
7. Rating of this hospice
8. Willing to recommend this hospice

The validity analyses sought to identify the extent of correlation between the HVWDII measure pairs and the CAHPS® Hospice Survey outcomes. Because visits at the end of life are presumed to indicate good quality of care, evidence for validity would present as a positive correlation between the provision of visits and the CAHPS® Hospice Survey outcomes. That is, hospices providing higher rates of visits at the end of life were expected to also have better quality outcomes as reported by caregivers for decedent patients. Such a pattern would yield a positive correlation, and the stronger the pattern, the more positive the correlation.

CMS calculated the coefficients of correlation between visits from individual disciplines and each CAHPS quality measure. The disciplines examined were those that comprised the predominate amount of visits included in each measure:

- *Measure 1*: Registered Nurse (RN) and Physician visits
- *Measure 2*: Medical Social Worker (MSW), Chaplain, Licensed Practical Nurse (LPN), and Home Aide visits.

For each discipline, CMS calculated the hospice-level percentages of patients who received a visit from that discipline, in two different timeframes: 1) within the last three days of life (the timeframe of Measure 1); and 2) within the last seven days of life (the timeframe of Measure 2). CMS calculated correlation coefficients using each discipline and timeframe. CMS also tested combinations of the disciplines included in Measure 1 and Measure 2, separately, for those time periods as well. Analyses used HIS data for patients discharged between April 1, 2017 and March 30, 2018. As will be shown,

[5] See: *Hospice Compare*, available via: https://www.medicare.gov/hospicecompare/.

analyses suggested that CMS could respecify the measure and achieve improved performance results using visit data collected wholly from Medicare hospice claim records. Further testing of a claims-based version of the end-of-life visits construct was conducted using claims data in Federal Fiscal Year 2018 (October 1, 2017 through September 30, 2018). Descriptive statistics of performance scores' distributions were used to assess variability and as with the HVWDII measure pair correlations between measure scores and CAHPS® Hospice Survey outcomes were used to assess validity.

Analysis & Results

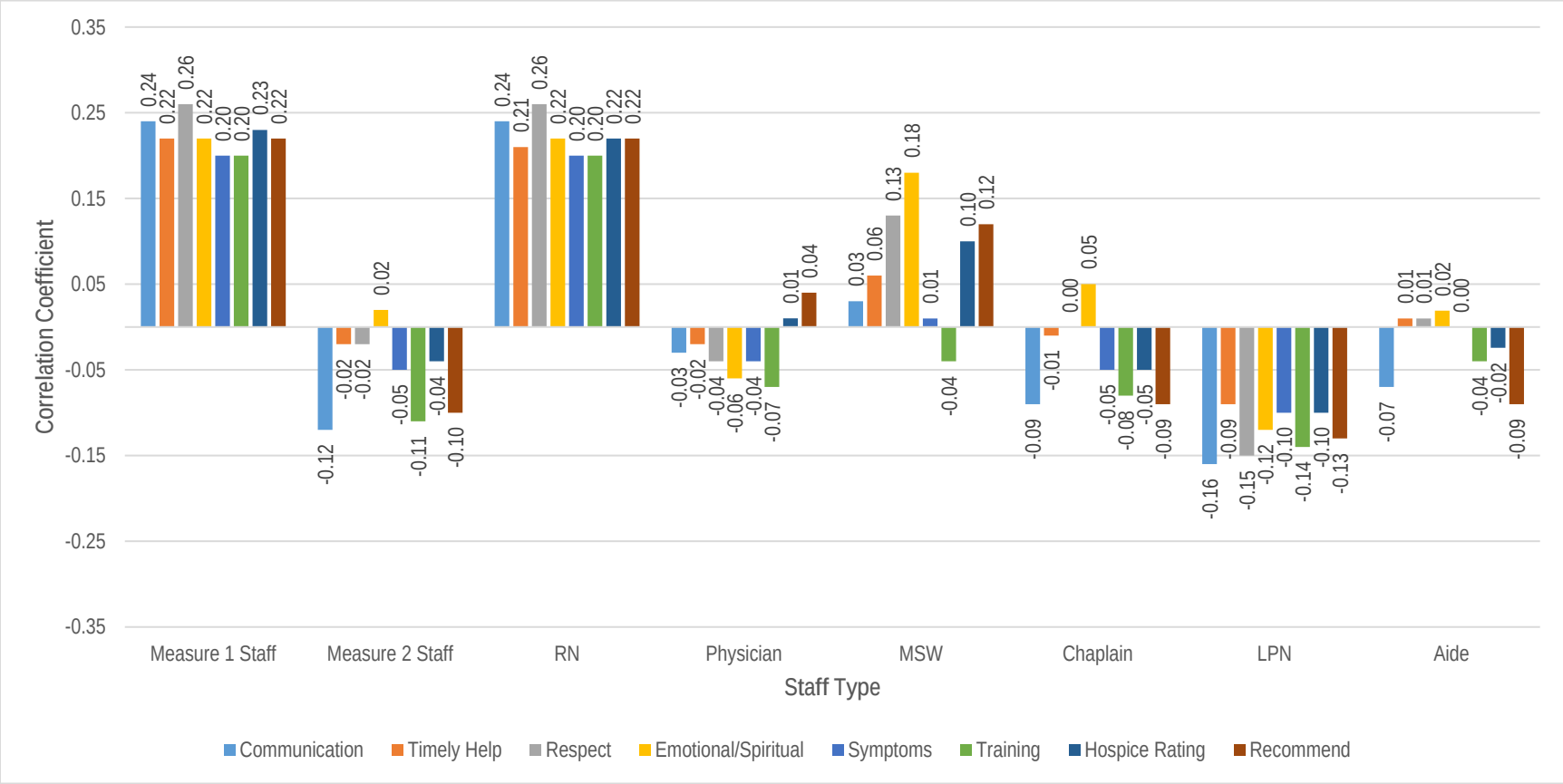
Hospice Visits When Death is Imminent: Measure 2 Testing Results

