Federation of American Hospitals®

Charles N. Kahn III
President & CEO

June 26, 2018

The Honorable Seema Verma, Administrator
Centers for Medicare and Medicaid Services
Department of Health and Human Services
Hubert H. Humphrey Building, Room 445-G
200 Independence Avenue, SW.
Washington, D.C. 20201

***RE: CMS-1688-P, Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2019; Proposed Rule (Vol. 83, No. 89), May 8, 2018***

Dear Administrator Verma:

The Federation of American Hospitals (FAH) is the national representative of more than 1,000 investor-owned or managed community hospitals and health systems throughout the United States. Our members include teaching and non-teaching hospitals in urban and rural America, as well as inpatient rehabilitation, psychiatric, long-term acute care, and cancer hospitals. The FAH appreciates the opportunity to comment to the Centers for Medicare & Medicaid Services (CMS) about the referenced Notice of Proposed Rulemaking on ***Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2019; Proposed Rule (Vol. 83, No. 89), May 8, 2018.***

**IV. Facility-Level Adjustment Factors**

The FAH strongly recommends that the CMS monitor and report on these factors annually and adjust them if a material change is noted. CMS should provide as part of its annual rulemaking a detailed analysis of the Agency's review justifying either a continued freeze or an update to the adjustment factors. Finally, we ask that CMS adjust all three factors at a minimum once every three years in order to maintain payment accuracy. This will help ensure a dynamic and accurate IRF payment system that recognizes and responds to change in the cost of care and promotes the delivery of efficient and effective IRF services.

## V.D. Proposed Wage Adjustment for FY 2019

Consistent with our comments to the FY 2018 IRF PPS proposed rule, the FAH requests changes to the IRF wage index policy that recognize the realities of and promotes a modern, competitive labor marketplace. This is particularly important in light of the accelerating movement towards alternative payment models that remove the barriers separating payment systems based on site of care. Along those lines, CMS should make the IRF wage index concurrent with other post-acute care settings as well as acute care hospitals including enabling IRFs to obtain geographic reclassifications and basing the IRF wage index on the same current year pre-classified wage index value as acute care hospitals rather than prior year values.

## VI. Proposed Update to Payments for High-Cost Outliers Under the IRF PPS for FY 2019

The proposed rule would increase the high-cost outlier threshold from $8,679 in FY 2018 to $10,509 in FY 2019, a significant 21 percent increase. As recently as FY 2017, the threshold was only $7,984. The FAH supports an outlier pool no greater than three percent and notes that in recent years CMS appears to have paid out the full amount indicating that the fixed loss threshold has achieved its target when viewed across all hospitals. According to an analysis of CMS FY 2019 rate-setting files, however, the top ten decile IRFs will receive over 57 percent of IRF outlier payments, a disproportionately high amount. In addition, there appears to be a disconnect between patient acuity, efficiency, and outlier payments, all of which suggests that CMS consider imposing a cap on outlier payments a hospital could receive.

# FIM™ INSTRUMENT AND CASE-MIX CLASSIFICATION SYSTEM

## VII. Proposed Removal of the FIM® Instrument and Associated Function Modifiers from the IRF-PAI Beginning with FY 2020 and Proposed Refinements to the Case-Mix Classification System Beginning with FY 2020

FAH opposes the CMS proposal to remove the FIM® instrument and revise the Case-Mix Classification System for FY 2020 as described. Revising the IRF PPS at this time to set rates and determine payments using unproven assessment items with known rater reliability problems threatens the validity and stability of the IRF PPS, increasing the likelihood of a need for future system rebasing or possibly legislative fixes and risking beneficiary access to restorative care in the meantime. While we appreciate CMS's efforts to reduce burden on clinicians and to begin to align measurement of patient functionality across post-acute care settings, we are troubled that the proposed changes could well put patients at risk by introducing instability into the IRF payment system and threatening access to IRF care for the clinically complex patients who require it. In addition to data quality issues, we are concerned about the lack of clinical validation and transparency in the process of creating the proposed Case-Mix Classification System.

CMS did not issue an impact file with this proposed rule, and the relevant assessment data are not otherwise publicly available. The FAH engaged consulting firm Dobson DaVanzo and Associates (Dobson | DaVanzo) to examine the potential impacts of the proposed changes

(see Attachment A) using deidentified FY 2017 case data from the Uniform Data System for Medical Rehabilitation (UDSMR). This data represents 84 percent of IRF cases used in the proposed FY 2020 Case-Mix Classification System revision. From our perspective, this report highlights the following:

1. Incongruence between the current and proposed system in that a significant number of cases that were in a particular CMG under the FIM® motor scores shift well beyond reasonable explanations to other CMGs under the proposed system. This leads to significant regrouping of patients and is accompanied by changes in case payments and average length of stay for the CMG, which implies that the same patient would be treated very differently under the proposed system. CMS did not provide adequate justification for the degree of shift in patient treatment protocols implied from the change from one system to the next.
2. Fundamental differences in the assessment of functional status between the current and proposed system, changing basic clinical implications.
    a. The proposed assessment items are based on patients' "usual performance" on a given item within the assessment period. In comparison, FIM® responses are based on patients' lowest performance, or highest burden of care, on a given item. CMS did not indicate or demonstrate that this change would improve payment system accuracy or address the difference in clinical or payment outcomes this imparts on the CMGs.
    b. Fewer scaled responses are available under the proposed assessment items compared to the FIM®. The proposed assessment response items are definitionally broader to attempt to capture a similar range of impairment as the FIM® instrument measures, which reduces the ability of the instrument to clinically distinguish between patients at admission. This also decreases the ability of the assessment tool to capture functional improvement over the course of treatment.
        i. This is a shift away from granularity in assessing patient functionality, as noted by a narrower, more concentrated distribution of the proposed motor scores. In the proposed system, patients may appear to be less severe than they show in the current FIM® system, and changes in functionality from admission to discharge may not be captured with as much specificity. This shift moves in the opposite direction that CMS has taken in other PPS systems, most notably the move to MS-DRGs in acute care hospitals, a move which FAH then commended. More granularity can enable providers to more carefully target resources to patient needs, leading to more efficient care.
    c. There is a much higher prevalence of non-scaled items under the proposed system than FIM® and this occurs more often per case. (Non-scaled responses are assessment item responses that are

subsequently rescaled to be the most impaired assessment response for the purposes of calculating a functional score for CMG assignment.) The existence of more missing items in the proposed motor score is of particular concern as this indicates potential gaps in and a general lack of understanding of the proposed system. It is concerning that for a particular patient, clinicians were able to derive a FIM® score for a case, but not for the equivalent functional item in the proposed system, signifying that the proposed items are not yet understood by clinicians.

d. The exclusion of key elements of the current system (namely, cognitive items and motor score item weights) and decision of which motor impairment items to include was not strongly justified and has unclear effects on the payment system and ultimately patient care.
    i. CMS has not proposed to weight the new motor score at this time despite having substantial justification for doing so under the current system. Implementing an unweighted system now, then adding weights in the future may add further confusion as well as disruptions in operations and patient care protocols, among other consequences, as providers adjust to different payment incentives under a rapidly evolving functional scoring system.
    ii. Further, CMS did not provide evidence of careful analysis to justify the removal of cognitive assessment items. Removal of the cognitive items is counter-intuitive as they are predictive of resource use under the current Case-Mix Classification System and contrary to prior technical expert panel meetings, which highlighted the importance of cognitive items.
    iii. The choosing of particular motor items for inclusion in the proposal was only weakly justified. For example, some shorter walking items were chosen because patients ostensibly have an easier time performing the test; however, the longer walk items are included in the FIM® motor score and often have a scaled response.

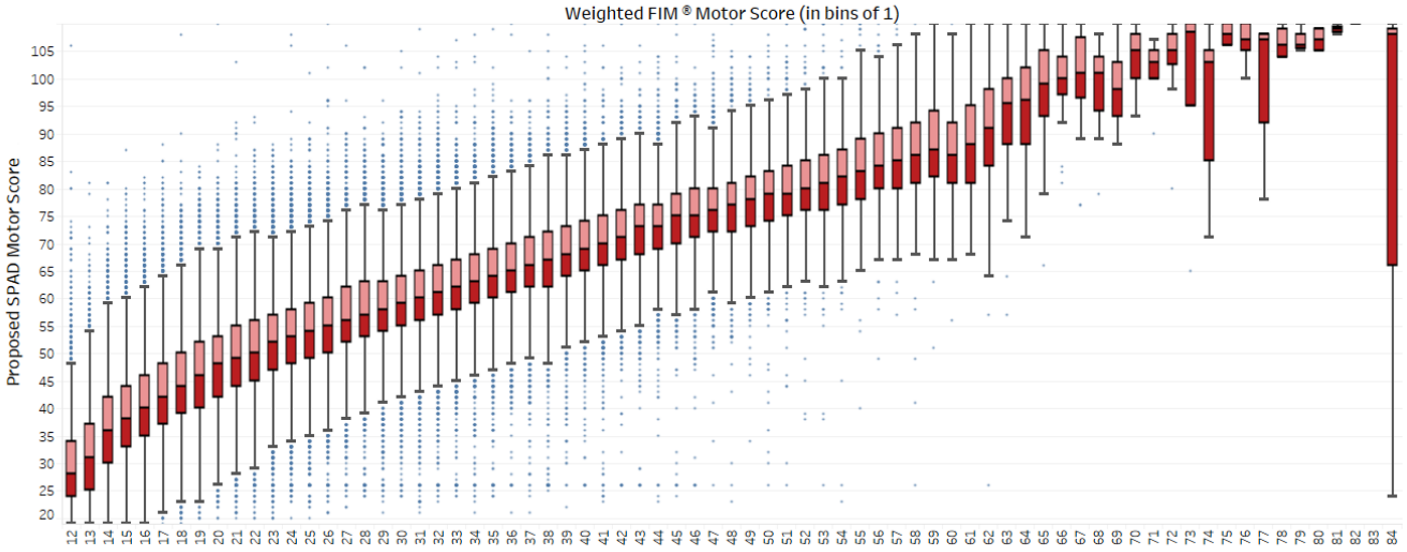
3. Multiple data quality issues.
    a. The assessment items on which the proposed system was built are new as of October 2016, the start of data collection of items included in the proposed revisions. Clinicians were not yet well-versed in scoring the new items, which use a different measurement scale and clinical construct than the familiar FIM® instrument. Indeed, CMS repeatedly issued interpretational guidance revising measure construct definitions throughout FY 2017, making it effectively impossible to have strong inter-rater reliability, particularly across rehabilitation facilities.

b. An FAH member examined a sample of medical records and found low inter-rater reliability and poor fidelity on the new assessment items in relation to the medical record despite strong reliability among FIM® responses.
c. As clinicians better understand the differences between "usual" and worst performance, this will have an impact on how they assess patients, and thus their patients' eventual assessment score. However, better assessment scoring may well reduce the number of non-scaled responses and would likely drive down payments. Given the newness of the assessment items and their likely continued changes, rate setting and case-mix groups could be a moving target over several years. This could compromise year-over-year budget neutrality for reasons not associated with coding quality.

As shown in Exhibit 1, the proposed assessment scale does not allow one to definitively link the existing level of functionality to that proposed. Using the proposed assessment items and scale, it appears to be possible for any given patient under a FIM® scale assessment to fall into a broad range of functional scores in the proposed system. We caution CMS against moving from the current form of functional assessment to the proposed system in the near term, as this has unknown impacts on patient care and quality outcomes.

