Research Report

Reviewing, Refining, and Validating Claims-Based Algorithms of Frailty and Functional Impairment

Final Report

Sara E. Heins, Denis Agniel, Jacob Mann, Melony E. Sorbero

For more information on this publication, visit **www.rand.org/t/RRA1871-1**.

Published by the RAND Corporation, Santa Monica, Calif.

About This Report

Frailty is a clinical syndrome that is associated with negative health outcomes and increased risk of mortality. Measuring frailty might be important to improving risk-adjustment for value-based payments or targeting interventions. This report reviews existing algorithms and describes newly developed algorithms to predict frailty and functional impairment using routinely collected administrative claims data. The findings of this report will be of interest to researchers, health care providers, measure developers, and other stakeholders interested in functional impairment.

This work was conducted under a contract with the Office of the Assistant Secretary for Planning and Evaluation (contract #HHSP233201500038I) with funding by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund and carried out within the Payment, Cost, and Coverage Program in RAND Health Care.

RAND Health Care, a division of the RAND Corporation, promotes healthier societies by improving health care systems in the United States and other countries. We do this by providing health care decisionmakers, practitioners, and consumers with actionable, rigorous, objective evidence to support their most complex decisions. For more information, see www.rand.org/health-care, or contact

RAND Health Care Communications
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
(310) 393-0411, ext. 7775
RAND_Health-Care@rand.org

Acknowledgments

Valuable feedback on this report from the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, was provided by Lok Wong Samson, Violanda Grigorescu, Rachel Zuckerman, Wafa Tarazi, and Kenneth Finegold. We thank the members of our Project Advisory Task Force (Table A.1) for their thoughtful input into developing and validating algorithms. We are grateful to Sukruth Shashikumar from Washington University and Justin Timbie, Christine Eibner, and Paul Koegel of the RAND Corporation for reviewing this report.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the U.S. Department of Health and Human Services.

Summary

Frailty is a clinical syndrome that is characterized by a constellation of symptoms, including loss of strength, low energy, and weight loss, and is associated with negative health outcomes, such as falls, disability, fractures, and increased risk of mortality (Fried et al., 2001). Frailty is also associated with increased health care utilization and spending, independent of other medical risk factors (Goldfarb et al., 2017; McNallan et al., 2013). Identifying and quantifying frailty might be an important component of risk-adjustment for value-based payments or might help target specific interventions. Despite its importance, measuring frailty is challenging because of the lack of consistent measurement of frailty-related concepts. One such concept is *functional impairment*, which captures limitations in one's ability to perform everyday tasks (Üstün and Kennedy, 2009). Functional impairment is sometimes used as a proxy for frailty and is measured in some specialty care settings but is not collected during most clinical encounters.

To help stakeholders and researchers identify individuals at greater risk of frailty and functional impairment using routinely collected data, the U.S. Department of Health and Human Services Assistant Secretary for Planning and Evaluation (ASPE) contracted with the RAND Corporation to review and refine claims-based algorithms. A project advisory task force of experts was convened to guide the project and review the findings. RAND reviewed the literature on existing algorithms, identified candidate algorithms for evaluation, and developed new algorithms using Medicare fee-for-service (FFS) claims that were validated using patient assessment data from two types of post-acute care (PAC) providers: home health agencies and skilled nursing facilities. This approach had several potential advantages over prior algorithm development efforts, such as use of a larger sample size, validation using clinician-assessed functional impairment outcomes, and inclusion of beneficiaries younger than 65. However, the new algorithms were developed solely on beneficiaries receiving PAC care, so it was unknown how they would perform relative to existing algorithms in a broader Medicare population. Thus, RAND compared the relative performance of the new and existing algorithms at predicting three claims-based outcomes in a data set representative of all Medicare FFS beneficiaries. Overall, RAND found that using algorithms previously developed by Kim et al., 2018, performed best for most outcomes and subpopulations, although the new algorithms performed slightly better at predicting a nursing home stay in the following year by some metrics, particularly among PAC patients.

Recommendations

The results of this study informed ASPE's recommendations to add Kim's claims-based frailty index scores to the Centers for Medicare and Medicaid Services (CMS) Chronic

Conditions Data Warehouse (CCW) to be used by CMS, researchers, and other stakeholders. These scores will be available in the Master Beneficiary Summary File chronic condition segment of the CCW. Ongoing work is examining the potential added value of incorporating electronic health record data to the Kim algorithm. Future work might explore the value and feasibility of adding PAC assessment outcomes to the CCW.

Contents

Figures and Tables

Figures

Tables

1. Introduction

Background

Frailty is a clinical syndrome that is characterized by a constellation of symptoms, including loss of strength, low energy, and weight loss (Fried et al., 2001; Xue, 2011). Frail individuals are vulnerable to negative health outcomes, such as falls, disability, and fractures (Ensrud et al., 2009) and have an increased risk of mortality (Graham et al., 2009). Frailty is also associated with increased health care utilization and spending (Kan et al., 2018; McIsaac et al., 2016; Office of the Assistant Secretary for Planning and Evaluation [ASPE], 2020), independent of other medical risk factors (Goldfarb et al., 2017; McNallan et al., 2013). Understanding and measuring frailty might be an important component of clinical risk adjustment for a variety of purposes—for example, identifying high-risk patients for interventions, value-based payment models, and research. Although increasing frailty with age is to be expected, it is possible to intervene upon and improve some aspects of frailty. For example, a systematic review found evidence that strength training interventions can improve physical functioning among older adults (Liu and Latham, 2009). Thus, identifying frail individuals might also be important for targeting specific interventions.

Although frailty is thought to be clinically recognizable, identifying frailty in secondary data is challenging due to the lack of consistent measurement of frailty-related concepts in routinely collected data and disagreement over which clinical measurements should be included (Rockwood, 2005; Sternberg et al., 2011). There is no single way to directly measure frailty, but a common approach is to count an accumulation of *deficits*, including such symptoms as hearing or vision loss, abnormal laboratory values, or difficulty engaging in activities of daily living (ADL; Mitnitski, Mogilner, and Rockwood, 2001). Lack of ADL independence alone has also been considered a proxy for frailty (Faurot et al., 2015), and having an ADL dependency is a strong predictor of early mortality (Keeler et al., 2010; Walter et al., 2001).

Although ADL measures are important proxies for frailty, they are not obtained in routine clinical encounters. To help identify individuals who are frail or at risk of frailty among a broader sample of older adults, several algorithms have been developed that use diagnoses and other data coded in routinely collected administrative claims data to predict proxies of frailty or functional impairment. Examples are claims-based algorithms by Faurot et al., 2015, and Kim et al., 2018. These algorithms both use data from the Medicare Current Beneficiary Survey (MCBS), which includes claims data along with self-reported ADL dependencies and other self-reported indicators related to functional impairment, such as mobility limitations or recurrent falls.

Purpose

The goal of this project was to develop or identify one or more measures of frailty that could be added to the Centers for Medicare and Medicaid Services (CMS) Chronic Conditions Data Warehouse (CCW) so that these data would be available for use by researchers, health care providers, measure developers, and other stakeholders. ASPE contracted with the RAND Corporation to review and refine algorithms using Medicare fee-for-service (FFS) claims to identify and predict frailty and functional impairment using patient assessment data from two types of post-acute care (PAC) providers: home health agencies (HHAs) and skilled nursing facilities (SNFs). This project also compared the various algorithms against claims-based outcomes to identify the best performing algorithm in predicting health outcomes.

Approach

To arrive at the claims-based frailty index (CFI) scores used in the report, we reviewed several existing algorithms and tested and refined these algorithms using new sources of data. The methods and results of this review, refinement, and validation work are detailed in the following chapters. Briefly, our approach was as follows:

- reviewed literature on existing algorithms to predict frailty and functional impairment
- convened a project advisory task force (PATF) consisting of ten individuals representing clinical, research, and data science backgrounds with expertise in identifying frailty and functional impairment to advise on algorithm development and testing (see Table A.1 for list of members)
- identified a pool of claims-based candidate predictors drawing on a literature review and feedback from the PATF and ASPE to potentially be included in new algorithms
- created a study population of Medicare FFS beneficiaries enrolled from 2014 to 2017 and consisting of those who did and did not receive care from PAC providers
- linked Medicare FFS claims with PAC assessment data, where applicable
- developed inverse probability weights to improve generalizability of the PAC population to the overall Medicare FFS population
- developed two outcome measures from PAC admission assessment items using factor analysis methods that captured aspects of functional impairment: (1) presence of a memory limitation and (2) a count (zero to six) of activity and mobility limitations
- used Lasso regression, a supervised machine learning technique, to identify which candidate predictors best predicted the main outcomes and tested alternate algorithm specifications
- validated the new algorithms in a holdout sample to compare predicted and actual outcomes in the PAC population
- compared the performance of models using scores from the new algorithms with models using scores from existing algorithms in predicting claims-based outcomes, such as hospitalizations and nursing facility stays in several populations of interest.