Correlation coefficients are summarized by staff type for the last three days of life and last seven days of life in **Exhibit 1** and **Exhibit 2**, respectively. Greater height of the bars (more positive values of the coefficients) indicates a stronger positive relationship between rates of visits when death is imminent and the CAHPS® Hospice Survey outcomes. Overall, Measure 1 exhibited the expected positive correlation with CAHPS® Hospice Survey outcomes. Measure 2, however, displays a negative correlation with CAHPS® Hospice Survey outcomes, suggesting receipt of these visits at the end of life is associated with poorer caregiver-reported experiences of care.

CMS also estimated correlation coefficients between the CAHPS® Hospice Survey outcomes and the individual staff types included in the measure pair. With regards to clinical staff, RN visits positively correlate with all of the with CAHPS® Hospice Survey outcomes, but physician visits show weak correlation with these outcomes. The combination of Measure 1 disciplines (RN, physicians, nurse practitioners or physician assistants), correlates positively with the CAHPS Hospice Survey outcomes – likely because a large percentage of total visits in the last days of life (52 percent in 2018) are from RNs.

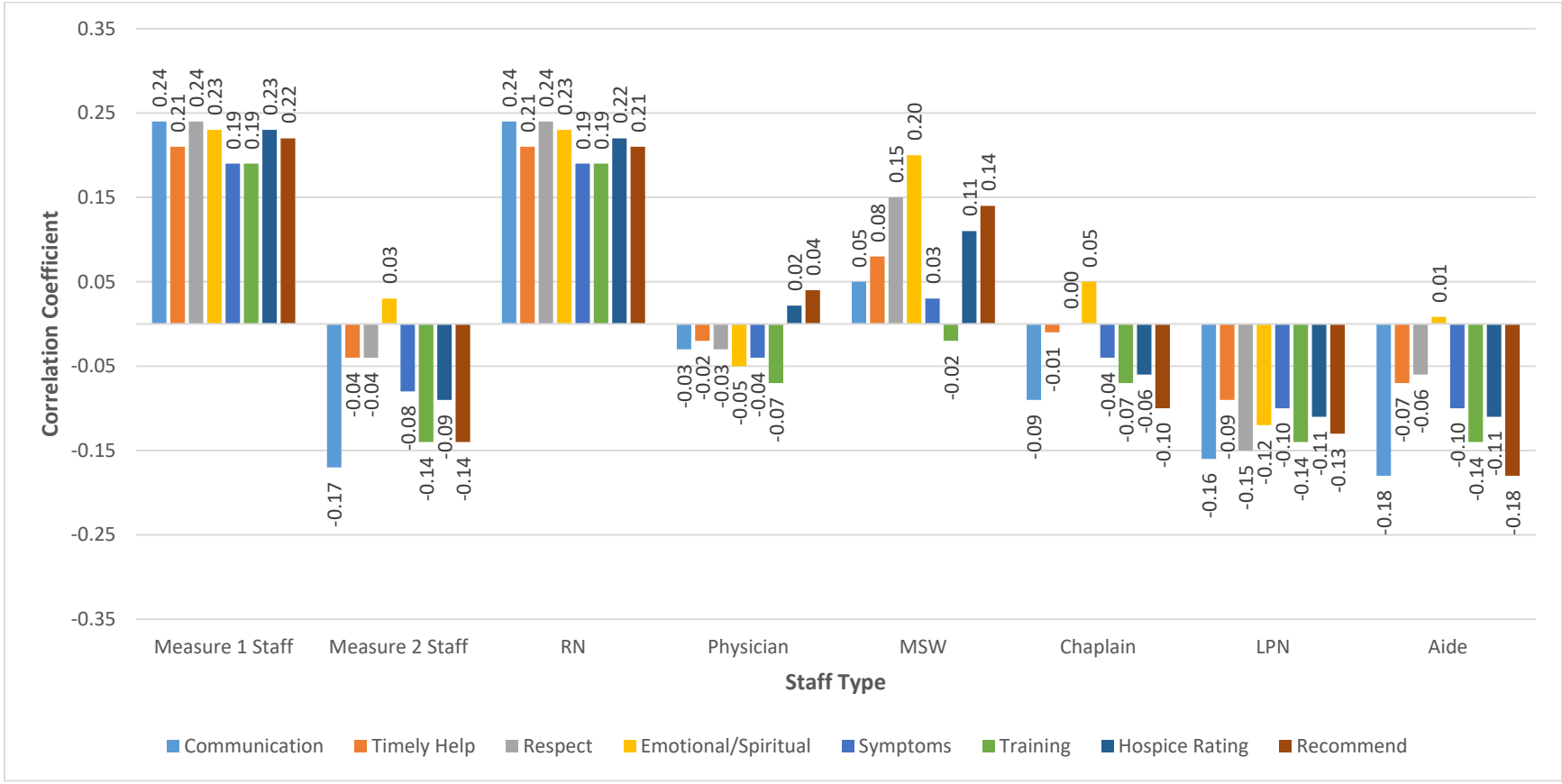
The other staff type with strong evidence of a positive correlation to CAHPS® Hospice Survey outcomes was MSW visits. This result differed from those for the other non-clinical Measure 2 staff, who generally displayed negative correlations with CAHPS® Hospice Survey outcomes individually, as they did for Measure 2 overall. For example, chaplain visits show negative correlations with