**Exhibit 1: Distribution of FY 2017 Cases by Weighted FIM® Motor Score and Proposed SPAD Motor Score**

*Each "box" shows the 25th and 75th percentile SPAD motor scores for each FIM® motor score (truncated to closest integer); the midpoint of each box (where the light and dark red meet within each box) shows the 50th percentile of the SPAD motor score for each IRF cases' FIM® motor score. The "whisker" (lower and upper-most bound of each box plot) extend to values within 1.5 times the inter-quartile range. The dots represent cases that fall outside of the range defined by + or – 1.5 times the interquartile range and show the minimum and maximum values.*




**Score**

Source: Dobson DaVanzo analysis of UDSMR FY 2017 cases.

Ultimately, it is the beneficiaries who will face the consequences of CMG, length of stay, and payment confusion and instability as the proposal would alter both how patients are perceived and how they access care as providers seek to adapt to these changes. As the new

system reflects the above considerations, the functional status score for any given individual is likely to change. The implications are that this patient will likely be given different treatment for their care and that payment for this care may or may not reflect the patient's acuity and the resources required to provide the care she needs.

If these elements are not addressed prior to implementing this new system, unintended consequences, including reduced access to care for IRF patients, are likely. Accordingly, the FAH urges CMS to delay the implementation of the proposed changes beyond the proposed FY 2020 start date, and instead carefully examine and address the data quality issues reported. The quality of the FY 2017 data underlying the calibration of the new assessment tool is simply not sufficient to support the movement to a new IRF PPS as proposed. Indeed, when CMS does move forward with a new patient assessment and case-mix system, it would be prudent to base it on no less than two years of data to help increase accuracy. The very recent implementation of these new items reveals a lack of clarity on the precise definition of "usual performance," as well as a preponderance of non-scaled items, and demonstrates that it is premature to adopt this new payment system in FY 2020.

We recommend that CMS continue to work with the provider community to ensure common and adequate understanding of the clinical measurement approach and to solicit input on their concerns regarding the proposed system. For instance, we recommend that CMS collaborate with clinical experts to explore how changing assessment item measurement constructs affects the clinical interpretation of the case-mix system. Any new measurement approach must yield trusted data that IRFs have confidence will yield the right results for patients. Assessment data whose quality and accuracy cannot meet that standard should not be included in the payment system and should not serve as the basis for a major overhaul. The data and analysis that CMS has put forward does not yet meet that test.

Finally, once the proposed changes have become ready for implementation, we encourage CMS to provide comprehensive and clear guidance (in the form of written documentation, power point presentations, and live webinars) to educate clinicians and answer their questions as they learn about and adjust to the proposed assessment items.

## COVERAGE REQUIREMENTS

### VIII. Proposed Revisions to Certain IRF Coverage Requirements Beginning With FY 2019

To reduce regulatory burden on rehabilitation providers and physicians, CMS is proposing to revise several IRF coverage criteria in response to comments submitted pursuant to the request for information (RFI) in the FY 2018 IRF PPS proposed rule. FAH's members appreciate CMS's careful consideration of the responses to that RFI, and the proposals in the FY 2019 IRF PPS proposed rule to reduce administrative burden associated with documentation requirements. Comments on each of the proposals are detailed below.

**VIII.A. Proposed Changes to the Physician Supervision Requirement**

CMS is proposing to allow the post-admission physician evaluation to count as one of the required three weekly face-to-face weekly physician visits. The FAH supports this proposal and agrees that the rehabilitation physician should have the flexibility to assess the patient as well as conduct the required post-admission physician evaluation.

**VIII.B. Proposed Changes to the Interdisciplinary Team Meeting Requirement**

Current regulations permit remote team conferencing supported by proper documentation. CMS is proposing to relax this documentation requirement in order to facilitate remote participation. The FAH supports this proposal allowing remote participation by the physician, when necessary, though we recommend CMS indicate a preference that the physician be physically present unless circumstances make that impractical.

**VIII.C. Proposed Changes to the Admission Order Documentation Requirement**

CMS is proposing to remove the requirement at §412.606(a) that, at the time the patient is admitted, an IRF must have physician orders for the patient's care during the time the patient is hospitalized. The FAH supports this proposal and concurs with CMS's assessment that the current requirement is duplicative, as this is already required under the hospital conditions of participation (CoPs) and the admission order payment requirements.

**VIII.D. Solicitation of Comments Regarding Additional Changes to the Physician Supervision Requirement**

When IRF coverage criteria were initially implemented in 2010, CMS believed that the rehabilitation physician visits should be completed face-to-face by a rehabilitation physician to ensure that the patient receive the most comprehensive care throughout the IRF stay. CMS is now interested in further exploring this requirement as part of its ongoing efforts to reduce unnecessary regulatory burden on IRFs. CMS specifically requests comments on whether the rehabilitation physician should have the flexibility to determine when the patient needs to be assessed face-to-face and when the assessment can be successfully accomplished remotely via an alternative mode of communication, such as video or telephone conferencing.

While the FAH appreciates and encourages CMS's efforts to reduce regulatory burden, we are concerned that remote face-to-face physician visits could dilute what is a hallmark and distinguishing characteristic of IRF care – frequent, "hands-on" patient care administered and overseen by highly-skilled rehabilitation physicians. This is the standard of care required to treat these clinically complex patients; a standard that IRFs are uniquely able to provide. CMS, however, could consider limited circumstances under which remote visits might be permitted in rural facilities that have difficulty securing physician coverage.

### VIII.E. Solicitation of Comments Regarding Changes to the Use of Non-Physician Practitioners

CMS is seeking information on whether non-physician practitioners, such as physician assistants, could fulfill some of the duties currently required to be completed by a physician.

As noted above, highly-trained and experienced rehabilitation physician engagement with the clinically complex patients who require IRF care is what differentiates IRFs from lower level of care settings such as SNFs. Permitting non-physician practitioners to perform face-to-face visits or to lead team conference discussions, for example, among other critical physician functions, would dilute and compromise the intensity and quality of IRF services to the detriment of patients who need that care.

## QUALITY DATA REPORTING

## IX. PROPOSED REVISIONS AND UPDATES TO THE IRF QUALITY REPORTING PROGRAM (QRP)

### IX.B. General Considerations Used for the Selection of Measures for the IRF QRP

Accounting for Social Risk Factors in the IRF QRP

The FAH supports CMS' commitment to continue working with ASPE, the public, and other key stakeholders to identify policy solutions that achieve the goals of attaining health equity for all beneficiaries and minimizing unintended consequences. The FAH continues to urge CMS to adopt the use of SDS factors into the risk adjustment methodology for resource use measures, such as Medicare Spending Per Beneficiary (MSPB), because there is a strong tie between SDS patient characteristics and the relative amount of resources required to provide adequate care.

Recent medical literature indicates that patient characteristics that are not included in any Medicare risk adjustment framework, such as SDS factors, drive much of the difference in readmission risk between patients (a major driver of episode spending variation, and therefore of variation in resource use).[1] Additionally, the National Quality Forum (NQF) in the midst of a two-year trial program of assessing risk adjustment of performance measures for SDS factors.

### IX.C. Proposed New Removal Factor for Previously Adopted IRF QRP Measures

The FAH commends CMS for its proposed application of the Meaningful Measures initiative to the Inpatient Rehabilitation (IRF) Quality Reporting Program (QRP). Prioritizing and reducing the number of quality measures across these programs addresses our previously

[1] *See* Michael L. Barnett, MD, John Hsu, MBA, J. Michael MacWilliams, MD, PhD, *Patient Characteristics and Differences in Hospital Readmission Rates*, JAMA INTERNAL MEDICINE, Vol. 175(11), 1803-1812 (Nov. 2015), http://archinte.jamanetwork.com/article.aspx?articleid=2434813; *see also* Herb Kuhn and David Nerenz, *It's Time to Add Socio-demographic Factors When Weighing Quality Performance*, Modern HealthCare (Feb. 6, 2016), http://www.modernhealthcare.com/article/20160206/MAGAZINE/302069978.

expressed concerns about the burden of managing many measures and the unnecessary duplication of measures across programs. The FAH supports a focus on measures designed for improving patient care and working towards outcomes that are meaningful to patients.

In the proposed rule CMS proposed to adopt an eighth quality measure removal factor for the IRF QRP. This new quality removal factor aligns across both programs and would serve to remove measures where "the costs associated with a measure outweigh the benefit of its continued use in the program". The FAH supports the proposal to add an eighth factor, identified as the cost associated with a measure outweighing the benefit of its continued use in the program, to the lists of factors used for considering removal of measures from the IRF QRP. This proposed new factor is appropriate for moving toward measure sets that meet the goal of streamlining measures with a focus on those that will work toward the best outcomes for patients. The FAH appreciates that CMS has identified costs beyond those associated with data collection and submission. The costs associated with tracking performance and investing resources for quality improvement should be considered as well. It would be useful for CMS to clarify in the final rule the nature of the burden that the removal of a measure relieves, the methods or criteria used to assess when the measure cost or burden outweighs the benefits of retaining it.

**IX.E. Proposed Removal of Two IRF QRP Measures**

The FAH supports the proposed removal of two measures from the IRF QRP measure set.

The FAH agrees with CMS that the Facility-wide Inpatient Hospital-onset Methicillin-resistant Staphylococcus aureus (MRSA) Bacteremia Outcome Measure (NQF #1716), from the IRF QRP measure set beginning with the FY 2020 IRF QRP as the number of IRFs with expected MRSA infections that have sufficiently high incident rates of MRSA infection incidence ratio is too low to calculate reliable measure rates worth publicly reporting.

The FAH supports the removal of the measures Percent of Residents or Patients Who Were Assessed and Appropriately Given the Seasonal Influenza Vaccine (Short Stay) (NQF #0680), from the IRF QRP beginning with the FY 2021 IRF QRP as performance among IRFs for this measure is so high and unvarying that meaningful distinctions in improvement in performance can no longer be made.

**IX.F. IMPACT Act Implementation Update**

In the FY 2018 IRF PPS final rule, CMS states it intends to specify two measures that would satisfy the domain of accurately communicating the existence and provision of the transfer of health information and care preferences under section 1899B(c)(1)(E) of the Act no later than October 1, 2018 and intended to propose to adopt them for the FY 2021 IRF QRP with data collection beginning on or about October 1, 2019. Given comments CMS received during a public comment period between November 10, 2016 and December 11, 2016, input provided by a technical expert panel (TEP) convened by the CMS-1688-P contractor, and pilot measure testing conducted in 2017, CMS is engaging in continued development work on these two

measures and is proposing to delay measure specification by up to one year. While the FAH supports continuous improvement in the nature of the quality measures, we request that CMS provide additional detail in the final rule with respect to the granting of extensions under the IMPACT Act when statutory deadlines are stipulated.[2]

**IX.H. Proposed Changes to Reconsiderations Requirements Under the IRF QRP Section**

Section 412.634(d)(1) of CMS's regulations states, in part, that IRFs found to be noncompliant with the quality reporting requirements for a particular fiscal year will receive a letter of non-compliance through the Quality Improvement and Evaluation System Assessment Submission and Processing (QIES-ASAP) system, as well as through the United States Postal Service.

CMS is proposing to revise §412.634(d)(1) to expand the methods by which they would notify an IRF of non-compliance with the IRF QRP requirements for a program year. Revised §412.634(d)(1) would state that CMS will notify IRFs of non-compliance with the IRF QRP requirements via a letter sent through at least one of the following notification methods: the QIES-ASAP system, the United States Postal Service, or via an email from the Medicare Administrative Contractor (MAC).

The FAH agrees that that this change will address the feedback from providers requesting additional methods for notification and supports adding additional notification systems to include the QIES-ASAP system. However, the FAH does not support using MACs to administer IRF QRP non-compliance notifications given the lack of expertise of MACs in the area of quality reporting. The FAH is concerned that this would add yet one more entity into an already complex infrastructure and likely will cause more problems than it will solve. Rather than inserting another entity into the QRP programs, the FAH urges CMS to work with its existing contractors on the QRP side (such as Cormac, which runs the QRP helpdesk) to ensure they can meet the demands of their current contracts while potentially expanding their role to encompass QRP non-compliance notification as well.