Summary of Findings

- New algorithms were developed and tested on a sample of 35,141,239 Medicare beneficiaries—18 percent of whom were in the HHA group, 6 percent of whom were in the SNF group, and 76 percent or whom had neither an HHA nor a SNF stay during the study period.
- In these new algorithms, age and dementia were significant predictors of both (1) memory limitation and (2) activity and mobility limitations outcomes.
- In a comparison of the new algorithms with existing algorithms from Kim et al., 2018 (hereafter, the *Kim algorithm*), and Faurot et al., 2015 (the *Faurot algorithm*), the Kim model (hereafter, the *Kim model*), which used scores from the Kim algorithm in deciles in combination with age and sex, had the best overall performance at predicting claims-based outcomes of interest in a separate sample of Medicare FFS beneficiaries for most metrics and subpopulations tested, although models including scores from the new algorithms performed slightly better at predicting a nursing home stay in the following year by some metrics, particularly among PAC patients.
- Our results did not indicate that differential health care utilization by race/ethnic group or neighborhood socioeconomic status negatively affected model performance.

Summary of Recommendations

Given the results of this work, the Kim algorithm was selected to calculate CFI scores to identify Medicare beneficiaries at risk for functional impairments and will be included in the CCW. These scores might also be useful to health systems for risk adjustment or for tracking quality of care and utilization for at-risk populations by stratifying measures. Researchers might consider including these scores as controls in evaluations of policies or using them to study the potential effect modification of frailty on different interventions. By making these scores available in the CCW, health systems, researchers, and other data users might test different applications of the CFI scores.

Organization of the Report

This report first describes the study population used to develop new algorithms predicting functional impairment and validated against PAC patient assessment data (Chapter 2), development of the new algorithms, including testing of multiple specifications (Chapter 3), and validation of the new algorithms (Chapter 4). The report then describes a comparison of the new algorithms with existing algorithms to predict claims-based outcomes in a separate sample of the broader Medicare population (Chapter 5). Drawing on these findings, we recommend the addition of CFI scores to the CCW and suggest future applications of the scores and avenues of research (Chapter 6).

2. Development of the Study Population

Overview

This chapter describes the study population used to develop and refine a new claims-based frailty algorithm by linking Medicare claims with PAC assessment data. Out of 35,141,239 Medicare FFS beneficiaries with at least six months of continuous enrollment, 24 percent received SNF or HHA care during the study period.

Linking Claims Data with Post-Acute Care Assessment Data

The primary study population consisted of Medicare beneficiaries with at least six months of Medicare FFS enrollment between July 1, 2014, and December 31, 2017. We required individuals with PAC assessment data to have HHA or SNF stays with both admission and discharge dates during the period spanning 2015 to 2017. We also required PAC patients to have at least six months of continuous enrollment prior to PAC admission to serve as a reference period—that is, the time period from which we extracted the predictors—to identify diagnoses for algorithm development. For individuals with multiple eligible PAC stays, we selected one random stay. For individuals with no PAC stays (hereafter *non-PAC*), we selected a random six-month period of continuous enrollment.

Our study sample had 35,141,239 Medicare FFS beneficiaries. Eighteen percent were in the HHA group, 6 percent were in the SNF group, and 76 percent were in the non-PAC group.

To better understand the differences between the HHA, SNF, and non-PAC populations, we examined descriptive statistics from the three groups. Nearly all (99.3 percent) HHA patients had complete PAC admission assessment data for all outcomes of interest (described in detail in Chapter 3), while only 79.4 percent of SNF beneficiaries had complete PAC admission assessment data for all outcomes of interest (Table 2.1). Beneficiaries might be missing all admission assessment variables if they are discharged, transferred, or die before the assessment can take place. Individual items might also be missing if the clinician was unable to observe the specific outcome being assessed. HHA and SNF beneficiaries were generally older, more likely to be dually eligible for Medicare and Medicaid and receive Part D low-income subsidies, and more likely to have an inpatient stay during the reference period. SNF patients were more likely to be White (85.8 percent) as compared with HHA or non-PAC beneficiaries (81.4 percent and 80.1 percent, respectively). Non-PAC beneficiaries were more likely to have an “other” or “unknown” race (6.4 percent) compared with HHA and SNF beneficiaries (2.1 percent and 2.5 percent, respectively). There were also regional differences in use of HHA and SNF services, with beneficiaries in the South disproportionately using HHA services and beneficiaries in the

Midwest disproportionately using SNF services. Neighborhood socioeconomic status, as measured by the area deprivation index (ADI; Kind et al., 2014), was similar for all three groups.

Table 2.1. Population Using Six-Month Reference Periods[a] 2014–2017

	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
N	6,421,516	2,082,220	31,942,637
Complete PAC assessment data,[d] *N* (%)	6,378,403 (99.3%)	1,654,484 (79.4%)	N/A
Inpatient stay within 14 days of end of reference period, *N* (%)	2,900,067 (45.2%)	1,751,367 (84.1%)	189,199 (0.6%)
Inpatient stay at any point during six-month reference period, *N* (%)	4,135,971 (64.4%)	1,901,052 (91.3%)	1,221,737 (3.8%)
Age,[e] mean (SD)	76.29 (11.95)	78.82 (11.04)	69.41 (12.14)
> 65,[d] *N* (%)	5,585,085 (87.0%)	1,885,367 (90.5%)	26,304,083 (82.3%)
Male, *N* (%)	2,583,854 (40.2%)	876,553 (42.1%)	15,397,889 (48.2%)
Medicaid enrolled, full or partial,[e] *N* (%)	1,622,044 (25.3%)	534,446 (25.7%)	5,157,126 (16.1%)
Low-income subsidy (Part D),[e] *N* (%)	1,867,263 (29.1%)	605,412 (29.1%)	5,950,722 (18.6%)
ADI,[e] mean (SD)	51.13 (18.69)	50.58 (18.32)	50.50 (18.64)
Region[e]			
Midwest, *N* (%)	1,313,801 (20.5%)	570,417 (27.4%)	6,947,719 (21.8%)
Northeast, *N* (%)	1,242,574 (19.4%)	433,882 (20.8%)	5,673,059 (17.8%)
South, *N* (%)	2,763,380 (43.0%)	733,221 (35.2%)	11,851,416 (37.1%)
West, *N* (%)	1,076,194 (16.8%)	334,988 (16.1%)	6,366,290 (19.9%)
Race			
Non-Hispanic White, *N* (%)	5,227,566 (81.4%)	1,786,034 (85.8%)	25,583,980 (80.1%)
Non-Hispanic Black, *N* (%)	747,633 (11.6%)	197,911 (9.5%)	2,932,306 (9.2%)
Hispanic, *N* (%)	166,814 (2.6%)	29,327 (1.4%)	698,960 (2.2%)
Asian, *N* (%)	119,514 (1.9%)	25,850 (1.2%)	680,602 (2.1%)
Other or unknown, *N* (%)	159,989 (2.5%)	43,098 (2.1%)	2,046,789 (6.4%)

NOTE: N/A = not applicable; SD = standard deviation.
[a] All beneficiaries included in population must have continuous enrollment data during reference period.
[b] A randomly selected stay out of those that have an eligible reference period.
[c] Six-month reference periods from beneficiaries with no HHA, SNF, or other type of PAC stay; randomly selected and matched on time only to a visit; no more than one reference period per beneficiary.
[d] See Table A.8 for a list of PAC assessment items of interest.
[e] As of the end of the six-month reference period.

Development of Inverse Probability Weights

To address the fundamental issue of developing algorithms applicable to the entire Medicare population when only those with PAC stays had outcome data, we used claims-based variables to develop inverse probability weights that were applied with the goal of making the beneficiaries in the algorithm development group more closely resemble the overall population. To develop the weights, we first predicted the probability of a patient having a PAC stay in which complete outcome data was collected, then used the resulting inverse probability weights in the models predicting functional impairment among those with PAC assessment data. These weights helped ensure that the PAC population better represented the overall population of Medicare beneficiaries and that beneficiaries who were highly unlikely to be in the control group (described next) had less influence on the estimation of regression coefficients. First, individuals with complete PAC admission assessment data for the outcomes of interest were considered to be in the *PAC outcome* group, and all others in the sample (i.e., non-PAC patients and PAC patients with incomplete outcome data) were considered to not have the PAC outcome and were included in the *control* group. Predictors in logistic regression analysis to predict the binary PAC outcome consisted of the full set of candidate predictors (described in detail in Chapter 3); age (included both as a continuous variable and in five-year age categories to allow for nonlinear relationships); race category; ADI; Medicaid enrollment and Part D low-income subsidy indicators; and state, month, and year fixed effects. To verify that there was adequate overlap of predicted probabilities of a PAC outcome between the SNF, HHA, and control groups, distributions were plotted and compared for each of the three groups (Figure 2.1).

Most of the PAC outcome group had higher predicted probabilities of a PAC outcome, while most of the non-PAC outcome group had lower ones. For example, 91 percent of the control group had below a 20 percent probability of a PAC outcome, while only 18 percent of the PAC outcome group had below a 20 percent probability of a PAC outcome (Table 2.2). However, we still found that there were sufficient HHA and SNF patients with reference group characteristics similar to control patients to make inverse probability weighting reasonable.