CAHPS® Hospice Survey outcomes, except for the emotional and spiritual support CAHPS® Hospice Survey outcome, as might be expected. LPN visits are negatively correlated with all CAHPS® Hospice Survey outcomes. Finally, home aide visits, except for some statistically insignificant correlations, are mostly negatively correlated, particularly in the last seven days of life compared to the last three days of life.

Exhibit 1: Correlation Coefficients between Visits by Hospice Staff in the Last Three Days of Life and CAHPS® Hospice Survey Outcomes



Note: Not all correlations are statistically significant. Analyses of CAHPS® Hospice Survey Outcomes and Hospice Item Set Records, 2017-2018.

Exhibit 2: Correlation Coefficients between Visits by Hospice Staff in the Last Seven Days of Life and CAHPS® Hospice Survey Outcomes



Note: Not all correlations are statistically significant. Analyses of CAHPS® Hospice Survey Outcomes and Hospice Item Set Records, 2017-2018.

Hospice Visits in the Last Days of Life

The primary conclusion from Measure 2 testing was that RN and MSW social worker visits yielded evidence of a positive correlations to CAHPS® Hospice Survey outcomes. Coincidentally, these are the two staff types CMS was seeking to incentivize through the Service Intensity Add-On (SIA) payment policy implemented in 2016. Therefore, these results presented CMS with the opportunity to re-specify the HVWDII measure and achieve improved testing results with a specification better aligned with existing payment policy. Moreover, records of RN and MSW visits are fully available from hospice claims, eliminating the need for data collection through the HIS, thereby reducing burden on providers.