If CMS proceeds with this proposal as proposed, the FAH urges CMS to refrain from having MACs engage in actual QRP non-compliance determinations and instead only keep their role limited to transmittal of non-compliance decisions made directly by CMS.

**IX.I. Proposed Policies Regarding Public Display of Measure Data for the IRF QRP**

CMS is proposing to begin publicly displaying data on *IRF Compare* on the following four assessment-based measures in CY 2020, or as soon thereafter as technically feasible: (1) Change in Self-Care (NQF #2633); (2) Change in Mobility Score (NQF #2634); (3) Discharge Self-Care Score (NQF #2635); (4) and Discharge Mobility Score (NQF #2636). The FAH requests that CMS provide more information with respect to the manner they intend on displaying these measures. Until more detail on these specifications is provided FAH cannot comment on this proposal.

[2] SSA §1899B(c)(1)(E); U.S.C. 42 U.S.C. §1395*lll*(c)(1)(E).

**IRF Public Reporting - Procedures for the Opportunity to Review and Correct Data and Information**

IRFs receive two different types of confidential feedback reports for publicly reported measures – one report for assessment-based measures (e.g., CAUTI, pressure ulcers) and another for claims-based measures (DTC, PPR, MSPB, etc.). The assessment-based measure reports would include patient-level data and be available on a monthly basis, while the claims-based measure reports would apparently not include patient-level data (only "aggregate hospital-level data") and would be available only once annually. As FAH has commented numerous times on previous proposed rules, claims-based measures are of little value for quality improvement if the data cannot be provided more often than once per year and also include patient-specific data.

Hospitals need patient-level data to be able to identify why their claims-based measure rates are what they are. Timely patient-level data on claims-based measures such as PPRs and MSPB would empower IRFs to perform root-cause analyses and determine what could have been done better in a certain case or set of cases. However, only providing an annual facility-level rate in comparison to the national rate does not provide an IRF with any indication as to where or how to begin quality improvement activities. The FAH therefore requests to begin sharing patient-level feedback data on claims-based measures with IRFs, as it does with other providers in the Medicare program, as soon as possible.

## PROMOTING INTEROPERABILITY AND ELECTRONIC INFORMATION EXCHANGE

**X. Request for Information on Promoting Interoperability and Electronic Healthcare Information Exchange Through Possible Revisions to the CMS Patient Health and Safety Requirements for Hospitals and Other Medicare- and Medicaid-Participating Providers and Suppliers**

CMS is seeking feedback in the proposed rule on how it could advance the electronic exchange of information in support of care transitions between providers using: CoPs, CfCs, and RfPs for Long-Term Care Facilities. Specifically, CMS is considering revising these to require providers to electronically perform a variety of activities, including: transfer of medically necessary information from a hospital to another facility upon a patient transfer or discharge; transfer of discharge information from a hospital or post-acute provider to a community provider, if possible; and providing patients access to certain information via electronic means, if requested, including directing that information to a third-party application.

The FAH has long supported efforts to achieve comprehensive interoperability and data liquidity – the free flow of meaningful, actionable information that supports and enhances patient care within and across settings. Our members have a vested interest in data flow to improve patient care, workflow efficiencies and clinician satisfaction, population health and payment models, and research. However, the FAH does not support the proposed revision of the CoPs, CfCs, and RfPs related to interoperability and the exchange of health information. The current

ecosystem is simply not mature enough to facilitate the movement of this information, as evidenced by the obstacles that currently prevent seamless information exchange and would make it exceedingly difficult for hospitals and other providers to comply with the requirements. The FAH appreciates CMS's acknowledgement of this in the proposed rule, noting that, "While both adoption of EHRs and electronic exchange of information have grown substantially among hospitals, significant obstacles to exchanging electronic health information across the continuum of care persist. Routine electronic transfer of information post-discharge has not been achieved by providers and suppliers in many localities and regions throughout the Nation." 83 Fed. Reg. 21004.

These obstacles are amplified in the patient discharge and transfer arenas because post-acute providers and behavioral health providers were ineligible for the EHR Incentive Programs under the Health Information Technology for Economic and Clinical Health (HITECH) Act, which have been instrumental in enabling acute care hospitals to achieve so much of the potential that EHRs specifically and HIT generally offer. As such, post-acute providers and behavioral health providers have not been able to adopt HIT to the extent of hospitals and CAHs. Thus, were CMS to move forward with revisions to the CoPs, CfCs, and RfPs, hospitals and CAHs would be unable to meet these requirements because of the lack of providers available to accept that information electronically. And, for post-acute care and behavioral health providers, it would be unfair, and tantamount to an unfunded mandate, to require that these providers adopt and maintain expensive EHRs and other HIT through CoPs, CfCs, and RfPs when they receive no corresponding financial assistance to do so.

The lack of providers in a position to accept this information electronically raises questions regarding how providers would be deemed in compliance with such requirements. How would providers prove during a survey process that they are "interoperable?" Would they need to send information to other providers electronically? Ensure those providers ultimately received the information? Receive information from other providers? And/or receive information and incorporate it into an actionable format in the EHR? These are just a sampling of the multitude of questions that would arise in determining compliance – and many of them would hinge not on the individual provider's action, but the actions of HIT vendors and other providers over whom the hospital has virtually no control. For example, on provider may be able to send the information electronically, but the receiving provider is unable to accept it. Or, a provider may be unable to incorporate the information it receives into its EHR in a format acceptable to the surveyors due to the limitations of the EHR itself, for example, the misaligned standards, semantics, and specifications that currently hinder data flow and useable data across vendor platforms. Additionally, the CoPs, CfCs, and RfPs are infrequently updated relative to the annual Medicare payment rules. As such, it is possible that the proposed revisions to these requirements could quickly become outdated and hinder future HIT-related innovation, and in many cases even before they are finalized.

Failure to comply with CoPs, CfCs, or RfPs, carries serious penalties for health care providers, including the potential inability to treat Medicare and Medicaid beneficiaries. Such penalties also have profound consequences for patients as well, as they may lose the ability to receive treatment in their communities. Imposing these penalties on providers and patients in the face of an immature health information ecosystem – and the significant implementation issues

raised above – would only restrict rather than facilitate patients' access to care and information exchange.

The FAH appreciates CMS's focus on interoperability and shares CMS's frustrations regarding the lack of actionable, accessible electronic information, as well as the desire to accelerate an interoperable health system that improves the safety and quality of care, enables innovations, and achieves the best possible outcomes for patients. To continue to address these concerns, the FAH recommends that CMS permit the numerous public and private initiatives in this area, some of which are nascent, time to mature and advance our shared goals. CMS and ONC should also continue to work to improve the capabilities of EHRs and other HIT, including: simplifying information exchange across HIT vendor platforms; identifying patients across vendor platforms; and simplifying clinician workflow related to sending, receiving, incorporating, and utilizing information.

As CMS states in the proposed rule, there are "several important initiatives that will be implemented over the next several years to provide hospitals and other participating providers and suppliers with access to robust infrastructure that will enable routine electronic exchange of health information." *Id.* These initiatives include the TEFCA, which is still in draft form; the revamped and refocused Promoting Interoperability Program, which was recently proposed; the Prevention of Information Blocking Attestation;3 and the MyHealthEData initiative, which was announced earlier this year, among others. There are also private-sector led efforts underway to advance other components of the interoperability puzzle, such as plug-and-play interoperability among devices and systems.4 The FAH provided feedback on these and other initiatives and looks forward to continuing to work with CMS, ONC, and other private-sector partners to realize the promise of HIT to improve our nation's health care system.

**********

The FAH appreciates the opportunity to submit these comments. If you have any questions, please contact me at 202-624-1534, or Steve Speil, Executive Vice President Policy, at sspeil@fah.org or 202-624-1529.

Sincerely,

[3] Eligible hospitals, eligible professionals, and CAHs participating in the Promoting Interoperability Programs must attest to the Prevention of Information Blocking Attestation. *The Medicare and Medicaid EHR Incentive Programs Prevention of Information Blocking Attestation Fact Sheet*, October 2017, available at: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/EHR_InformationBlockingFact-Sheet20171106.pdf.

[4] Center for Medical Interoperability, *Fact Sheet*, available at: http://medicalinteroperability.org/wp-content/uploads/2017/01/CMI-Overview-Jan2017.pdf.

Attachment A

# Proposed Refinements to the IRF PPS Case Mix Classification System and Removal of FIM® Items from the IRF PAI

## *Impact Analysis and Potential Considerations*

Submitted to:
Federation of American Hospitals

Allen Dobson, Ph.D.
Alex Hartzman, M.P.A, M.P.H
Kimberly Rhodes, M.A.
Sung Kim
Joan DaVanzo, Ph.D., M.S.W.

Tuesday, June 26, 2018 — *Final Report*

## INTRODUCTION

This report summarizes analyses of the potential impact of proposed changes to the Inpatient Rehabilitation Facility Prospective Payment System (IRF PPS) functional status categorization case-mix system as proposed in the technical report, "Analyses to Inform the Potential Use of Standardized Patient Assessment Data Elements in the Inpatient Rehabilitation Facility Prospective Payment System"[1] (referred to herein as "RTI technical report" or "technical report") and the FY2019 IRF PPS proposed rule. We found that while the proposed system is budget neutral and leaves many facets of the current case mix system in place, it would fundamentally alter the clinical foundation of the IRF PPS through changing the functional assessment items and not weighting them. We note several data and interpretation issues with the new items and characterize the effect of the changes through additional analyses of payments and CMG changes from the current to the proposed system. Dobson DaVanzo and Associates (Dobson | DaVanzo) was commissioned by the Federation of American Hospitals (FAH) on behalf of its membership to conduct this analysis.

CMS has not made the assessment data (required to fully analyze the impacts of the proposed changes) available to date. Although IRF functional assessment data are generally available for use more broadly, the data for the proposed assessment items in question are not yet included in publicly available data sets. Furthermore, an impact file was not made available with the publication of the proposed rule. For purposes of these analyses, Uniform Data System for Medical Rehabilitation (UDSMR) provided Dobson | DaVanzo with a deidentified case-level database contains roughly 84% of IRF cases in the relevant period (FY2017).[2] With these data, in cooperation with UDSMR, we were able to conduct an independent impact assessment of the proposed functional status scoring system.

If implemented without first addressing several important data and interpretation issues pointed to in this report, we anticipate a variety of contradictory financial incentives that may affect the payment accuracy and long-term financial stability of the IRF field.

[1] Morley, Melissa, Benjamin Silver, Anne Deutsch and Melvin Ingber, "Analyses to Inform the Potential Use of Standardized Patient Assessment Data Elements in the Inpatient Rehabilitation Facility Prospective Payment System" RTI, April 2018, page 1-5 https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/Downloads/IRFPPSAnalysis2018RTI.pdf.

[2] The data for this study was obtained and used with permission from the Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities, Inc. The service marks and trademarks associated with the FIM® instrument are all owned by Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities Inc.

## Overview of Proposed Changes to the Inpatient Rehabilitation Facility (IRF) Prospective Payment System (PPS) Case-Mix System

### Current Case-Mix System

The current law IRF PPS case-mix system uses Functional Independence Measure (FIM®) motor and cognitive assessment items to predict case resource use, set rates and determine payments. The motor scores are derived from 12 functional assessment items (listed in Exhibit 1) and weighted to reflect their relative contribution to the costs of care.