Figure 2.1. Predicted Probability of Post-Acute Care Assessment Data by Group

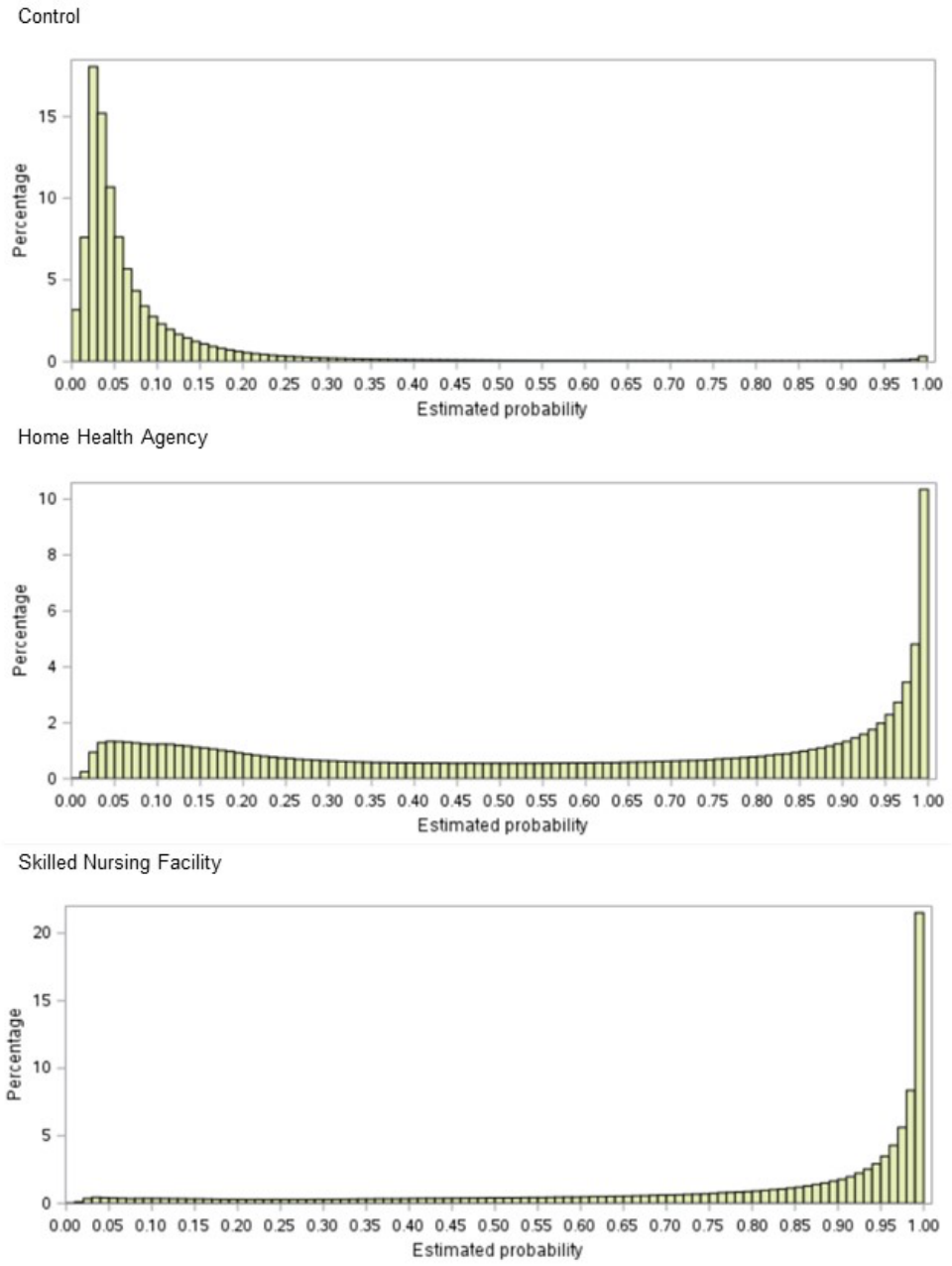


Table 2.2. Predicted Probability of Post-Acute Care Outcome by Group

Group	< 0.2 Predicted Probability of Post-Acute Care Outcome (*N*, %)	> 0.8 Predicted Probability of Post-Acute Care Outcome (*N*, %)
Control (*N* = 32,374,928)	29,545,998 (91%)	442,836 (1%)
HHA with PAC outcome (*N* = 6,378,403)	1,353,782 (21%)	2,652,624 (42%)
SNF with PAC outcome (*N* = 1,693,042)	107,919 (6%)	1,131,222 (67%)
PAC outcome (HHA and SNF combined)	1,461,701 (18%)	3,783,846 (47%)

3. Development of Algorithms to Predict Functional Impairment

Overview

This chapter describes how RAND developed new claims-based frailty algorithms to predict functional impairment outcomes. The predictors used in the algorithms were empirically selected from a pool of candidate predictors that consisted of potentially disabling conditions from the CCW and predictors from the Faurot algorithm and Kim algorithm. The key functional impairment outcomes—(1) activity and mobility limitations and (2) memory limitation—were developed from multiple patient assessment items using factor analysis. The new RAND algorithms retained 111 out of 186 candidate predictors associated with activity and mobility limitations and 144 out of 186 candidate predictors associated with memory limitation.

Selection of Candidate Predictors

Drawing on a review of the literature, discussions with ASPE, and feedback from the PATF, we developed a pool of potential claims-based predictors to empirically select for inclusion in our algorithms. These consist of the International Classification of Diseases, versions 9 and 10 (ICD-9 and ICD-10)– and the Healthcare Common Procedure Coding System (HCPCS)–based sets of indicators developed by Faurot (Faurot et al., 2015) and Kim (Kim et al., 2018; Gautam et al., 2021; Table A.2 and Table A.3, respectively) and indicators of chronic and disabling conditions from the CCW (Table A.4). Using feedback from the PATF, we tested two versions of each ICD-9– and ICD-10–based predictor: one requiring at least one inpatient diagnosis within the last 14 days of the reference period (referred to as *proximal*) and a second requiring at least two diagnoses on different days from any place of service during the reference period, *except* for the last 14 days (referred to as *multiple prior*). This approach is somewhat different from that used by Kim et al., 2018, and Faurot et al., 2015, because of the difference in timing of outcome measurement in our study. The frailty outcome measures used by Kim et al., 2018, and Faurot et al., 2015, come from the MCBS, which is self-reported as part of a panel survey of participants and not tied to a specific acute event. As previously shown in Table 2.2, beneficiaries usually (though not always) received admission PAC assessments following hospitalization for an acute event. This acute event might be related to the beneficiary's functional impairment at assessment, independent of other chronic conditions also affecting functional impairment. We adopted the approach of including both proximal and multiple prior versions of predictors to distinguish between acute conditions warranting medical attention just prior to the assessment and chronic presentations of the predictors because these might be differentially related to functional impairment at the time of assessment. Durable medical equipment (DME) and other HCPCS-based indicators only required one instance during the

reference period. In total, our candidate predictors consisted of 35 from Faurot et al., 2015, 43 from Kim et al., 2018, and 90 from the CCW. Frequencies of Faurot, Kim, and CCW candidate predictors by PAC group are displayed in Tables A.5, A.6, and A.7, respectively.

Development of Outcomes Based on Post-Acute Care Assessment Data

Before selecting variables among our pool of candidate predictors, we first needed to define the functional impairment outcome variables that could later be used in our Lasso regressions. We used PAC assessment data from SNF patients who were assessed by a clinician using Minimum Data Set (MDS) Version 1.15 admission assessment items and HHA patients who were assessed using Outcome and Assessment Information Set (OASIS) Version C2 (CMS, 2019a; CMS, 2019b). These two instruments have similar but not identical assessment items, necessitating the creation of a crosswalk between items of interest in the two instruments based on conceptual similarities (Table A.8). Frequencies of items were examined to determine whether relative item difficulty was consistent between instruments (i.e., the rank of percentages of beneficiaries with each specific type of functional limitation was the same between instruments). We also verified that beneficiaries with SNF stays were more likely than the HHA beneficiaries to have functional limitations as identified by the crosswalked items, in line with expectations about the two populations.

To determine whether the items represented distinct aspects of frailty or functional impairment, we first compared polychoric correlations (a measure of agreement between ordinal variables) between all seven items using consolidated response levels (Table A.8). Next, we conducted exploratory factor analyses on the seven items using oblique rotation and principal factor analysis in SAS 9.4. The appropriate number of factors was determined by considering eigenvalues greater than one, greater than 70 percent variance explained, visual inspection of scree plots, and expert input from the PATF on the theoretical number of domains. Items were considered part of a factor if they had communalities above 0.4. All analyses were conducted separately in HHA and SNF groups to determine whether similar factor structures were observed in the two settings.

After examining the previously mentioned criteria, we determined that the items were best represented by two factors. Rotated factor patterns for a two-factor solution are presented in Table 3.1. *Toilet use*, *transferring*, *dressing*, *mobility*, and *grooming or hygiene* all clearly load onto Factor 1 while *memory and recall* clearly loads onto Factor 2, with results very consistent between settings. The *feeding and eating* item has communalities greater than 0.4 for Factor 1 in both the HHA and SNF settings and for Factor 2 in the SNF setting. Feedback from the PATF obtained prior to conducting the factor analysis suggested that *memory and recall* was a conceptually distinct item from the activity limitation items. Internal consistency for the mobility and activity limitations items was high (ordinal alpha of 0.95 and 0.94 for HHA and SNF populations, respectively). Using the results of the factor analysis and feedback from the PATF,

we determined that assessment items were best represented by two factors: *activity and mobility limitations*, aggregated as a count of limitations ranging from zero to six, and *memory limitation*, a single item dichotomized as no memory deficit or some degree of memory deficit.