With the possibility for improvement, CMS sought to determine whether it could re-specify the HVWDII measure pair to better meet readiness standards for future public reporting and to minimize provider burden by using a claims-based approach. CMS solicited subject matter expertise on an SIA-aligned claims-based Hospice Visits measure. Clinical experience suggested that CMS explore the receipt of RN and/or MSW visits near death (the last 2, 3, or 4 days of life), when patients are most likely exhibiting the extreme symptoms of actively dying. This time interval is also consistent with studies analyzing the onset of clinical signs of dying, which confirm the increased symptom burden observed in the last 2 to 3 days of life.[6] Additionally, in their experience with caregiver focus groups, our experts noted consistent visits at the end of life were viewed as better care by hospice patient's families; e.g., a visit each day in the last 3 days of life is better than just one visit in the last 3 days of life.

CMS proceeded to test multiple versions of a re-specified claims-based Hospice Visits measure to determine which version performed the best on reportability, variability, validity, and reliability analyses. CMS tested 11 versions of a claims-based Hospice Visits measure:

- Measures 1 & 2: claims-based versions of the respective current HVWDII measure pair[7]
- Measures 3-5: receiving any RN or MSW visits in the last 2/3/4 days of life, respectively
- Measures 6-8: receiving an RN or MSW visit every day for the last 2/3/4 days of life, respectively
- Measures 9-11: receiving an RN or MSW visit every day for the last 2/3/4 days of life, respectively, allowing for one missed day in each (e.g., visits would only be required in two of the last three days of life to meet the measure criteria)

All versions of the measure yielded similar reportability and reliability results, and as a result, these analyses are omitted from this report. Variability and validity analyses are presented below.

Variability analyses

As shown in **Exhibit 3**, analyses indicated improved variability to distinguish between high- and low-quality hospices, compared to the HVWDII measure pair. The table lists each of the eleven measure and their respective numerator, denominator, and exclusion statements. The right-most column of the table indicates the interquartile range of the distribution of measure scores. A greater value indicates there is

---

[6] Hui D, dos Santos R, Chisholm G, Bansal S, Buosi Silva T, Kilgore K et al. 2014. Clinical Signs of Impending Death in Cancer Patients. *The Oncologist*; 19(6):681-687. Accessible via: http://theoncologist.alphamedpress.org/content/19/6/681.long.

[7] Our claims-based versions are imperfect replications of the HIS construct, as physician assistant and chaplain visits are not available in claims records. However, HIS data shows that the receipt of such services is relatively small, so the impact of these omissions should be minimal.

more spread in measure scores between the 25th and 75th percentiles of the distribution of scores. An increased range for a measure specification indicates that the measure can better differentiate one hospice from another. Scores are asterisk-coded for general reference, *** = high variability, ** = fair variability, and * = low variability. Groupings were assigned by relative position in the distribution (four in the high variability group, four in the fair variability group, and three in the low variability group).

The claims-based hospice visits measures defined with visits every day (or every day allowing for one missed day) had higher variability scores than the HIS or one RN/MSW visit versions, suggesting their superiority as differentiators. Measures 6, 7, 10, and 11 have the highest scores, ranging from 0.277 to 0.332. In contrast, the HVWDII measures yield poorer variability scores, 0.122 and 0.198, respectively.

Exhibit 3: Alternative Hospice Visit Measure Specifications, Variability Testing Results

Measure Version	Numerator	Denominator	Exclusions	Interquartile Range
Measure 1 (HVWDII)	Any physician/RN visits last 3 days	All patients discharged to death	Small hospices (<20 reporting patients); Any non-RHC days in last 3	0.122*
Measure 2 (HVWDII)	At least two (2+) MSW/LPN/Aide visits last 7 days	All patients discharged to death	Small hospices (<20 reporting patients); Any non-RHC days in last 7; one-day LOS	0.198**
Measure 3	Any RN/MSW last 2 days	All patients discharged to death	Small hospices (<20 reporting patients); LOS<2 days; Any non-RHC days in last 2	0.184**
Measure 4	Any RN/MSW last 3 days	All patients discharged to death	Small hospices (<20 reporting patients); LOS<3 days; Any non-RHC days in last 3	0.109*
Measure 5	Any RN/MSW last 4 days	All patients discharged to death	Small hospices (<20 reporting patients); LOS<4 days; Any non-RHC days in last 4	0.069*
Measure 6	Consistent visits last 2 days by RN or MSW	All patients discharged to death	Small hospices (<20 reporting patients); LOS<2 days; Any non-RHC days in last 2	0.332***
Measure 7	Consistent visits last 3 days by RN or MSW	All patients discharged to death	Small hospices (<20 reporting patients); LOS<3 days; Any non-RHC days in last 3	0.277***
Measure 8	Consistent visits last 4 days by RN or MSW	All patients discharged to death	Small hospices (<20 reporting patients); LOS<4 days; Any non-RHC days in last 4	0.217**
Measure 9	Visits by RN or MSW at least 1 of last 2 days of life	All patients discharged to death	Small hospices (<20 reporting patients); LOS<2 days; Any non-RHC days in last 2	0.184**
Measure 10	Visits by RN or MSW at least 2 of last 3 days of life	All patients discharged to death	Small hospices (<20 reporting patients); LOS<3 days; Any non-RHC days in last 3	0.279***
Measure 11	Visits by RN or MSW at least 3 of last 4 days of life	All patients discharged to death	Small hospices (<20 reporting patients); LOS<4 days; Any non-RHC days in last 4	0.300***