**Exhibit 1: FIM® Motor Score Functional Assessment Items**

| Item | IRF PAI Number |
|---|---|
| Eating | 39Aa |
| Grooming | 39Ba |
| Bathing | 39Ca |
| Dressing, upper body | 39Da |
| Dressing, lower body | 39Ea |
| Toileting | 39Fa |
| Bladder management | 39Ga |
| Bowel management | 39Ha |
| Transfers, bed/chair/wheelchair | 39Ia |
| Transfers, toilet | 39Ja |
| Walk/wheelchair | 39La |
| Stairs | 39Ma |

Motor scores are used in conjunction with the patients' rehabilitation impairment category (RIC), and sometimes age and the sum of the cognitive score items to place patients into Case Mix Groups (CMGs). Each CMG carries a set of relative payment weights; the exact weight for a given patient is selected based on his or her comorbidity tier (0-3). Each CMG and comorbidity tier combination has an associated average length of stay which is used as part of the payment determination for early transfer cases to another institutional setting. If a patient's actual length of stay was less than the CMG-comorbidity tier average and the patient was transferred to another institutional setting, the case is paid on a per diem basis (calculated by multiplying the standard payment amount by the case's relative payment weight and dividing the product by the CMG-comorbidity tier's average length of stay. This per diem amount is then multiplied by the case length of stay plus 0.5 days). The CMG is the primary mode by which case payments are differentiated by patient in the IRF PPS.

### Proposed Case-Mix System

The proposed case-mix and patient classification system would replace FIM® items with standardized patient assessment data (SPAD) elements from the Quality Indicators section of the IRF-PAI. These SPAD elements proposed for the new case-mix system are a subset

of standardized assessment items that were implemented to align functional outcome measurement across IRFs, SNFs, LTCHs and HHAs as mandated by the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014. IRF clinicians began collecting patient data on these items for the first time starting in October 2016.

In the proposed system, a collection of 19 items (listed in Exhibit 2) are summed to create an unweighted functional score. Cognitive items were examined for inclusion but not used in the proposed system.

**Exhibit 2: Standardized Patient Assessment Data (SPAD) Elements Motor Score Functional Assessment Items**

| Item | IRF PAI Number |
|---|---|
| Eating | GG0130A1 |
| Oral hygiene | GG0130B1 |
| Toileting hygiene | GG0130C1 |
| Shower/bathe self | GG0130E1 |
| Upper-body dressing | GG0130F1 |
| Lower-body dressing | GG0130G1 |
| Putting on/taking off footwear | GG0130H1 |
| Roll left and right | GG0170A1 |
| Sit to lying | GG0170B1 |
| Lying to sitting on side of bed | GG0170C1 |
| Sit to stand | GG0170D1 |
| Chair/bed-to-chair transfer | GG0170E1 |
| Toilet transfer | GG0170F1 |
| Walk 10 feet | GG0170I1 |
| Walk 50 feet with two turns | GG0170J1 |
| Walk 150 feet | GG0170K1 |
| One-step curb | GG0170M1 |
| Bladder continence | H0350 |
| Bowel continence | H0400 |

Neither the proposed rule nor the RTI technical report provide detail on the rationale for selecting these particular 19 items as opposed to others. And while the RTI technical report does explain that assessment items that were not included "are more challenging and less likely to be assessed on admission",[3] neither document elaborates as to why specific items

[3] Morley, Melissa, Benjamin Silver, Anne Deutsch and Melvin Ingber, "Analyses to Inform the Potential Use of Standardized Patient Assessment Data Elements in the Inpatient Rehabilitation Facility Prospective Payment System" RTI, April 2018, page 1-5 https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/Downloads/IRFPPSAnalysis2018RTI.pdf.

were excluded, such as those pertaining to wheelchair locomotion, which are not necessarily more challenging items than the included items related to walking.

The unweighted functional score, along with the patients' RIC and sometimes age, are used to assign patients to a CMG. As with the current system, each CMG carries a set of relative weights which are selected for each patient depending on his or her comorbidity tier, and each CMG and comorbidity tier combination has an associated average length of stay. Payment calculations for early transfer cases remain the same as in the current system, but the average length of stay values are different given the changes to the CMG grouper.

If finalized, these proposed changes would be implemented in FY2020 (discharges on or after October 1, 2019). In both the technical report and proposed rule, CMS has indicated that changes will be implemented in a budget neutral manner indicating the projected aggregate payment amount to all IRFs for a given case mix would not change. As such, the effect of changes manifests on the distribution of payments across CMGs as well as in the distribution of cases in each CMG, CMG payment rate and average length of stay.

### Differences between Proposed SPAD and Current FIM® Items

While the assessment window for the SPAD and FIM® items is the same (within the first 3 days of IRF admission), SPAD items measure patient functionality in a clinically and quantitatively different way, changing the basic clinical implications (if not treatment implications) of the CMGs. These changes are likely exacerbated by the high prevalence of non-scaled SPAD responses (assessment item responses which are subsequently rescaled to be the most impaired assessment response for the purposes of calculating a functional score for CMG assignment), discussed further below.

- SPAD assessment items are based on patients' "usual performance" on the given item within the assessment period. In comparison, FIM® are based on patients' lowest performance, or highest burden of care, on a given item. Both measures attempt to capture the intensity of assistance patients require, but from different perspectives. While it is not yet clear whether this proposed change to items measuring "usual performance" is inherently problematic, we note numerous implementation issues with the SPAD assessment items.
    - We have received anecdotal reports from FAH members that "usual performance" thus far is not adequately defined, leading to potential inconsistency and problems in inter-rater reliability and stability over time. FAH members reported audits of medical records and found that in more than half of cases the technical reviewer would revise SPAD item responses substantially, suggesting that the inter-rater reliability on SPAD items is currently low. This is particularly concerning given the

high prevalence of non-scaled SPAD assessment responses used in analyses supporting the proposed rule as they substantially affect rate setting and payment determination but may represent coding confusion rather than reasonable non-assessment.

  - Further, CMS repeatedly issued significant updated guidance on the definitions of SPAD items in relation to "usual performance" and other aspects during FY2017, the period in question. As SPAD response meaning has shifted over the first year since implementation to reflect updated guidance, the data used in this proposed payment system change cannot be representative of final or consistent coding practices by definition.

- 6-point scale on SPAD items vs. 7-point scale on FIM®. SPAD items characterize a narrower range of functional impairment but are definitionally broader to attempt to capture a similar range of impairment. See Exhibits 3 and 4, below.
- Both the FIM® and the SPAD assessment items contain possible responses which are subsequently rescaled to be the most impaired assessment response for the purposes of calculating a functional score for CMG assignment. In this report, we are calling such responses "non-scaled responses." However, there are four possible non-scaled responses on the SPAD items while there is just one non-scaled response on FIM®. For SPAD items, these additional non-scaled responses are "patient refused", "not applicable", "not attempted due to safety concerns" and blank; for the FIM® items, the non-scaled response is "activity did not occur". See Exhibit 3 for descriptions of each response items.
- Unlike the FIM®, SPAD motor score items are not weighted, and cognitive items are not present in the proposed CMG grouper.

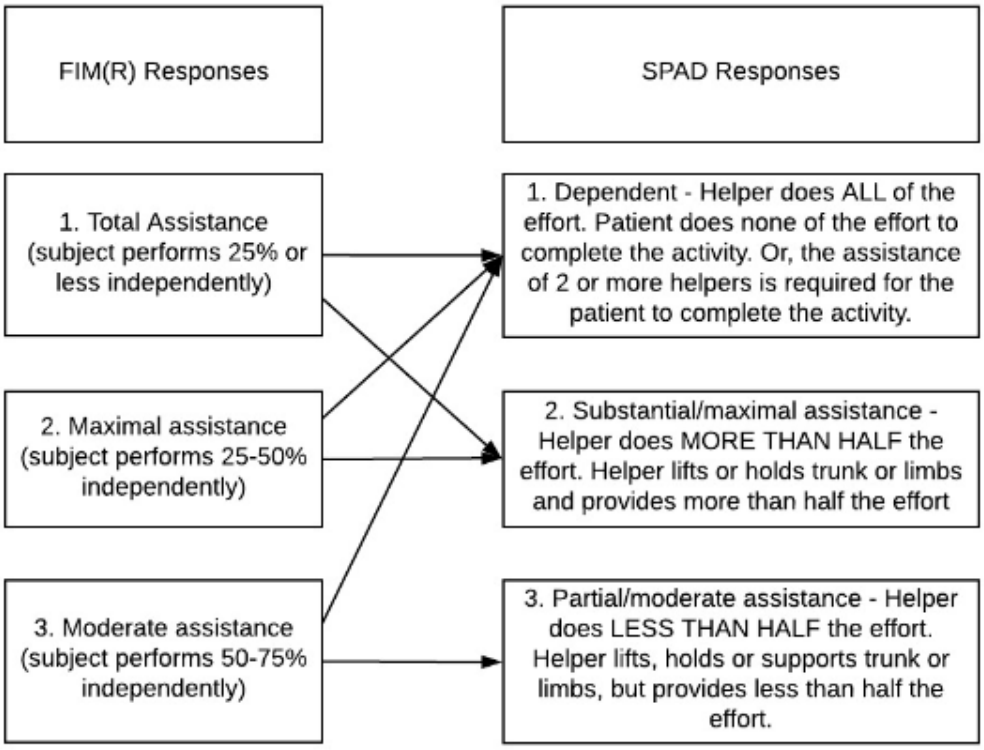
**Exhibit 3: Comparison of FIM® and SPAD Item Scales**

*Non-scaled responses are in italic font.*

| Score | FIM® Levels | SPAD Levels |
|---|---|---|
| **0** | *Activity does not occur (admission only)* | |
| **1** | Total assistance (subject performs 25% or less independently) | Dependent - Helper does ALL of the effort. Patient does none of the effort to complete the activity. Or, the assistance of 2 or more helpers is required for the patient to complete the activity. |
| **2** | Maximal assistance (subject performs 25-50% independently) | Substantial/maximal assistance - Helper does MORE THAN HALF the effort. Helper lifts or holds trunk or limbs and provides more than half the effort |
| **3** | Moderate assistance (subject performs 50-75% independently) | Partial/moderate assistance - Helper does LESS THAN HALF the effort. Helper lifts, holds or supports trunk or limbs, but provides less than half the effort. |
| **4** | Minimal assistance (subject performs 75% or more independently) | Supervision or touching assistance - Helper provides VERBAL CUES or TOUCHING/STEADYING assistance as patient completes activity. Assistance may be provided throughout the activity or intermittently. |
| **5** | Supervision (subject performs 100% independently) | Setup or clean-up assistance - Helper SETS UP or CLEANS UP; patient completes activity. Helper assists only prior to or following the activity. |
| **6** | Modified independence (device used) | Independent - Patient completes the activity by him/herself with no assistance from a helper. |
| **7** | Complete independence | *Patient refused* |
| **88** | | *Not attempted due to medical concern or safety issue* |
| **9** | | *Not applicable* |

Leaving aside that the FIM® and SPAD items measure fundamentally different constructs of patient impairment (worst versus "usual" performance), we note that the scales used to describe impairment correspond across assessment instruments in ways which may have contributed to case payment redistributions. In Exhibit 4, we examine the three scaled items indicating most severe functional impairment under FIM® and SPAD and indicate how FIM® ratings may be reinterpreted under the broader SPAD response scale. For example, the highest severity SPAD response can be inclusive of a broad range of FIM® performance because of the clause "including patients requiring two or more helpers"; further, the SPAD response of 2 spans the range of impairment (50% or more work done by assistant) included in FIM® responses 1 and 2.