Table 3.1. Rotated Factor Patterns in Home Health Agency and Skilled Nursing Facility Populations

	Home Health Agency		Skilled Nursing Facility	
	Factor 1	Factor 2	Factor 1	Factor 2
Toilet use	**0.959**	–0.184	**0.957**	–0.086
Transferring	**0.946**	0.025	**0.968**	0.006
Dressing	**0.921**	0.066	**0.933**	0.059
Mobility	**0.917**	0.018	**0.861**	–0.053
Grooming or hygiene	**0.889**	0.101	**0.859**	0.167
Feeding and eating	**0.631**	0.234	**0.410**	**0.450**
Memory and recall	0.002	**0.979**	–0.070	**0.961**

NOTE: Bold text indicates the factor(s) onto which the item loaded.

The weighted and unweighted frequency of each type of limitation and the number of weighted and unweighted activity and mobility limitations in the HHA and SNF populations are presented in Tables 3.2 and 3.3, respectively.

Table 3.2. Frequency of Beneficiaries with Different Types of Limitations by Post-Acute Care Setting[a]

	Home Health Agency (Unweighted, %)	Skilled Nursing Facility (Unweighted, %)	Home Health Agency (Weighted, %)	Skilled Nursing Facility (Weighted, %)
Toilet use[b]	90.58	98.41	89.23	97.58
Transferring[b]	94.53	98.33	93.03	97.39
Dressing[b]	95.15	98.38	94.36	97.68
Mobility[c]	5.89	10.80	5.00	9.70
Grooming or hygiene[b]	90.05	96.41	88.78	95.91
Feeding and eating[b]	64.49	71.19	64.24	71.68
Memory and recall[d]	16.24	33.10	15.32	35.26

[a] Frequencies and percentages are out of patients with complete outcome data.
[b] *Partial* or *full limitation* as defined in Table A.8.
[c] *Unable to walk with assistance* and *unable to wheel self in chair* as defined in Table A.8.
[d] *Some degree of memory deficit* as defined in Table A.8.

Table 3.3. Frequency of Beneficiaries with Partial or Total Dependence in Activity or Mobility Limitation by Post-Acute Care Setting

Number of Limitations	Home Health Agency (Unweighted, %)	Skilled Nursing Facility (Unweighted, %)	Home Health Agency (Weighted, %)	Skilled Nursing Facility (Weighted, %)
0	1.38	0.36	2.85	0.93
1	1.74	0.28	2.67	0.69
2	2.27	0.51	3.03	0.99
3	4.24	2.13	5.31	2.54
4	26.54	23.83	25.99	23.14
5	56.66	63.22	55.63	63.86
6	7.16	9.67	4.52	7.85

Alternate Definition of Activity and Mobility Limitations Outcome

We found poor discrimination in the number of activity and mobility limitations; most beneficiaries had four or five limitations. We hypothesized that the lack of variation in the activity and mobility limitations outcome might have occurred because of the crosswalk between functional impairment as measured by (1) OASIS items for HHA patients and (2) MDS items for SNF patients. Because of differences in item wording defining *partial limitation* categories between OASIS and MDS, we collapsed several response categories into dichotomous responses that simply indicated presence or absence of any limitation. Although this allowed us to create more analogous items between instruments, considerable detail in level of functional limitation was lost. We sought to encode more information about the severity of activity and mobility limitations by creating a data-driven summary of both the number and severity of activity and mobility limitations via factor analysis.

First, we identified a subset of 573,703 beneficiaries (9 percent of PAC patients) who were assessed using MDS and then using OASIS within a span of fourteen days. The median time between MDS and OASIS assessments in these patients was 2.3 days. We thus had data on all items from both assessments on this group of patients measured within a reasonably short timeframe.

Second, we used the mirt package (Chalmers et al., 2021) in R to fit a maximum likelihood Multidimensional Item Response Theory (MIRT) factor analysis model to the data. We used a single factor, graded response model and a standard expectation maximization estimation method. The goal of this step was to identify whether the items from both instruments measured a single underlying activity and mobility limitations domain that could be used as a new outcome for algorithm development and capture more variation in the population. We found that all item factor loadings were greater than or equal to 0.45 and McDonald's omega was 0.93, indicating high internal consistency reliability. However, other fit statistics did not indicate a good model

fit. The root mean squared error of approximation (RMSEA) was 0.225, the Tucker-Lewis index (TLI) was 0.467, and the comparative fit index was 0.598. Generally, RMSEA less than 0.06 and comparative fit index and TLI greater than 0.95 are considered to indicate good model fit (Hu and Bentler, 1999). A two-factor model would provide a better fit, but we found that this resulted in a MDS factor and an OASIS factor; thus, the model could not be used in subsequent analyses. Given these limitations, we chose to proceed with a one-factor model.

Next, using the model created in the prior factor analysis step, we computed a single factor score estimate using the expected *a posteriori* estimation method for beneficiaries with any PAC assessment outcomes. This factor score maximally captured the information in all items from both assessments, without dichotomizing or collapsing the information in individual items. This also allowed us to obtain factor scores representing a single activity and mobility limitations domain for all patients, even those with only one assessment. We assessed whether the factor score was adequately related to all items by examining correlations between the score and individual MDS and OASIS items. This step told us whether the factor score was capturing relevant information from all items or whether it was related to only a few of the items and less related to the others. We plotted distributions of the factor scores for HHA and SNF patients, as well as patients with both assessments.

We found that the estimated factor score was well correlated with all of the individual items (Table 3.4), and correlations ranged from 0.42 (MDS item G0110H, eating) to 0.75 (OASIS item M1810, current ability to dress upper body). The factor score was comparatively less related to both eating items (correlations of 0.42 and 0.48) than to all others. The rest of the items had at least a 0.63 correlation with the limitation score, and five of the items had correlations over 0.7. This suggested that the estimated factor score captured information relatively equally from all items, and—because these correlations are to the uncollapsed items—it captured information about the full spectrum of severity.

Factor scores by default are calculated with a mean of zero and a standard deviation of one, with higher scores reflecting greater functional impairment. Although we calculated factor scores for the entire PAC population, to develop a model using factor score as the outcome, we restricted the population to those with complete PAC assessment data (94 percent of the total PAC population). This allowed us to more directly compare our new model with our baseline model.

Table 3.4. Correlation of Estimated Factor Score with Individual Activity and Mobility Limitation

Assessment	Item	Correlation with Factor Score
OASIS (HHA)	M1800. Grooming	0.73
	M1810. Current Ability to Dress Upper Body	0.75
	M1820. Current Ability to Dress Lower Body	0.69
	M1840. Toilet Transferring	0.71
	M1845. Toileting Hygiene	0.73
	M1850. Transferring	0.64
	M1870. Feeding or Eating	0.48
	M1860. Ambulation/Locomotion	0.65
MDS (SNF)	G0110J. Personal hygiene	0.66
	G0110G. Dressing	0.67
	G0110I. Toilet use	0.70
	G0110B. Transfer	0.71
	G0110H. Eating	0.42
	G0110E. Locomotion on unit	0.66
	G0110F. Locomotion off unit	0.63

SOURCE: Assessment items are from CMS, 2019a, and CMS, 2019b.

The PAC patients that we used to develop our algorithms had factor scores with a mean of –0.014 and a standard deviation of 0.96. Distributions of estimated factor scores are given in Figure A.1 for three samples: the sample of individuals with both SNF and HHA assessments within 14 days of one another that were used to construct the factor score; patients with only an HHA assessment; and patients with only a SNF assessment. The figure reassuringly shows what one would expect: Factor scores were higher for SNF patients (median score 0.1) than for HHA patients (median –0.1). The figure also shows that the distribution for patients with both assessments is quite similar to the distribution for HHA patients, who make up approximately 75 percent of the PAC sample. In subsequent steps developing and validating the algorithms, we tested both count of activity limitations using a conceptual crosswalk and the alternate factor score outcome.

Reducing the Pool of Candidate Predictors

Our overall approach was to use Lasso regression (Tibshirani, 1996), a supervised machine learning technique, to empirically reduce our pool of candidate predictors to those that best predict each of the previously described outcomes: memory limitation, count of activity and mobility limitations, and a factor score serving as an alternate to the activity and mobility limitations outcome. Variables were retained in the order of their importance to the explanatory

power of the model. For both outcomes, we introduced the full list of variables included in our inverse probability weights model with the exception of race and ADI decile. We included these variables in the development of the inverse probability weights to ensure generalizability of the sample but excluded them from the frailty model because we wanted estimates of frailty to focus primarily on clinical indicators. However, there might be important differences in the completeness and accuracy of diagnoses used in our claims-based indicators by race and area socioeconomic status, so these variables are examined in subsequent sensitivity analyses.

All models included the previously described inverse probability weights to make the PAC population more representative of the general population and were run on the PAC population with complete outcome data. The data set was divided into two sets: 70 percent for training and 30 percent for validation. Lasso regressions applied a shrinkage factor to coefficients based on the L1-norm of the set of coefficients to prevent model overfitting. This shrinkage factor caused some coefficients to be set to exactly zero, which also allowed Lasso to be used to reduce the number of variables included in the model. The amount of shrinkage was chosen by estimating a series of models on the training set and selecting the best model in terms of minimizing the prediction error on the validation set. Although models were developed on the PAC population, predicted probabilities were calculated for control, HHA with PAC outcome, and SNF with PAC outcome groups, and distributions for each were examined.