Note: Analyses of Medicare Hospice claims for patients with dates of death during Federal Fiscal Year 2018. *** = high variability, ** = fair variability, and * = low variability.

Validity Analyses

Validity analyses included tests of correlations between measure specifications and provider-level CAHPS® Hospice Survey outcomes, as summarized in **Exhibit 4**. For brevity, the exhibit only displays correlation coefficients between each tested specification and the hospice rating and recommendation

CAHPS® Hospice Survey outcomes, although all CAHPS® Hospice Survey outcomes were tested and results were generally consistent.

The measures defined with hospice visits every day (or every day allowing for one missed day) had higher correlations with CAHPS® Hospice Survey outcomes, with Measures 10 and 11 having the highest correlations of all. Correlations with the overall rating CAHPS® Hospice Survey outcome was 0.26 for Measure 10 and 0.28 for Measure 11, and for the recommendation outcome was 0.24 for Measure 10 and 0.26 for Measure 11. Note that the validity scores of these measures exceed those of the HVWDII Measures 1 and 2. Measure 2 remains the lone specification to have a negative correlation with CAHPS® Hospice Survey outcomes, as previously shown in **Exhibits 1 & 2**.[8]

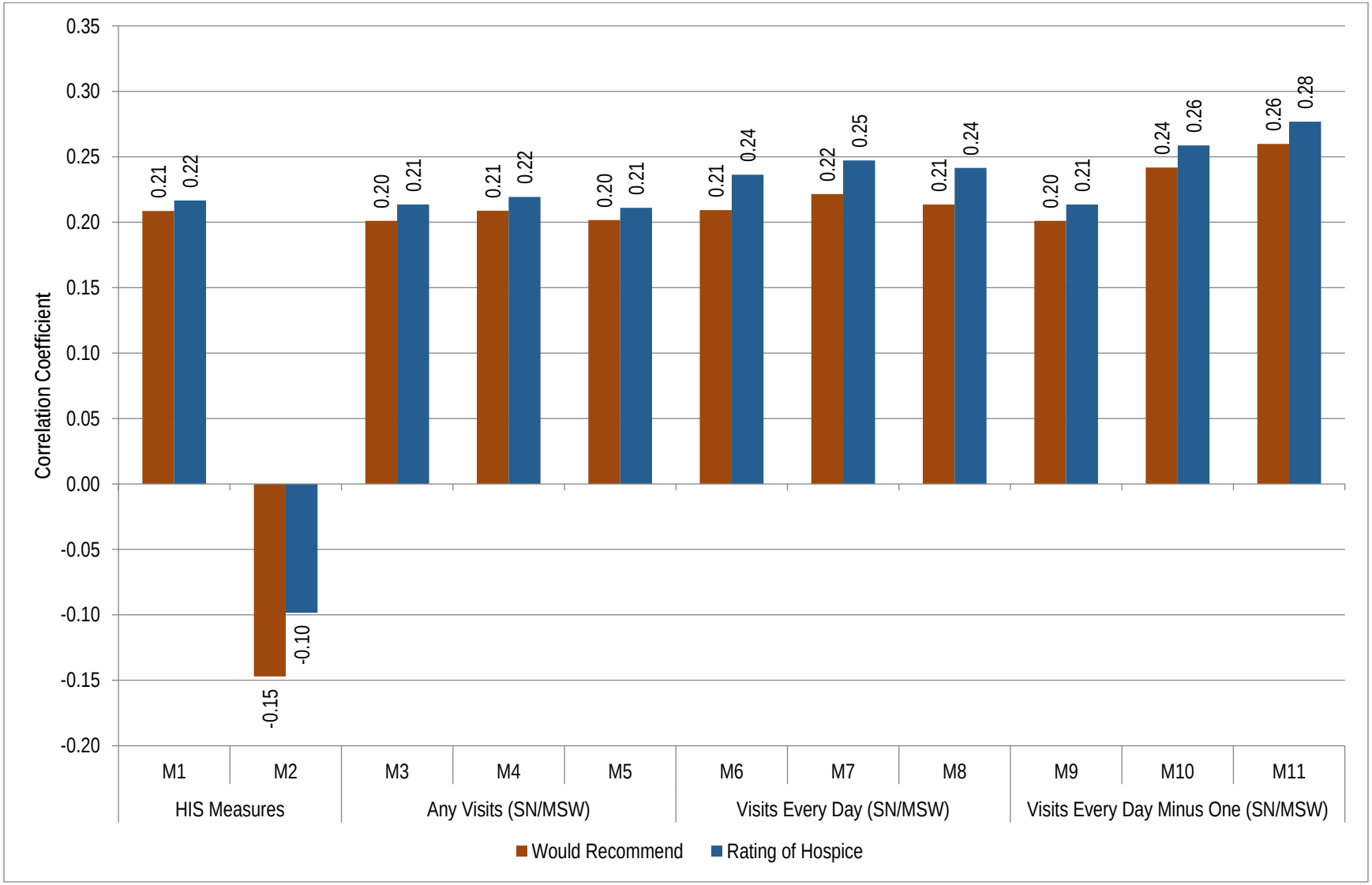
Selecting a New Specification

Measure versions 10 or 11 of the re-specified claims-based Hospice Visits measure (consistent visits in the last 3 or 4 days of life by an RN and/or MSW, allowing for one missed day) best meet the testing results. To be consistent with the evidence of increased highly specific physical signs associated with death occurring in the last 3 days of life,[9] CMS chose to proceed with the development of Measure 10, Hospice Visits by RN or MSW in at least 2 of the last 3 days of life. To distinguish this measure from the HVWDII measure pair, CMS named the new specification "Hospice Visits in the Last Days of Life" (HVLDL).

[8] Note that correlation estimates are slightly different between **Exhibits 1 & 2** and **Exhibit 3** due to different data sets and time periods.

[9] Hui D, dos Santos R, Chisholm G, Bansal S, Buosi Silva T, Kilgore K et al. 2014. Clinical Signs of Impending Death in Cancer Patients. *The Oncologist*; 19(6):681-687. Accessible via: http://theoncologist.alphamedpress.org/content/19/6/681.long.