**Exhibit 4: Comparison of FIM® and SPAD Responses on Highest Severity Ratings**




### Resulting Changes to Payment System

To incorporate selected SPAD items into CMGs, CMS and RTI took a relatively similar approach to grouping patients as in the original IRF PPS CMG setting. That is, CMGs are defined by clinical condition and combined motor score (sometimes accounting for patient age as well). This process resulted in a different set of CMGs with updated definitions reflecting the SPAD functional assessment items and updated combined motor score. The proposed system would have fewer CMGs (88 rather than the current 92). This would alter the clinical and operational interpretation of each group as relative payment weights and average group lengths of stay also changed to reflect the new CMGs.

Key aspects of the payment system would not change under this proposal. Definitions and usage of rehabilitation impairment categories (RICs), comorbidity tier groups and age were not altered in the proposed rule.

### CMS Rationale for Proposed Changes

In the proposed rule and technical report, CMS offered two arguments in support of the proposed change:

- Reduce IRF administrative burden by removing the FIM® items and associated functional modifiers; and
- Support the broader movement to align data collection across PAC settings.[4]

CMS did not include a specific rationale for changing the clinical meaning of the CMGs or elaborate on how removing the FIM® would support the alignment of the SPAD items across settings.

## Dobson | DaVanzo's Replication of the Current and Proposed Case-Mix Groupings and Payments

Given that an impact file was not made available by CMS, Uniform Data System for Medical Rehabilitation (UDSMR) provided us with a deidentified case-level database that accounts for approximately 84% of traditional Medicare IRF cases. The database included the current and proposed assessment items, patients' age, along with the RIC, comorbidity tier, and proposed and current CMG, CMG average length of stay and unadjusted case payment amounts. Unadjusted case payment amounts are defined as payments that were not adjusted for outlier payments, wage adjustment, rural vs. urban status, low-income status or teaching status. Using unadjusted payments in both the current and proposed systems creates an 'apples to apples' comparison and allows us to analyze the effects of the proposed system changes. Similarly, the impact analyses presented in the RTI report are based on unadjusted payments.[5]

For the purposes of this study, we restricted the dataset to FY2017 cases (patients discharged between October 1, 2016 and September 30, 2017). This is the same period used in the setting of relative weights for the FY2020 proposed CMGs and in the proposed rule and technical report impact assessments.

Using the assessment item, patient age, comorbidity tier and length of stay variables provided in the UDSMR dataset, Dobson | DaVanzo replicated the current and proposed CMG groupings and unadjusted payments. Dobson | DaVanzo also verified that the CMG

[4] Proposed Rule: Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2019 (83 FR 20972 - 21015 // CMS-1688-P). Pages 20988 - 20989. Centers for Medicare & Medicaid Services. May 8, 2018. https://www.federalregister.gov/d/2018-08961.

[5] Morley, Melissa, Benjamin Silver, Anne Deutsch and Melvin Ingber, "Analyses to Inform the Potential Use of Standardized Patient Assessment Data Elements in the Inpatient Rehabilitation Facility Prospective Payment System" RTI, April 2018, page 3-1. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/InpatientRehabFacPPS/Downloads/IRFPPSAnalysis2018RTI.pdf.

average length of stay variables in the UDSMR dataset matched the FY2020 values provided in the FY2019 proposed rule[6] and the FY2017 values in the FY2017 final rule.[7]

To replicate the CMG groupings, Dobson | DaVanzo used the UDSMR dataset variables and the FY2017 CMG relative weights from the FY2017 final rule[8] and the FY2020 CMG relative weights from the FY2019 proposed rule.[9]

With this robust sample size and independent replication of current and proposed CMG groupings and payments, we are confident the following findings are generalizable to IRFs generally.

## FINDINGS

### Budget Neutrality

Budget neutrality between the current and proposed IRF case mix system has been asserted by CMS and contractor RTI through the proposed rule and preceding technical report. However, CMS did not make data publicly available to verify neutrality or to examine the impacts of proposed changes. As such, independent analysts assessed budget neutrality and redistributive effects via samples and data subsets without a complete database.

Within our sample, we find $332M in case payment reductions and $304M in case payment increases, for a net change of negative $28M (-0.46%) in revenue on a basis of the UDSMR dataset of 310,175 discharges and $6,067M revenue in FY2017. In combination with results shared by eRehab (which is representative of much of the remainder of the industry), we conclude payment simulation results are consistent with the CMS statement of a budget neutral case mix system change. However, as discussed below, there are legitimate questions about the budget neutrality of this system moving forward if implemented as outlined in the proposed rule.

Though budget neutrality is a typical condition of payment system changes unless otherwise explicitly altered by legislation, it is not a wholly sufficient measure of the appropriateness of a proposed system change. We also examine case payment and facility

[6] Proposed Rule: Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2019 (83 FR 20972 - 21015 // CMS-1688-P). Table 9 on pages 20992 – 20994. Centers for Medicare & Medicaid Services. May 8, 2018. https://www.federalregister.gov/d/2018-08961.

[7] Final Rule: Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 (42 CFR 412 pages 52055-52141). Table 1 on pages 52063-52070. Centers for Medicare & Medicaid Services. August 5, 2016. https://www.federalregister.gov/d/2016-18196.

[8] Final Rule: Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 (42 CFR 412 pages 52055-52141). Table 1 on pages 52063-52070. Centers for Medicare & Medicaid Services. August 5, 2016. https://www.federalregister.gov/d/2016-18196.

[9] Proposed Rule: Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2019 (83 FR 20972 - 21015 // CMS-1688-P). Table 9 on pages 20992 – 20994. Centers for Medicare & Medicaid Services. May 8, 2018. https://www.federalregister.gov/d/2018-08961.

revenue change as these are indicative of the potential redistribution of revenue and may be useful in predicting industry stress and other threats to beneficiary access when considering proposed system changes. Regardless of the assessment outcome of the proposed change effect on total or average facility revenue, proposed system changes must align with clinical practice and be based on accurate measures – tests which the proposal does not seem to pass.

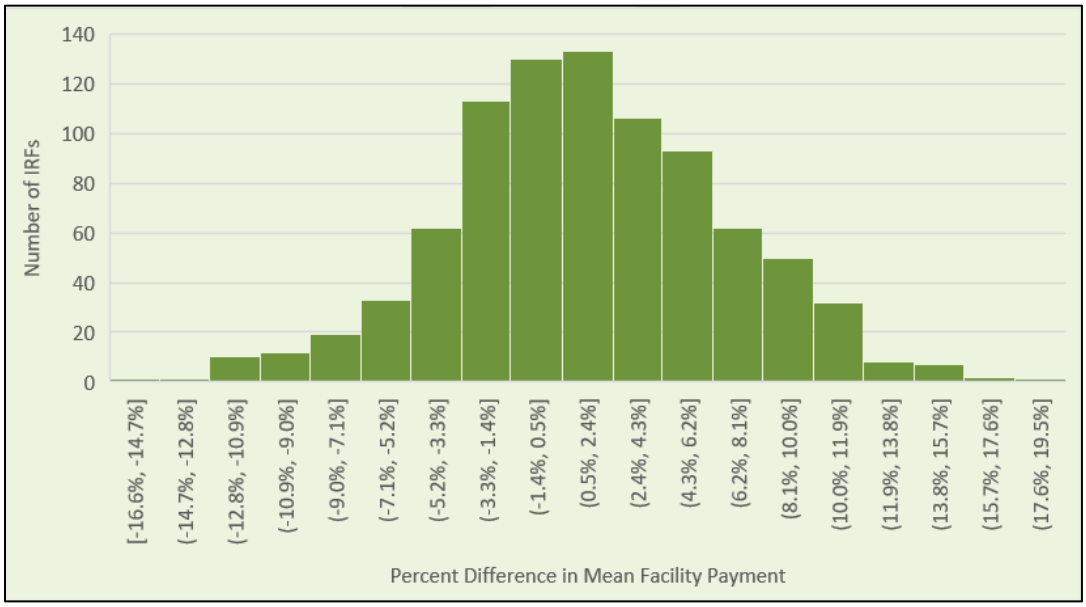
### Redistribution of Cases and Case Payments

Though the proposed system change appears to be budget neutral, we found the proposed rule would redistribute a substantial portion of payments as:

- Case Mix Group definitions are changed, and thus
- CMG payment weights and length of stay are revised to reflect the new groupings.

Given this, we detected substantial changes in overall facility revenue with 54% of facilities in the sample experiencing a +/-3% change; the full distribution of facility percent change in revenue is displayed in Exhibit 5, below.

**Exhibit 5: Distribution of Facility Mean Percent Change in Facility Payment**




Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

Exhibit 6, below, demonstrates the regrouping of beneficiaries from current law CMG (down the left) to proposed CMG (across the top) for stroke cases. In general, the group makeups shift appreciably, suggesting incongruence between the current and proposed sys-

tems. That is, a case that is currently CMG 105 could move to CMG 102 or 103 in the proposed system. In general, cases have migrated towards the middle CMGs in the proposed system, suggesting less variation than is observed under the current system. While to some extent this is expected given the proposed removal of CMGs 107-110, the amount of case CMG movement from one system to the next is significant.

**Exhibit 6: Regrouping of Beneficiaries between Current and Proposed CMGs (RIC 1, Stroke, n = 60,956)**

| | | FY2020 Proposed Stroke CMGs | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 101 | 102 | 103 | 104 | 105 | 106 | Total |
| FY2017 Current Stroke CMGs | 101 | 1,358 | 297 | 22 | 1 | | | 1,678 |
| | 102 | 1,762 | 1,578 | 328 | 7 | 1 | 3 | 3,679 |
| | 103 | 458 | 465 | 114 | 6 | | 2 | 1,045 |
| | 104 | 1,187 | 3,183 | 2,156 | 60 | 2 | 15 | 6,603 |
| | 105 | 424 | 1,901 | 3,300 | 257 | 5 | 24 | 5,911 |
| | 106 | 195 | 917 | 3,531 | 811 | 27 | 96 | 5,577 |
| | 107 | 84 | 432 | 2,760 | 1,737 | 80 | 347 | 5,440 |
| | 108 | 8 | 80 | 783 | 1,289 | 4,415 | | 6,575 |
| | 109 | 36 | 173 | 1,397 | 1,783 | | 1,108 | 4,497 |
| | 110 | 25 | 115 | 1,136 | 2,957 | | 15,718 | 19,951 |
| | Total | 5,537 | 9,141 | 15,527 | 8,908 | 4,530 | 17,313 | 60,956 |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

This change in grouping cases redistributes payments in ways that are predictable from a data perspective, but perhaps not from a clinical one. That is, intuitively beneficiaries that have a higher severity should be reimbursed at higher levels (so far as this is predictive of resource use); however, it is not intuitive why a beneficiary who falls into one CMG under the current system may fall into a very different functional grouping under the proposed system. See Exhibit 7, which shows the aggregate payment change associated with the regrouping of stroke CMGs as an example.