For the first model, we conducted logistic regressions using the dichotomous memory limitation outcome to predict probabilities of having some memory impairment. For the other models, we conducted linear regressions using a count of activity and mobility limitations (zero to six) and continuous factor score as the outcome, respectively. The variables retained for each model are summarized in Table 3.5. The memory limitation model retained the greatest number of variables, followed by the factor score model and the count of activity and mobility limitations model. Age as a continuous variable and dementia (multiple prior) were among the first five variables retained in all three models, while intellectual disabilities and related conditions (multiple prior) were among the five variables with the largest magnitude coefficient for all three models. In general, a larger proportion of variables from the Kim and Faurot algorithms were retained than from the CCW variables. Across all three models, the five variables with the largest magnitude coefficients were all positive with the exception of neurotic disorders, personality disorders, and other nonpsychotic mental disorders (proximal), which was negative. However, all models included predictors with negative coefficients. Notably, the memory limitation model included several predictors with highly negative coefficients (e.g., arthritis, arthropathies and related disorders, lipid abnormalities). Theoretically, we would not expect such variables to be protective of memory limitation, and this might be indicative of a limitation of using the PAC population. Because patients who enter a PAC setting should be expected to have some degree of mental or physical limitation, the absence of certain indicators in claims indicating physical limitations might mean that the patient is more likely to have a memory limitation—a relationship that might not hold in a non-PAC population. Alternatively, patients with high-

severity cognitive limitations might be less likely to have lower-severity conditions diagnosed, coded, or both in claims. Because some of these conditions have very high prevalence in the PAC population (e.g., hypertension), they might look protective in the models. Thus, in a subsequent sensitivity analysis, we fit a separate model only including indicators with positive coefficients in the initial model iteration. We describe this along with other alternate model specifications in Chapter 4.

Table 3.5. Summary of Variables Retained in RAND Frailty Algorithms

	Memory Limitation	Activity and Mobility Limitations	Factor Score
CCW variables retained (of 90)	70	43	58
Faurot variables retained (of 35)	33	27	29
Kim variables retained (of 43)	35	29	35
Total variables retained (of 173)[a]	143	104	127
First five variables retained	1. Dementia (MP) 2. Age, continuous 3. Hypertensive disease (MP) 4. Arthropathies and related disorders (MP) 5. Arthritis (MP)	1. Age, continuous 2. Ambulance/life support 3. Transportation services including ambulance 4. Dementia (MP) 5. Beneficiary dual status	1. Age, continuous 2. Ambulance/life support 3. Transportation services 4. Dementia (MP) 5. Hereditary and degenerative diseases of the central nervous system (MP)
Five variables with largest magnitude[b]	1. Dementia (MP) 2. Intellectual disabilities and related conditions (MP) 3. Dementia (P) 4. Stroke/brain injury (P) 5. Alzheimer's disease and related disorders or senile dementia (P)	1. Neurotic disorders, personality disorders, and other nonpsychotic mental disorders (MP) 2. Cerebral palsy (MP) 3. Mobility impairments (MP) 4. Paralysis (MP) 5. Intellectual disabilities and related conditions (MP)	1. Cerebral palsy (MP) 2. Intellectual disabilities and related conditions (MP) 3. Hospital beds and associated supplies 4. Mobility impairments (MP) 5. Paralysis (MP)

NOTE: MP = multiple prior; P = proximal.
[a] In addition to the CCW, Faurot, and Kim predictors, five candidate predictors consisting of dual eligibility status, low-income subsidy indictor, beneficiary sex, age in years, and age in five-year categories were included.
[b] Excluding continuous age and age group.

4. Validation of the Developed Algorithms

Overview

This chapter describes refinement of specifications and validation for RAND's new claims-based frailty algorithm—a combination of select predictors from the list of potentially disabling conditions in the CCW, the Faurot algorithm, and the Kim algorithm. We tested several alternate model specifications of the RAND algorithms and found minimal improvement to the prediction performance, except for increasing the reference period to capture predictors from six months to 12 months prior to the PAC admission date.

Initial Validation of Algorithms

We used multiple approaches to validate our algorithms. We compared the root mean squared error (RMSE) for each of our fitted models and compared predicted limitation outcomes to actual outcome results in the PAC population. For the memory limitation model, we used the predicted probabilities to calculate sensitivity, specificity, and positive predictive value (PPV) for different cutoff values for assessing whether a person is predicted to have a memory limitation. For the activity and mobility limitations model, we rounded predicted numbers of limitations to the nearest integer and compared them with the actual number of limitations experienced by the PAC population. For each of these approaches, calculations were only performed on the 30 percent of the data used for validation. Inverse probability weights described in Chapter 2 were applied for all validation analyses.

Table 4.1 presents sensitivity, specificity, and PPV for different cutoff values in the memory limitation model. Sensitivity decreases steadily as the cutoff increases, while specificity and PPV increase. If a cutoff of 0.2 for the predicted probability of memory limitation were used, both the majority of beneficiaries without memory limitation and the majority of beneficiaries with memory limitation would be identified correctly. This cutoff might be useful if the goal is to identify a large number of beneficiaries with memory limitation. However, the PPV is only 38.7 percent for this cutoff, meaning that less than half of individuals identified as having a memory limitation would actually have one. If the goal of this model is to identify individuals to target with more-intensive resources, a higher cutoff, such as 0.8, would ensure that most of the patients targeted (76.7 percent) actually had a memory limitation.

Table 4.1. Sensitivity, Specificity, and Positive Predictive Value of Memory Limitation Model for Different Predicted Probability Cutoffs

Cutoff	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)
0.2	56.52	81.87	38.68
0.3	37.22	93.15	52.37
0.4	27.41	96.63	62.19
0.5	22.42	97.73	66.61
0.6	17.09	98.48	69.48
0.7	10.73	99.19	72.75
0.8	4.69	99.71	76.70

Table 4.2 displays the number of mobility and activity limitations predicted for the model as compared with the actual number of limitations. The model only predicted between three and six limitations for the validation population. For those with four predicted limitations, the actual number of limitations was correct only 27.9 percent of the time. For those with five predicted limitations, the actual number of limitations was correct 61.7 percent of the time. Overall, the correct number of limitations was predicted 33.2 percent of the time. The predicted and actual decile of the alternative factor score outcome is presented in Table A.9. Performance of this model was similar, with the correct decile predicted only 13.7 percent of the time (compared with 10 percent of the time as expected by chance).

Table 4.2. Predicted and Actual Activity and Mobility Limitations, Weighted Number of Post-Acute Care Patients (Row Percentage)

Predicted Number of Limitations	Actual Number of Limitations						
	0	1	2	3	4	5	6
0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
3	1,804 (30.38%)	234 (3.94%)	408 (6.86%)	552 (9.30%)	998 (16.80%)	1,939 (32.65%)	4 (0.06%)
4	281,702 (3.66%)	234,414 (3.05%)	257,365 (3.35%)	449,354 (5.84%)	2,142,746 (27.87%)	4,046,739 (52.63%)	277,330 (3.61%)
5	13,763 (0.96%)	18,033 (1.26%)	19,952 (1.40%)	39,488 (2.77%)	291,846 (20.45%)	880,643 (61.69%)	163,724 (11.47%)
6	6 (0.05%)	4 (0.03%)	8 (0.07%)	41 (0.35%)	781 (6.60%)	5,639 (47.66%)	5,353 (45.24%)

NOTE: Gray shading indicates weighted number of beneficiaries with accurately predicted number of limitations.

Testing of Alternate Predictor Specifications

The baseline algorithms described in Chapter 3 and validated against functional impairment outcomes are potentially sensitive to various modeling decisions. Here we explore whether changes in the specification of predictors affect the prediction performance of the model. Specifically, we examined the following changes:

1. One-year reference periods: Our baseline model measured predictors for each patient over the six months prior to PAC admission. We fit models using predictors measured over the 12 months prior to PAC admission. Including a longer reference period has the potential to capture more-relevant comorbidities for each patient, but excludes beneficiaries with less than 12 months of continuous enrollment.
2. Removing indicators of DME and services: These predictors are based on HCPCS and procedure codes rather than diagnosis codes, which are used for the majority of the other predictors (see Table A.2 and Table A.3 for Kim and Faurot DME and other services indicators). If the algorithm were used to generate a measure of limitation that would be used as a case-mix adjuster in downstream analyses for costs or utilization, then one would not want it to be based on predictors that also directly measure utilization.
3. Remove indicators of mobility limitation: Mobility limitations might be disproportionately coded in the PAC population because they can be a prerequisite for PAC admission. Thus, diagnoses indicating mobility limitations might be coded more often for someone seeking PAC than someone in the general population not intending to use these services, even if they actually experience the same mobility limitations. We thus excluded these predictors from the model.
4. Only include positive coefficients: There are drawbacks to using a PAC sample to predict functional limitations. One issue is that nearly everyone admitted to a PAC has some type of physical or cognitive limitation, and this might induce selection bias when estimating certain predictor effects. For example, those without mobility limitations might have higher rates of memory limitation than those with mobility limitations because PAC patients are selected on having at least one limitation. This might incorrectly make having a mobility limitation appear protective against having a memory limitation. A separate but related issue is that a less severe condition, such as arthritis, might not be diagnosed when a more severe condition, such as paralysis, is present. In this scenario, less severe conditions might also incorrectly appear to be protective. We sought to limit this type of behavior in the model by eliminating predictors that were estimated to have a negative coefficient in the baseline models described in Chapter 3.
5. Include age and condition indicator interaction terms: The relationship between conditions and functional impairment might vary by age. For example, some indicators for acute injury might have a weaker relationship with functional impairment in younger populations that recover more quickly. Conversely, given that the under-65 population typically qualifies for Medicare because of disability, certain condition indicators might represent more severe or disabling conditions in the under-65 population than in the over-65 population. Thus, we performed the Lasso regression technique described in Chapter 3 and included interaction terms between an over-65 indicator and each of the condition indicators. We then tested the resulting model.