Exhibit 4: Alternative Hospice Visit Measure Specifications, Validity Testing Results



Note: Analyses of CAHPS® Hospice Survey Outcomes and of Medicare Hospice claims for patients with dates of death during Federal Fiscal Year 2018. "M1" = Measure 1 (specification), etc.

Conclusion

Overall, positive correlations were present between CAHPS® Hospice Survey outcomes and scores from the HVWDII Measure 1 (visits provided by clinical staff) but not with Measure 2 (visits provided by non-clinical staff). With the exception of MSW visits, higher receipt of Measure 2 staff visits were found to be associated with poorer caregiver experiences of hospice care, as reported in the CAHPS® Hospice Survey outcomes scores.

Based on the above analyses, CMS sought to develop a newly specified visits measure, focusing on RN and MSW visits, which were the staff types found to have the greatest positive correlation with CAHPS® Hospice Survey outcomes. Coincidentally, the RN and MSW visits are incentivized by the SIA Medicare hospice payment policy. Because the visits of both staff types can be obtained from existing claims records alone, this eliminates the need for data collection through the HIS, and would thereby reduce provider burden. The new claims-based version of the concept was named the Hospice Visits in the Last Days of Life (HVLDL) measure. Testing during the development of the HVLDL measure replicated the strong correlation between CAHPS® Hospice Survey outcomes and both RN and MSW visits near death, and poorer correlations with home aides. The HVLDL exhibited greater potential as a differentiator than the HVWDII measure pair, suggesting a greater potential for usefulness to consumers.

Ultimately, measure testing of HIS, hospice claims, and CAHPS® Hospice Survey outcomes led CMS to a claims-based, re-specified version of the HVWDII measure pair. This new measure shows performance improvement over Measure 2 (which did not meet public reporting standards) but even over Measure 1 (which did meet these standards). The adoption of the HVLDL measure will add value to the HQRP in terms of providing a stronger-performing process measure calculated at reduced burden to the public.