**Exhibit 7: Average Case Payment Change ($) in (FY2020 – FY2017) by Regrouped CMG (RIC 1, Stroke, n = 60,956)**

| FY2017 Current Stroke CMGs | FY2020 Proposed Stroke CMGs | | | | | | |
|---|---|---|---|---|---|---|---|
| | 101 | 102 | 103 | 104 | 105 | 106 | Total |
| 101 | 2,896.23 | 6,188.27 | 9,772.13 | 11,890.21 | | | 3,574.42 |
| 102 | 310.29 | 3,605.54 | 7,654.18 | 12,660.77 | 516.01 | 23,796.57 | 2,421.14 |
| 103 | (1,638.94) | 1,563.83 | 5,197.25 | 9,409.72 | | 19,560.39 | 635.99 |
| 104 | (2,662.95) | 658.07 | 4,533.90 | 8,340.42 | 7,370.91 | 7,788.76 | 1,414.63 |
| 105 | (5,024.98) | (1,655.53) | 2,316.59 | 6,972.70 | 4,732.77 | 9,796.79 | 747.38 |
| 106 | (7,180.49) | (3,700.71) | 327.51 | 4,952.71 | 5,212.52 | 10,421.43 | 272.64 |
| 107 | (9,251.57) | (5,836.38) | (1,738.79) | 3,108.92 | 4,179.50 | 8,852.05 | 130.27 |
| 108 | (15,199.82) | (10,904.35) | (6,920.04) | (1,752.42) | 878.80 | | (728.72) |
| 109 | (12,504.03) | (8,841.12) | (4,450.84) | 582.14 | | 6,425.65 | (8.87) |
| 110 | (20,946.46) | (16,457.65) | (11,577.77) | (6,477.76) | | 393.19 | (1,430.66) |
| Total | (873.23) | (328.99) | (495.48) | (955.26) | 969.96 | 1,030.12 | (59.61) |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

## Non-Scaled Assessment Responses and their Impact on Payment

Missing, non-applicable, and other non-scaled assessment responses are included in both the FIM® (response = 0) and proposed SPAD (response = 7, 88, 9 or blank) motor scores which are used in the current and proposed CMG systems. Missing or otherwise non-scaled responses are automatically given the highest severity rating for the measure, which tends to receive higher payments once aggregated in the combined functional score and assigned a CMG.

Though missing items recoded as high severity scores are included in both assessments' composite score approaches, we found them to be much more prevalent in the SPAD-based motor scores. In general, we found:

- 86% of cases were missing at least one SPAD item compared to 56% of cases missing at least one FIM® motor item.
- 69% of cases had at least two missing SPAD items compared to 10% missing at least two FIM® items.
- 81% of cases had a greater number of missing SPAD items than FIM® items; 2% had more missing FIM® items.

We also analyzed missing scores at the item level and found that when we attempted to compare similar items across both systems (current vs. proposed), certain items that had low prevalence in missing items in the FIM® had much higher prevalence of missing items in the SPAD. The largest discrepancies between the share of missing items from the FIM® to SPAD were in the locomotion category. For instance, we compared the FIM® Locomotion Walk/Wheelchair item to the SPAD Walk 150 feet item, as shown in Exhibit 8 below. We believe that the most appropriate one-to-one comparison with the FIM® Walk/Wheelchair item would be 'Admission Walk 150 feet (GG0170K)', given that the IRF-PAI manual indicates that the FIM® Walk distance of interest is 150 feet, which mirrors the selected equivalent SPAD item distance.

**Exhibit 8: Share of Non-Scaled Item Responses (at the Case-Level) for Walk Items in the Current and Proposed Systems**

| Percent of Cases Missing Admission FIM® Locomotion - Walk/Wheelchair (39L) (Where item = "0") | Percent of Cases Missing Admission Walk 150 feet (GG0170K) (Where item = "7", "88","9", or blank) |
|---|---|
| 6.3% | 75.5% |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

The existence of proportionally more missing items in the SPAD motor score indicates potential 'blind spots' in the proposed system; if an item was able to be scored under FIM® for a particular patient, but the 'same' item on the SPAD is missing a response, this could be a signal for gaps in patient assessment.

Cases with more missing SPAD items tended to receive higher payments under the proposed system.

- 15% of cases had 5 or more items with a non-scaled SPAD response. These cases had an average 4% increase in payment under the proposed system.

- Exhibit 9 demonstrates that, on average, with increasing number of missing SPAD items, proposed FY2020 payments increase relative to current law payments.

**Exhibit 9: Distribution of Payment Changes by Count of Non-Scaled SPAD Item Responses**

| Number of Non-Scaled SPAD Item Responses | Percent of Cases | Average LOS | Sum of FY2017 Unadjusted FPP | Sum of Difference in Payment (FY2020 minus FY2017 Unadjusted FPP w/FY2017 Base) | Percent Change of Difference in Payment (FY2020 minus FY2017 Unadjusted FPP w/FY2017 Base) |
|---|---|---|---|---|---|
| 0--3 | 66.8% | 11.6 | $ 3,832,670,636 | $ (80,270,777) | -2.10% |
| 4--8 | 31.5% | 14.7 | $ 2,163,404,614 | $ 46,790,388 | 2.16% |
| 9--17 | 1.7% | 8.5 | $ 71,135,196 | $ 5,434,895 | 7.86% |
| Total | 100% | 12.5 | $ 6,067,210,447 | $ (28,045,495) | -0.46% |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

Thus, CMGs, payment weights, average CMG length of stay, and payment determinations appear to be influenced by the large portion of items with non-scaled responses. This may be indicative of SPAD data quality issues and warrants closer examination due to the potential for adding instability to the system (payments could decrease as measurement practices improve) or conversely inadvertent incentives to the IRF PPS. **Overall, it is not clear that the FY2017 SPAD data reflecting a significant proportion of non-scaled responses is appropriate for inclusion in the IRF PPS as proposed.**

Fundamentally, the SPAD non-scaled responses represent a broader variety of reasons (compared to the FIM non-scaled response) for why a clinician may not have assigned a scalable response to a given assessment item, including patient refusal or inability to perform the activity. However, these responses are also used much more often under SPAD and it is not clear if the results created by these non-scaled responses are adequate reflections of the patients' functional abilities. As in the FIM® -based CMGs, this tends to compress the effect of sets of patients with somewhat differing functional limitations. If the high level of SPAD non-scaled responses is indeed correct, it may indicate that the 6-point response scale inadequately captures the range of functional impairment, compressing on the end of the least functional beneficiaries. Further, it is also apparent that the greater number of items proposed to be included in the combined SPAD motor score (19 measures, up from 12 FIM® measures in the current motor score) also increases the likelihood of including non-scalable items in rate setting and payment determinations.

Implementing SPAD-based CMGs at this time may incentivize the use of non-scaled assessment responses as items scored as non-responsive tend to result in higher payments. As there are four such responses per SPAD assessment question, there are more

opportunities to compress beneficiaries into groups in unexpected ways (when interpreting the array of possible responses which could lead to the same CMG).

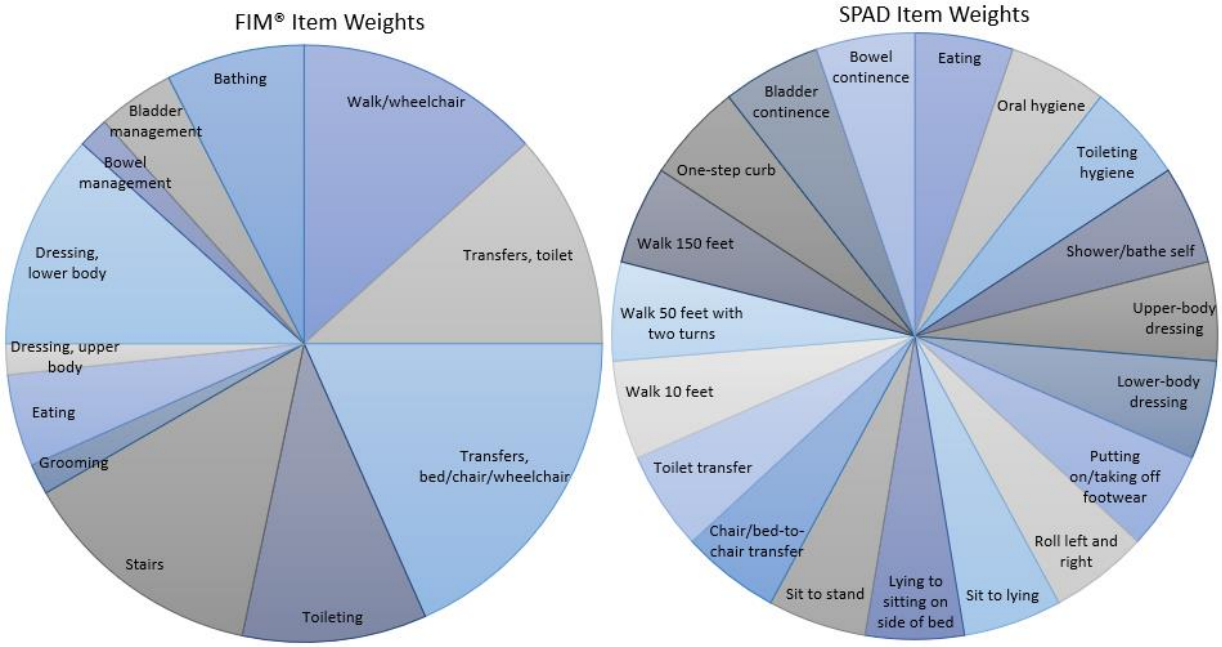
The ability of the SPAD items and motor score to predict resource usage may be eroded over time if the incentive to code items as missing or choosing a non-scaled response remains. As payment weights are reassessed over time, the payment system could become less sensitive to beneficiary impairment should a greater preponderance of non-scaled responses be used (as incentivized by the payment system). This would be detrimental to long-term data quality and make it difficult to assess patient improvement or conduct accurate risk adjustment for quality measures and Alternative Payment Models. Should the assessment data quality decrease, it would also undermine CMS's long-term goal of being able to compare patient functional status across PAC settings.

Conversely, if the prevalence of non-scaled item responses decreases through improved understanding of SPAD ratings over time, it would likely raise combined motor scores (indicating lower impairment) which could then systematically lower payments for providers under the proposed CMGs. The presence of these non-scaled responses and the associated uncertainty of how the prevalence of these responses will change over time is indicative of a system that may not be ready for implementation. Given this, if the proposed system changes were implemented we would anticipate the need for future changes to both the CMG weights and perhaps the base rate as well to assure continued budget neutrality over time.

### Implications of an Unweighted Motor Score

CMS has not proposed to weight the new motor score at this time. As outlined in the proposed rule, the motor score assessment items would each have equal weight in contrast to the current motor score, which is weighted according to each items' associated burden of care. Exhibit 10 depicts the current weighting of the motor score compared to the proposed, unweighted motor score composition.

**Exhibit 10: FIM® vs. SPAD Item Weights**




We have been unable to locate related documentation or analyses to identify whether CMS carefully evaluated the potential implications of removing the weighting system tied to the motor score. In the past, considerable research has gone into identifying and applying optimal motor score weights. This research (led by RAND for CMS) found improvements in explanatory power of models to predict actual costs when applying weights to the motor FIM® items, thus adding accuracy to the payment system.[10] Regardless, given that the burden of care represented by some items (oral hygiene, for instance) do not carry the same weight or burden as others (such as chair/bed to chair transfer, or stairs), the exploration of an item-level weighting system would be beneficial.

It is unclear whether CMS plans to implement a weighting system for the proposed motor score items in the future, as the proposed rule states that CMS is "not proposing to apply a weighting methodology to the motor score *at this time*" [Emphasis added].[11] However, if CMS were to implement a weighting system, the impacts of this would flow through the entire payment system and thus should be carefully considered.

[10] Relles, Daniel, Gregory Ridgeway, Grace Carter, and Melinda Beeuwkes Buntin. "Possible Refinements to the Construction of Function-Related Groups for the Inpatient Rehabilitation Facility Prospective Payment System" RAND Health supported by the Centers for Medicare and Medicaid Services. 2005. https://www.rand.org/pubs/technical_reports/TR207.html.

[11] Proposed Rule: Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2019 (83 FR 20972 - 21015 // CMS-1688-P). Page 20990. Centers for Medicare & Medicaid Services. May 8, 2018. https://www.federalregister.gov/d/2018-08961.

### Differences in Evaluating Patient Functional Status

The items and scale on which patients would be assessed in the proposed system fundamentally changes the way patient functionality is measured. One of the effects this change has is that it redistributes which patients are considered the most functionally impaired. This could be potentially harmful to patients considered highly impaired under the current system as it would reduce treatment budgets for these patients. As the change in measurement set and motor score were not well-justified clinically as well as the prevalence of numerous data reporting problems, we cannot say whether the change in payment for the most severely impaired patients currently under FIM® is appropriate.