We refit all models using these alternative predictor specifications, and we compared their prediction performance with the baseline model. Table 4.3 lists the prediction performance for the baseline model and each of the sensitivity models with alternative predictor specifications. None of the models performed substantially better than the baseline model, nor did any of them perform substantially worse. In fact, only the model using one-year reference periods showed improvement for predicting memory limitation (RMSE of 0.331 compared with 0.343 for the baseline model), and none of the models improved on the baseline model for predicting the number of activity and mobility limitations. After reviewing the results of the alternate model specifications, we decided to use the one-year reference period going forward. The one-year reference period resulted in improved prediction of memory limitation as compared with the baseline model. Additionally, the one-year reference period is consistent with the reference period used by Kim and several CCW indicators. In the next chapter, we compare the performance of the newly developed RAND models with existing models from Kim and Faurot at predicting claims-based outcomes on a separate population of Medicare FFS beneficiaries.

Table 4.3. Comparison of Model Performance for Alternative Model Specifications

Group	RMSE for Memory Limitation Model	RMSE for Activity and Mobility Limitations Model
Baseline model	0.343	1.232
One-year reference periods	0.331	1.238
Remove DME and other utilization-based indicators	0.343	1.246
Remove *mobility limitations*	0.343	1.232
Positive-only coefficients	0.346	1.237
Age-condition interactions	0.343	1.237

5. Comparison with Existing Algorithms

Overview

This chapter describes comparative performance of models using scores from the new RAND algorithms, the Faurot algorithm, and the Kim algorithm in a separate set of Medicare FFS beneficiaries. The algorithms were used to predict three claims-based outcomes: number of hospitalizations, nursing facility stay, and number of days at home (alive, not in a hospital or nursing facility) in the year following the claims reference period. Overall, we found that the Kim model had the best performance at predicting the three claims-based outcomes of interest in the general Medicare population across most metrics and subpopulations. The RAND model was slightly better at predicting a subsequent nursing home stay among PAC patients than other algorithms.

Comparison of RAND, Kim, and Faurot Algorithms

After testing multiple specifications of our algorithms, we next sought to understand how our algorithms compared with existing frailty algorithms at predicting claims-based outcomes in the general Medicare FFS population. To test these algorithms, we first created a new sample population. For all Medicare FFS beneficiaries with at least 12 months of continuous enrollment during 2014 to 2016, we selected a random 12-month reference period of continuous enrollment. For each selected individual's claims reference period, we created RAND, Kim, and Faurot condition indicator variables. The RAND and Kim algorithms used claims from the entire 12-month reference periods, while the Faurot algorithm used claims from the last eight months of the reference period. The characteristics of each algorithm are summarized in Table 5.1.

Table 5.1. Characteristics of RAND, Kim, and Faurot Models

	RAND Memory Limitation	RAND Activity and Mobility Limitations	Kim Frailty	Faurot Frailty
Total number of predictors[a]	134	114	93	29
Types of predictors	Age, sex, *proximal* and *multiple prior* versions of indicators from Kim, Faurot, and CCW		Indicators of diagnoses and services	Age, sex, race, indicators of diagnoses, and services
Timeframe, frequency, and setting of diagnoses	***Proximal* indicators**: At least one inpatient claim in last two weeks of reference period ***Multiple prior* indicators**: At least two claims in the rest of the 12-month reference period		At least one claim in 12 months	At least one claim in eight months
Population for development	Medicare FFS beneficiaries with a PAC stay following ≥ 12 months of continuous enrollment		MCBS participants ≥ 65	MCBS participants > 65
Outcomes	Dichotomous memory limitation item	Count of activity and mobility limitations zero to six	Survey frailty index (deficit accumulation approach)	ADL dependency

[a] The total number of predictors retained differ slightly from those presented in Table 3.5 because the updated models use a 12-month reference period.

Using the coefficients generated by the algorithms described in Tables A.10 and A.11 and the coefficients reported by Kim and Faurot in their manuscripts, we calculated the following *predicted scores* for each beneficiary in a random 80 percent of the new sample to serve as our *training* set:

- predicted probability of having a memory limitation; possible values ranged from zero to one (RAND)
- predicted number of activity and mobility limitations; possible (noninteger) values ranged from zero to six (RAND)
- predicted factor score (an alternative to number of activity and mobility limitations, described in Chapter 4); values for the development population have a mean of zero and a standard deviation of one (RAND)
- predicted survey frailty index—representing the proportion of abnormalities present out of a total of 56 possible self-reported symptoms, diagnoses, and functional limitations in the MCBS—using the Kim algorithm; possible values ranged from zero to one (Kim)
- predicted probability of having at least one self-reported dependency in the six ADLs in the MCBS using the Faurot algorithm; possible values ranged from zero to one (Faurot).

Because their relationship with outcomes might be nonlinear, we also grouped each of these predicted scores based on deciles (hereafter, *decile* or *decile-based* predictor) and based on categories suggested by Kim (Kim et al., 2020) representing the less than or equal to 10th, 11th to 25th, 26th to 75th, 76th to 90th, and greater than 90th percentiles (hereafter, *categorical* predictor). Each percentile and decile was calculated from the population of Medicare FFS beneficiaries. Decile and categorical predictors were coded as categorical variables represented by nine and four dummy variables, respectively.

Next, we constructed three claims-based outcome measures for each beneficiary in the sample. Outcomes were based on the 12 months following the randomly selected 12-month reference period. These outcomes consisted of the following:

- number of hospitalizations
- nursing facility stay (yes or no)
- days at home—number of days beneficiary was alive, not in the hospital, and not in a nursing facility.

We next performed regressions in our training set using the three claims-based outcomes described earlier. For each outcome, we used the following model specifications:

- age (both as a continuous variable and in five-year age categories to allow for nonlinear relationships) and sex (this defines the baseline model)
- baseline model predictors plus continuous predicted probability of having a memory limitation (RAND)
- baseline model predictors plus continuous predicted number of activity and mobility limitations (RAND)
- baseline model predictors plus continuous predicted factor score (RAND)
- baseline model predictors plus continuous predicted Kim score (Kim)
- baseline model predictors plus continuous predicted Faurot probability (Faurot)
- baseline model predictors plus continuous predicted probability of having a memory limitation *and* continuous predicted number of activity and mobility limitations (RAND)
- baseline model predictors plus continuous predicted probability of having a memory limitation *and* continuous predicted factor score (RAND).

For each of the specifications described earlier (with the exception of the baseline model) we also developed

- versions of each of these models where the continuous predictor is replaced by the decile-based version of the predictor (e.g., continuous predicted Kim score replaced by decile-based predicted Kim score)
- versions of each of these models where the continuous predictor is replaced by the categorical version of the predictor.

For each model, we calculated RMSE and area under the curve (AUC) in the other 20 percent *validation* set of our sample. To calculate AUC for the number of hospitalizations outcome, we dichotomized the outcome as zero and one or more hospitalizations. This is a measure of the ability of a model to correctly identify a dichotomous outcome. AUC is measured on a scale of zero to one; higher numbers indicate better diagnostic ability, and 0.5 indicates a model that is no better than chance.

RMSE and AUC for each of the models tested are presented in Table 5.2. The Kim decile model performed the best overall. This model had the lowest RMSE for the number of hospitalizations outcome (0.718 compared with 0.747 for the baseline model) and the Kim continuous model had the highest AUC for the nursing facility outcome (0.882 compared with 0.754 for the baseline model). The RAND *memory + activity and mobility limitations* decile

model had the lowest RMSE for the nursing facility outcome (0.200) followed closely by the Kim decile model (0.201). The Kim continuous model and the Kim decile model had the lowest RMSE for the days home outcome and the highest AUC for the greater than or equal to 1 hospitalization outcome (0.174 and 0.734, respectively for both models). A key reason the Kim decile model is preferred to the Kim continuous model is that the Kim continuous model had a high RMSE (0.810) for the number of hospitalizations; this was in fact higher than the baseline model RMSE (0.747) for the same outcome. This was likely because there were a small number of Kim scores that were outliers with respect to the median. The median Kim score was 0.14 while the 99th percentile score was 0.39 and the maximum score was 0.70. These outlier scores created more variability and a higher RMSE for the continuous model but exerted less influence over the results in the decile and categorical models. This phenomenon could be seen, to a lesser extent, in some of the other models where those using continuous scores generally performed worse than those containing categorical or decile versions of the scores.