In a sense, the proposed system may be biased against the most severe patients, who are most in need of the care that IRFs provide. Exhibit 11 shows how payments are redistributed from many of the most impaired beneficiaries to less impaired beneficiaries, as rated by the weighted FIM® composite score.

**Exhibit 11: Distribution of Payment Changes by Case Composite Weighted FIM® Motor Score**

| Case Composite Weighted FIM® Motor Score | Percent of Cases | Average LOS | Sum of Difference in Payment (FY2020 minus FY2017 Unadjusted FPP w/FY2017 Base) | Average of Difference in Payment (FY2020 minus FY2017 Unadjusted FPP w/FY2017 Base) |
|---|---|---|---|---|
| 12 | 5.10% | 16.35 | $ 1,057,508 | $ 401 |
| 14 | 7.04% | 15.75 | $ 84,158 | $ 62 |
| 16 | 7.86% | 15.25 | $ (470,442) | $ (327) |
| 18 | 7.31% | 14.86 | $ (1,378,710) | $ (1,118) |
| 20 | 6.97% | 14.43 | $ (2,186,542) | $ (1,908) |
| 22 | 6.67% | 13.73 | $ (1,852,793) | $ (1,718) |
| 24 | 6.35% | 13.18 | $ (1,597,472) | $ (1,570) |
| 26 | 6.13% | 12.52 | $ (533,189) | $ (559) |
| 28 | 5.88% | 11.94 | $ 136,690 | $ 141 |
| 30 | 5.77% | 11.53 | $ 271,276 | $ 297 |
| 32 | 5.43% | 11.19 | $ 60,653 | $ 67 |
| 34 | 5.08% | 10.80 | $ 277,638 | $ 350 |
| 36 | 4.76% | 10.44 | $ 353,509 | $ 480 |
| 38 | 4.35% | 9.91 | $ 783,082 | $ 1,162 |
| 40 | 3.90% | 9.52 | $ 911,382 | $ 1,472 |
| 42 | 3.27% | 9.21 | $ 671,089 | $ 1,287 |
| 44 | 2.47% | 8.89 | $ 559,016 | $ 1,448 |
| 46 | 1.89% | 8.45 | $ 433,272 | $ 1,436 |
| 48 | 1.31% | 8.09 | $ 407,749 | $ 2,006 |
| 50 | 0.96% | 7.59 | $ 389,160 | $ 2,540 |
| 52 | 0.56% | 7.15 | $ 271,456 | $ 2,923 |
| 54 | 0.37% | 6.68 | $ 176,066 | $ 2,907 |
| 56 | 0.22% | 6.30 | $ 104,167 | $ 2,640 |
| 58 | 0.13% | 5.86 | $ 56,857 | $ 2,404 |
| 60 | 0.09% | 5.41 | $ 43,803 | $ 2,376 |
| 62 | 0.05% | 5.06 | $ 22,889 | $ 2,134 |
| 64 | 0.03% | 4.49 | $ 12,068 | $ 1,940 |
| 66 | 0.02% | 4.02 | $ 5,871 | $ 1,467 |
| 68 | 0.01% | 4.16 | $ 2,735 | $ 1,466 |
| 70 | 0.01% | 2.84 | $ 2,536 | $ 966 |
| 72 | 0.01% | 3.33 | $ 3,419 | $ 1,576 |
| 74 | 0.00% | 3.29 | $ 5,186 | $ 2,420 |
| 76 | 0.00% | 2.90 | $ 840 | $ 691 |
| 78 | 0.00% | 2.43 | $ 1,270 | $ 809 |
| 80 | 0.00% | 1.40 | $ 339 | $ 339 |
| 82 | 0.00% | 2.00 | $ 679 | $ 339 |
| Grand Total | 100.00% | 12.53 | $ (28,045,495) | $ (90) |

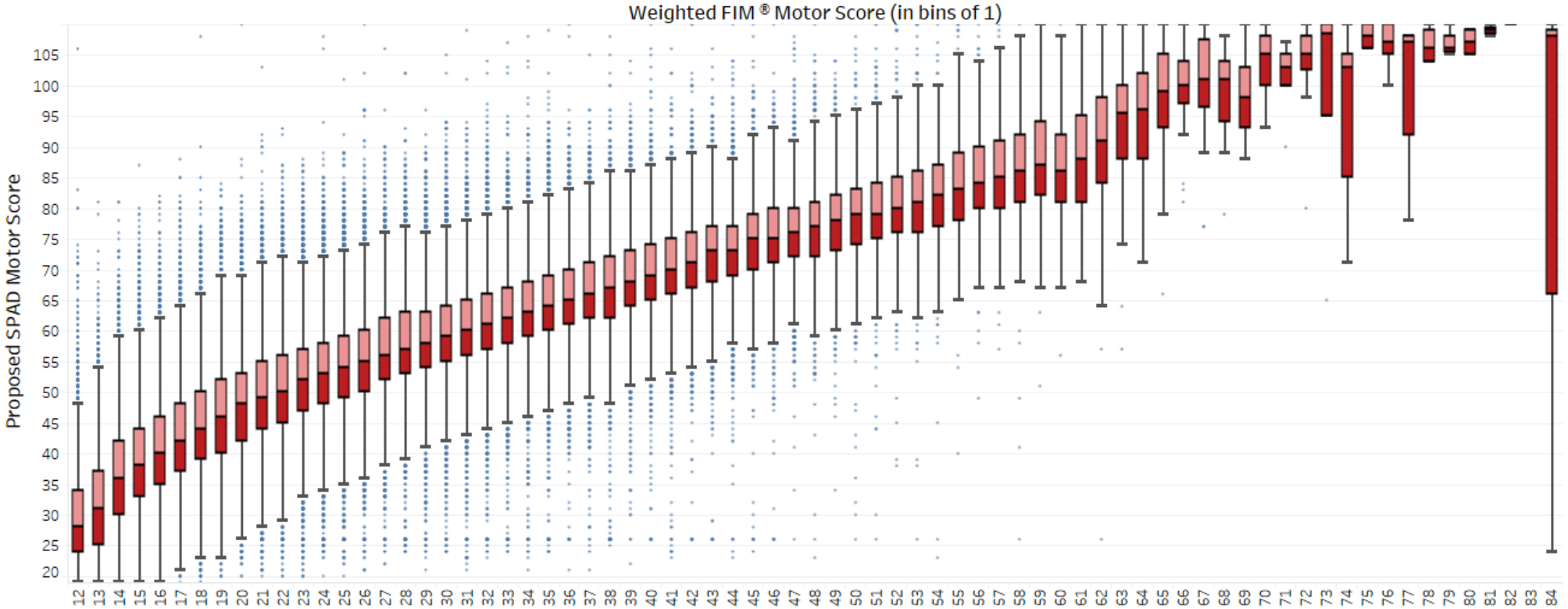
Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

Exhibit 12 shows the wide variability in patients' SPAD motor scores as compared to their weighted FIM® motor score. We also see overlapping interquartile ranges of SPAD motor scores from one FIM® score to the next. Both of these findings suggest that the SPAD items and their associated motor function scale blur the distinction between functional status levels, which are clearly distinguished from one another in the current FIM® scoring system. Ultimately, it appears that the SPAD scale has some amount of "noise" and makes it difficult to clearly discern different levels of functionality, reducing the distinctness of one SPAD level from another.

We also looked at the same distribution restricted to stroke (RIC 1) cases (not shown) and compared it to the FIM® motor score thresholds for CMGs 102 and 103. We found that a broad range of SPAD scores fit into the relatively narrow range of motor score cut points under the current CMG system.

**Exhibit 12: Distribution of FY2017 Cases by Weighted FIM® Motor Score and Proposed SPAD Motor Score**

*Each "box" shows the 25th and 75th percentile SPAD motor scores for each FIM® motor score (truncated to closest integer); the midpoint of each box (where the light and dark red meet within each box) shows the 50th percentile of the SPAD motor score for each IRF cases' FIM® motor score. The "whisker" (lower and upper-most bound of each box plot) extend to values within 1.5 times the inter-quartile range. The dots represent cases that fall outside of the range defined by + or – 1.5 times the interquartile range and show the minimum and maximum values.*




Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

### Differences in Assessing Change in Patient Functional Status

While the admission functional assessment scores are the only ones used for payment determination purposes, we also analyzed the discharge assessment scores to get a sense for how patients' change in functionality is captured in the current and proposed systems. To do this, we constructed admission and discharge motor functional scores and calculated the percent change in functionality from admission to discharge for each case. We used the unweighted motor FIM® score in order to see the raw change in functionality from admission to discharge, allowing for the most direct "apples to apples" comparison from the current to the proposed scores.

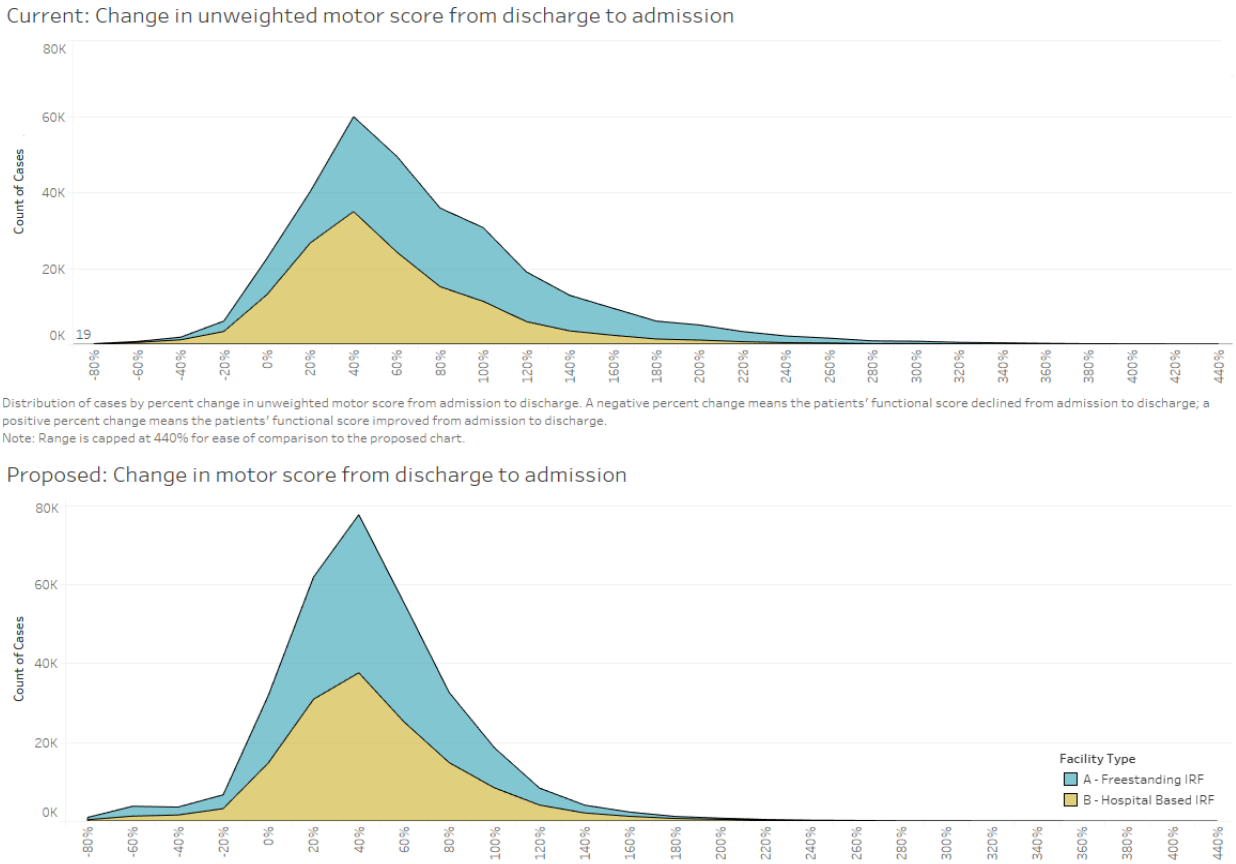
Based on this analysis, we found that the proposed motor SPAD items and scale may dilute measurement of "true" changes in patient functionality. If the new system causes patients to appear as less functionally impaired compared to the results of the case FIM® assessment, then less improvement over time is likely to be measured.