Overall, the RAND algorithms had several potential advantages over the Kim and Faurot algorithms: development using a large sample size, detailed functional impairment outcomes, and inclusion of beneficiaries younger than 65. The RAND algorithms might also better allow for concurrent risk-adjustment because our data included all claims up until the assessment date, whereas the Kim and Faurot algorithms used MCBS panel data resulting in up to a four-month gap between the end of the claims reference period and outcome measurement. However, our population also suffered from more-limited generalizability to the overall Medicare population because outcome data were only available for beneficiaries with PAC assessment data.

Table 5.2. Model Comparison on Claims-Based Outcomes in Full Medicare Populations

Model	Number of Hospitalizations (RMSE)	≥ 1 Hospitalization (AUC)	≥ 1 Nursing Facility Stay (RMSE)	≥ 1 Nursing Facility Stay (AUC)	Days Home (RMSE)	Number of Days Home (AUC)
Baseline (age and sex)	0.747	0.616	0.213	0.754	0.184	0.719
Memory, continuous	0.747	0.615	0.203	0.821	0.180	0.744
Memory, categorical	0.744	0.647	0.207	0.815	0.182	0.745
Memory, decile	0.743	0.649	0.207	0.818	0.182	0.748
Activity and mobility, continuous	0.743	0.645	0.204	0.852	0.178	0.780
Activity and mobility, categorical	0.737	0.656	0.202	0.854	0.178	0.785
Activity and mobility, decile	0.736	0.658	0.202	0.857	0.177	0.788
Factor score, continuous	0.743	0.635	0.206	0.842	0.178	0.778
Factor score, categorical	0.739	0.656	0.203	0.850	0.177	0.785
Factor score, decile	0.738	0.658	0.203	0.852	0.177	0.787
Memory + activity and mobility, continuous	0.741	0.650	0.202	0.857	0.178	0.780
Memory + activity and mobility, categorical	0.732	0.680	0.201	0.860	0.177	0.787
Memory + activity and mobility, decile	0.731	0.685	**0.200**[a]	0.864	0.177	0.790
Memory + factor score, continuous	0.742	0.638	0.203	0.848	0.178	0.778
Memory + factor score, categorical	0.734	0.675	0.202	0.856	0.177	0.786
Memory + factor score, decile	0.733	0.679	0.202	0.858	0.177	0.790
Kim score, continuous	0.810	**0.734**[a]	0.203	**0.882**[a]	**0.174**[a]	**0.825**[a]
Kim score, categorical	0.720	0.723	0.202	0.875	0.175	0.819
Kim score, decile	**0.718**[a]	**0.734**[a]	0.201	0.881	**0.174**[a]	0.823
Faurot score, continuous	0.757	0.675	0.205	0.869	0.176	0.804
Faurot score, categorical	0.730	0.672	0.201	0.864	0.176	0.801
Faurot score, decile	0.728	0.681	0.201	0.872	0.175	0.807

[a] The best performing model as indicated by the lowest RMSE or the highest AUC.

Through our analyses, we found that the Kim decile model had the best overall performance at predicting three claims-based outcomes of interest in the general Medicare population. However, the RAND *memory + activity and mobility limitations* decile model had a slightly lower RMSE for predicting a nursing facility stay in the following year. Thus, we next tested

both the Kim decile model and the RAND *memory + activity and mobility limitations* decile model to compare performance in subpopulations of interest.

Comparison on Models in Subpopulations of Interest

For some stratifications, we hypothesized that one algorithm might perform better than another in specific subpopulations. For example, the Kim algorithm did not include individuals under 65 in its development, whereas the RAND algorithm did, so we thought that the RAND algorithm might perform better in this group. For other stratifications, we were more interested in comparing the differences between a baseline model that just included demographic information with the Kim and RAND models by group. In these cases, we hypothesized that underreporting of diagnoses might be worse for certain subpopulations and would appear as variation in the magnitude of the difference between the baseline algorithm and the other algorithms. The definitions and rationale for examining these subpopulations are as follows:

- **ICD version**: In October 2015, the diagnosis coding system used in claims transitioned from ICD-9 to ICD-10. ICD-10 codes tend to contain significantly more granularity. Although crosswalking between diagnoses is fairly straightforward using CMS general equivalence mapping, in practice, certain diagnoses might be used more or less commonly following the transition, irrespective of true condition prevalence. We therefore examined performance separately among beneficiaries with reference periods entirely prior to October 1, 2015 (ICD-9 only), entirely on or after October 1, 2015 (ICD-10 only), and spanning the ICD-9 to ICD-10 transition (both).
- **Race/ethnic group**: A documented shortcoming of using claims-based indicators for proxies of health status is the tendency to underreport conditions among racial and ethnic minorities and other groups with more-limited health care utilization relative to health status (Obermeyer et al., 2019). We thus wanted to determine whether there were differences in algorithm performance between race/ethnic groups.
- **ADI**: As discussed previously, the ADI provides rankings of zip codes by socioeconomic disadvantage based on such factors as education and income (Kind et al., 2014). We were interested in examining algorithm performance by beneficiary ADI decile for reasons similar to those for examining differences by race/ethnic group. It is possible that beneficiaries in more-disadvantaged deciles might have lower health care utilization relative to health status and, as a result, that model performance might be poorer in these deciles.
- **Age group**: As discussed previously, there might be differences in the relationship between conditions and functional limitations at different ages. Of particular interest is the difference between the over- and under-65 populations. Medicare beneficiaries under the age of 65 typically qualify based on disability, whereas beneficiaries 65 and older can qualify solely on age, so we might expect to see differences in these populations. Furthermore, the RAND algorithms included beneficiaries under 65 in their development, whereas the Kim algorithm did not, so it is important to assess how these algorithms perform in the under-65 population.

- **PAC status**: Earlier in the chapter, we saw that the Kim algorithm generally performed better than the RAND algorithm at predicting three claims-based outcomes. We hypothesized that this was due to development on a more representative population of Medicare beneficiaries. Here, we assess the performance of the RAND and Kim algorithms separately among PAC patients. For this analysis, PAC patients were defined as those with a claim with a date of service during the reference period in an HHA or SNF setting as defined by place of service code.

We next performed regressions using the three claims-based outcomes as previously described separately in each of the subpopulations of interest. We did this for the baseline (age and sex model), the Kim decile model (i.e., the *Kim model*), and the RAND *memory + activity and mobility* limitations decile model (i.e., the *RAND model*). For each of the three models, we calculated RMSE and AUC on the 20 percent validation set.

Results showing model performance in subpopulations of interest for the number of hospitalizations, nursing facility stay, and days at home outcomes are shown in Table 5.3, Table 5.4, and Table 5.5, respectively. Overall, our results showed differences in RMSE and AUC between subpopulations of interest. In general, the results for subpopulations were consistent with our main findings: Models that performed better at predicting an outcome overall tended to also predict that outcome better across subpopulations.

Black beneficiaries had the lowest AUC for the Kim and RAND models across all three outcomes and the highest RMSE in both the RAND and Kim models for all but the *days at home* outcome. However, similar trends were seen in the baseline model, which did not include any indicators of health care utilization. We found a similar trend for ADI where more advantaged deciles tended to have better model performance, but the improvement from the baseline to models including RAND and Kim predictors was relatively consistent across ADI.

ICD-9 and ICD-10 versions of the models performed similarly. ICD-10 versions had a slightly lower RMSE across all three outcomes. There was a similar gap between the RAND and Kim models for ICD-9 and ICD-10 versions. There were no differences in which model performed better when stratifying by age group.

Although the Kim model performed better overall, the RAND model performed better on a few metrics, primarily with the nursing facility outcome. The RAND model had slightly lower RMSE for all ICD categories, among Black and White Medicare beneficiaries, among the seven most-disadvantaged deciles, among both age groups, and among both PAC and non-PAC patients for the nursing facility outcome. The Kim model had a higher AUC for all subpopulations using the nursing facility outcome with the exception of the PAC group. The only other metric where the RAND model was superior was for the *missing* ADI decile with the *days at home* outcome.

Table 5.3. Model Subpopulation Comparisons, Number of Hospitalizations

		RMSE			AUC		
Category	**Subpopulation**	**Baseline**	**Memory + Activity and Mobility Decile**	**Kim Decile**	**Baseline**	**Memory + Activity and Mobility Decile**	**Kim Decile**
All		0.747	0.731	**0.718**[a]	0.616	0.685	**0.734**[a]
ICD version	ICD-9	0.791	0.776	**0.761**[a]	0.613	0.676	**0.731**[a]
	ICD-10	0.694	0.678	**0.667**[a]	0.619	0.699	**0.734**[a]
	Both	0.729	0.713	**0.700**[a]	0.610	0.685	**0.731**[a]
Race/ethnic group	Asian	0.591	0.583	**0.566**[a]	0.640	0.706	**0.761**[a]
	Black	0.960	0.933	**0.914**[a]	0.586	0.687	**0.744**[a]
	Hispanic	0.789	0.771	**0.752**[a]	0.604	0.692	**0.767**[a]
	White	0.727	0.713	**0.700**[a]	0.617	0.682	**0.728**[a]
	Other	0.617	0.603	**0.589**[a]	0.604	0.687	**0.747**[a]
ADI decile (1 = most advantaged decile; 10 = least advantaged decile)	1	0.701	0.687	**0.674**[a]	0.645	0.701	**0.746**[a]
	2	0.747	0.732	**0.718**[a]	0.632	0.695	**0.740**[a]
	3	0.717	0.703	**0.691**[a]	0.624	0.687	**0.732**[a]
	4	0.736	0.721	**0.708**[a]	0.621	0.687	**0.731**[a]
	5	0.748	0.734	**0.720**[a]	0.621	0.684	**0.730**[a]
	6	0.746	0.731	**0.718**[a]	0.614	0.683	**0.729**[a]
	7	0.765	0.749	**0.736**[a]	0.612	0.681	**0.728**[a]
	8	0.771	0.755	**0.741**[a]	0.606	0.676	**0.724**[a]
	9	0.770	0.754	**0.741**[a]	0.600	0.672	**0.720**[a]
	10	0.801	0.783	**0.769**[a]	0.593	0.673	**0.722**[a]
	Missing	0.369	0.318	**0.252**[a]	0.541	0.680	**0.858**[a]
Age group	< 65	0.948	0.929	**0.903**[a]	0.552	0.642	**0.735**[a]
	≥ 65	0.700	0.685	**0.675**[a]	0.623	0.692	**0.734**[a]
PAC	Yes	1.363	1.306	**1.280**[a]	0.510	0.554	**0.586**[a]
	No	0.635	0.629	**0.617**[a]	0.602	0.653	**0.707**[a]

[a] The best performing model as indicated by the lowest RMSE.