Exhibit 13 shows the distribution of cases by their percent change in functional status as measured by the unweighted motor FIM® score (top graph) and by the proposed motor SPAD score (bottom graph).

A slightly wider and flatter distribution is observed in the FIM® motor score distribution compared to the SPAD motor score distribution, suggesting that there is more pre-post (admission to discharge) variation (or sensitivity) in the FIM® than in the SPAD.

The narrower, more concentrated distribution of the proposed SPAD motor scores indicates that patients may appear to be less severe in the proposed system than they show in the current FIM® system, and that changes in functionality from admission to discharge may not be captured with as much granularity in the SPAD as they are in the FIM®. This could affect future quality measurement (and potentially payment), which focuses on the outcomes of therapy and the provision of other IRF services.

**Exhibit 13: Distribution of FY2017 Cases by Percent Change in Motor Score from Discharge to Admission: Current (Unweighted FIM®) vs. Proposed (SPAD)**




Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

## Differences between FIM® and SPAD Response Meanings

The proposed functional status items are not direct equivalents to the current FIM® items and in some cases may make patients look less functionally impaired on similar items. For example:

- "Easier" items were chosen.
    - FIM® walk/wheelchair item was replaced with three walking items.
    - Locomotion 12 step FIM® item was replaced with a 1 step SPAD item.
    - Bowel and bladder FIM® items captured more nuance in acuity than SPAD items.

Other changes to the motor score include:

- Additional assessment items that do not replace existing FIM® items:
    - Roll left and right, sit to lying, lying to sitting on side of bed, and sit to stand
    - The removal of cognitive assessment items.
    - Removal of weights in the motor score calculation.

Last, the core concept of what is being measured is fundamentally different. SPAD items measure "usual performance" at admission as opposed to FIM® which captures worst admission performance. This is a change to the concept of how patient need is considered in relation to resource needs. This change, as well as the inclusion of more assessment items to inform the motor score, inherently alters which patients are viewed as having any given level of functional impairment. Exhibit 14a shows the redistribution of beneficiary scores from the FIM® to SPAD on the similar eating measure.

**Exhibit 14a: Distribution of Cases by their Current and Proposed Assessment Item Score**

| | | Admission SPAD Eating | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 9 | 88 (missing) | | Total |
| Admission FIM® Eating | 0 | 0.01% | 0.00% | 0.01% | 0.02% | 0.03% | 0.02% | 0.02% | 0.01% | 0.16% | 0.01% | **0.28%** |
| | 1 | 2.34% | 1.05% | 0.58% | 0.67% | 0.90% | 0.32% | 0.03% | 0.23% | 1.07% | 0.01% | **7.20%** |
| | 2 | 0.06% | 0.69% | 0.36% | 0.34% | 0.50% | 0.12% | 0.01% | 0.01% | 0.03% | 0.00% | **2.11%** |
| | 3 | 0.02% | 0.13% | 1.04% | 0.52% | 0.81% | 0.23% | 0.01% | 0.01% | 0.03% | 0.00% | **2.79%** |
| | 4 | 0.03% | 0.07% | 1.99% | 3.02% | 3.18% | 1.08% | 0.04% | 0.03% | 0.08% | 0.00% | **9.51%** |
| | 5 | 0.04% | 0.09% | 0.67% | 8.49% | 37.64% | 15.64% | 0.18% | 0.16% | 0.41% | 0.01% | **63.32%** |
| | 6 | 0.00% | 0.00% | 0.03% | 0.14% | 1.18% | 5.88% | 0.03% | 0.03% | 0.05% | 0.00% | **7.34%** |
| | 7 | 0.00% | 0.00% | 0.02% | 0.10% | 0.70% | 6.46% | 0.07% | 0.03% | 0.07% | 0.00% | **7.45%** |
| | **Total** | **2.51%** | **2.04%** | **4.69%** | **13.28%** | **44.93%** | **29.74%** | **0.38%** | **0.51%** | **1.89%** | **0.03%** | **100.00%** |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

This redistribution of scores has several consequences for scoring. For example, as Exhibit 14b (derived from 14a) shows, a 13.2-percentage point difference exists between the average percentage of maximal function in eating when looking at the FIM® item versus the SPAD eating item, causing the same cases to appear measurably more functionally independent in the SPAD item.

**Exhibit 14b: Percentage of Maximum Functional Score of Eating Measures (Average Functional Score divided by most Independent Scale Item)**

| FIM® | SPAD |
|---|---|
| 67.3% | 80.5% |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

## Altered Incentives for Providers and Coding Clinicians

The proposed system changes the average reimbursement and expected average length of stay for many cases. If implemented, the new case-mix groupings would reshape resource usage as these CMG characteristics effectively set the case payment amount budget. Thus, the same patient may be associated with a higher or lower level of resource use and a longer or shorter expected length of stay under the new system, as Exhibit 15 shows.

**Exhibit 15: Average Change in CMG Average Length of Stay for Stroke Cases between the Current and Proposed CMGs**

| FY2017 Current Stroke CMGs | FY2020 Proposed Stroke CMGs | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 101 | 102 | 103 | 104 | 105 | 106 | 5001 | 5103 | 5104 | Grand Average |
| 101 | 1.72 | 3.70 | 6.59 | 10.00 | | | | | | 2.14 |
| 102 | (0.00) | 2.00 | 4.98 | 8.86 | 10.00 | 14.00 | | | | 1.33 |
| 103 | (1.97) | 0.03 | 3.03 | 7.00 | | 11.50 | | | | (0.46) |
| 104 | (1.98) | 0.03 | 3.00 | 7.02 | 8.50 | 11.33 | | | | 0.73 |
| 105 | (3.95) | (1.97) | 1.00 | 5.00 | 6.20 | 9.67 | | | | (0.10) |
| 106 | (5.00) | (3.00) | (0.03) | 3.95 | 5.33 | 8.63 | | | | 0.06 |
| 107 | (7.00) | (4.99) | (2.01) | 2.00 | 3.39 | 6.62 | | | | (0.41) |
| 108 | (10.50) | (8.38) | (5.38) | (1.39) | 0.01 | | | | | (1.02) |
| 109 | (8.69) | (6.57) | (3.56) | 0.45 | | 5.06 | | | | (0.00) |
| 110 | (14.00) | (12.19) | (9.20) | (5.26) | | (0.61) | | | | (1.87) |
| 5001 | | | | | | | 11.38 | | | 11.38 |
| 5103 | | | | | | | | 11.72 | | 11.72 |
| 5104 | | | | | | | | | 0.42 | 0.42 |
| Grand Average | (0.89) | (0.82) | (0.86) | (0.90) | 0.12 | (0.03) | 11.38 | 11.72 | 0.42 | (0.41) |

Source: Dobson DaVanzo analysis of UDSMR FY2017 cases.

This could be problematic if the new scale is in some ways inaccurate and/or misleading. That is, if the new clinical groupings are not appropriate, this may introduce instability and perverse incentives in the system – as measurement improves, composite functional scores may decrease leading to potential system underpayments. As described above, new clinical groups represent a different view of patient functional status that is somewhat incongruent with the current assessment methods. As proposed CMGs consist of different groupings of beneficiaries compared to the current CMGs, case budgets and expected lengths of stay will change. This may alter how resources are apportioned – the course of care – across patients.

## Data Availability and Transparency of Analytic Methods Behind CMS' Impact Analysis

The technical report lacked key details about analytic approach, clinical advice, and rationale for broad changes to the clinical interpretation of the case mix system. These

details would have been reassuring that the process was conducted with the same rigor as the preliminary IRF PPS rate setting technical report or the highly detailed proposed update to the SNF PPS (which included a facility-specific impact file) and could ensure that relative weight and length of stay values were not drastically altered. Furthermore, we are not aware that CMS or RTI conducted a technical expert panel with IRF clinicians to gain input from the field on the potential clinical and operational impacts of these changes.

Though the technical report does include some descriptions of the process for revising CMG definitions, it excludes information about the CART algorithm, clinical input and results. In particular, there were numerous clinical decisions we have found here which were not addressed in the report, such as the removal of item weights, the addition of new items to the motor score, etc. In the initial IRF PPS technical report conducted by RAND that was published in 2002,[12] numerous combined assessment score models were attempted – each supported by published literature – and results including goodness of fit statistics were shared for each alternative. This due diligence may have been conducted by RTI, but the lack of prior rigorous evaluation of the SPAD items may have made this more challenging. In any case, the RTI report did not attempt to justify the change in scale or change in the number and types of assessment items selected beyond describing some as burdensome.

## CONCLUSION

The FY2019 proposed system changes for FY2020 would fundamentally change the way IRF patients' functional abilities and impairments are determined. The change in assessment items requires updates to the CMGs which, now representing different groupings of the same cases, change. As patients are recategorized and incentives realigned to meet this categorization, payments will be redistributed across case types relative to current law. This may cause unpredictable changes in how patients are cared for in inpatient rehabilitation facilities as well as in other settings, such as home health and home with a family or friend caregiver.

Further, there are implementation challenges inherent in using the first operational period of data from a new clinical assessment tool to reset a payment system. The performance of the new SPAD items has not yet been widely evaluated for benchmarking and comparative purposes and seem to have numerous coding and clinical interpretation problems. If the proposed system changes are implemented without first ensuring the data assessment tool performs adequately, CMS may introduce unintended incentives and year-to-year payment

[12] Carter, Grace, Melinda Beeuwkes Buntin, Orla Hayden, Jennifer Kawata, Susan Paddock, Daniel Relles, Gregory Ridgeway, Mark Totten and Barbara Wynn. Analyses for the Initial Implementation of the Inpatient Rehabilitation Facility Prospective Payment System. RAND Health prepared for the Centers for Medicare and Medicaid Services. 2002. https://www.rand.org/content/dam/rand/pubs/monograph_reports/MR1500/MR1500.pdf.

system instability as providers seek to understand the new case mix system. Ultimately, it is the beneficiaries who will face the consequences of CMG, length of stay and payment confusion and instability as the proposal would alter both how patients are perceived and how they access care as providers seek to adapt to these changes.

There remain essential unanswered questions both clinically and operationally. Data and CMG construction considerations indicate this model may not be an appropriate step at this time. Finalizing the proposed change at this time may introduce payment instabilities or perverse incentives that could weaken various aspects of functional assessment measurement as well as create an unbalanced system that may require further updates.

Ultimately, we have found evidence suggesting that the proposed system is not ready for implementation as of FY2020. First, the data on which the proposed system is new (as of October 2016) and through FAH member studies has shown evidence of low inter-rater reliability. Second, budget neutrality would be compromised if the non-scaled item responses (now set to be the least independent value) were more frequently coded at a different value than the default value. Third, as clinicians better understand the differences between usual and worst performance, this will have an impact of how they assess patients, and thus their patients' eventual assessment score. Fourth, the impact of the exclusion of key elements of the current system (namely, cognitive items and motor score item weights) is yet to be determined. As the new system reflects the above considerations, the functional status score for any given individual is likely to change – the implications are that any patient will be given different treatment for their care and that this care will be paid for differently. If these elements are not addressed, unintended consequences, including potential harm to IRF patients, could result.