Table 5.4. Model Subpopulation Comparisons, Nursing Home Facility Stay

		RMSE			AUC		
Category	**Subgroup**	**Baseline**	**Memory + Activity and Mobility Decile**	**Kim Decile**	**Baseline**	**Memory + Activity and Mobility Decile**	**Kim Decile**
All		0.213	**0.200**[a]	0.201	0.754	0.864	**0.881**[a]
ICD version	ICD-9	0.238	**0.224**[a]	0.225	0.747	0.852	**0.872**[a]
	ICD-10	0.180	**0.169**[a]	0.171	0.763	0.882	**0.887**[a]
	Both	0.204	**0.191**[a]	0.192	0.749	0.865	**0.880**[a]
Race/ethnic group	Asian	0.178	0.175	**0.168**[a]	0.764	0.862	**0.894**[a]
	Black	0.229	**0.211**[a]	0.214	0.684	0.856	**0.869**[a]
	Hispanic	0.185	0.179	**0.176**[a]	0.744	0.856	**0.895**[a]
	White	0.216	**0.203**[a]	0.204	0.761	0.865	**0.880**[a]
	Other	0.139	0.133	**0.131**[a]	0.727	0.857	**0.892**[a]
ADI decile (1 = most advantaged decile; 10 = least advantaged decile)	1	0.213	0.203	**0.202**[a]	0.755	0.850	**0.881**[a]
	2	0.217	0.206	**0.206**[a]	0.751	0.855	**0.879**[a]
	3	0.208	0.197	**0.196**[a]	0.763	0.864	**0.884**[a]
	4	0.211	**0.199**[a]	0.200	0.760	0.862	**0.880**[a]
	5	0.205	**0.194**[a]	0.195	0.766	0.866	**0.882**[a]
	6	0.208	**0.195**[a]	0.196	0.761	0.874	**0.886**[a]
	7	0.214	**0.200**[a]	0.202	0.760	0.870	**0.880**[a]
	8	0.230	**0.213**[a]	0.215	0.751	0.868	**0.878**[a]
	9	0.222	**0.206**[a]	0.209	0.749	0.867	**0.875**[a]
	10	0.216	**0.201**[a]	0.204	0.734	0.864	**0.872**[a]
	Missing	0.100	0.073	**0.052**[a]	0.646	0.772	**0.908**[a]
Age group	< 65	0.170	**0.162**[a]	0.164	0.611	0.809	**0.833**[a]
	≥ 65	0.221	**0.207**[a]	0.208	0.761	0.868	**0.884**[a]
PAC	Yes	0.424	**0.395**[a]	0.399	0.599	**0.699**[a]	0.687
	No	0.172	**0.162**[a]	**0.162**[a]	0.764	0.850	**0.867**[a]

[a] The best performing model as indicated by the lowest RMSE or the highest AUC.

Table 5.5. Model Subpopulation Comparisons, Days at Home

		RMSE			AUC		
Category	**Subgroup**	**Baseline**	**Memory + Activity and Mobility Decile**	**Kim Decile**	**Baseline**	**Memory + Activity and Mobility Decile**	**Kim Decile**
All		0.184	0.177	**0.174**[a]	0.719	0.790	**0.823**[a]
ICD version	ICD-9	0.245	0.233	**0.228**[a]	0.712	0.786	**0.818**[a]
	ICD-10	0.119	0.118	**0.116**[a]	0.715	0.795	**0.814**[a]
	Both	0.144	0.141	**0.140**[a]	0.704	0.781	**0.810**[a]
Race/ethnic group	Asian	0.164	0.162	**0.158**[a]	0.738	0.793	**0.817**[a]
	Black	0.187	0.179	**0.176**[a]	0.670	0.772	**0.808**[a]
	Hispanic	0.171	0.168	**0.165**[a]	0.721	0.780	**0.811**[a]
	White	0.187	0.180	**0.177**[a]	0.724	0.792	**0.824**[a]
	Other	0.138	0.134	**0.132**[a]	0.699	0.771	**0.807**[a]
ADI decile (1 = most advantaged decile; 10 = least advantaged decile)	1	0.174	0.168	**0.164**[a]	0.745	0.806	**0.837**[a]
	2	0.179	0.173	**0.170**[a]	0.732	0.798	**0.828**[a]
	3	0.176	0.170	**0.167**[a]	0.732	0.799	**0.832**[a]
	4	0.181	0.174	**0.171**[a]	0.728	0.795	**0.829**[a]
	5	0.182	0.175	**0.173**[a]	0.727	0.795	**0.827**[a]
	6	0.185	0.178	**0.175**[a]	0.718	0.790	**0.825**[a]
	7	0.188	0.180	**0.177**[a]	0.716	0.789	**0.823**[a]
	8	0.190	0.183	**0.179**[a]	0.711	0.784	**0.819**[a]
	9	0.192	0.184	**0.181**[a]	0.706	0.782	**0.816**[a]
	10	0.194	0.187	**0.184**[a]	0.692	0.773	**0.808**[a]
	Missing	0.175	**0.171**[a]	0.173	0.715	0.727	**0.731**[a]
Age group	< 65	0.148	0.146	**0.143**[a]	0.614	0.692	**0.767**[a]
	≥ 65	0.190	0.183	**0.180**[a]	0.725	0.799	**0.827**[a]
PAC	Yes	0.361	0.339	**0.330**[a]	0.603	0.642	**0.675**[a]
	No	0.150	0.147	**0.145**[a]	0.709	0.750	**0.783**[a]

[a] The best performing model as indicated by the lowest RMSE or the highest AUC.

6. Summary and Recommendations

In this report, we reviewed existing algorithms to predict frailty and functional impairment and attempted to improve on previous algorithms using a novel source of data. We developed new claims-based algorithms to predict two functional impairment outcomes: (1) memory limitation and (2) activity and mobility limitations. We tested numerous algorithm specifications and compared models using scores from our best performing algorithm with models including scores from the existing Kim and Faurot algorithms. Finally, we examined the performance of the two best models stratified by subgroups of interest. Overall, we found that the Kim model was the best at predicting the claims-based outcome measures of hospitalizations, nursing facility stays, and days at home across most measures of model fit and subgroups.

The RAND algorithms had several potential advantages over the Kim and Faurot algorithms in their development. The data set used to develop the RAND algorithms, which integrates Medicare claims data with PAC assessment data, contained substantially more beneficiaries than the MCBS data used to develop the Kim and Faurot algorithms. Furthermore, the RAND population included beneficiaries under the age of 65, allowing us to examine an important population with potentially high rates of functional impairment. Our data structure allowed us to include all claims up until the assessment date, whereas the MCBS used panel data that might result in up to a four-month delay between the end of the claims reference period and outcome measurement.

However, our data also had disadvantages. The primary disadvantage was that outcome data were only available for beneficiaries with PAC assessment data, who are not representative of the overall Medicare population. We attempted to mitigate this issue by weighting the population to be more representative, although our examination of weighted outcome distributions indicated that adding weights provided only modest improvements to representativeness. Another disadvantage was the use of different assessment items in the MDS and OASIS. We attempted to reconcile these differences using a crosswalk based on item face validity and through a factor analysis approach.

A final difference between the outcomes in the PAC assessment data used to develop the RAND algorithms and the outcomes in the MCBS used to calculate the Kim algorithm was that the former were assessed by a health care professional while the latter were self-reported. There is not a clear consensus in the literature as to whether self-reported or observed ADL outcomes are more accurate, and relative performance of the measures might depend on the assessors and study population. One study comparing the relationship between self-reported ADLs and observed ADLs with subsequent mortality found similar explanatory power for both measures: The self-reported measures had a slightly stronger association with mortality at low levels of disability, and observed measures had a slightly stronger association with mortality at high levels

of disability (Kuhn, Rahman, and Menken, 2006). Another study suggested that self-reported ADLs might be inaccurate among individuals with cognitive impairment (Sager et al., 1992).

The RAND algorithms incorporated several variables from the Kim and Faurot algorithms along with variables from the CCW and used similar methods to Kim et al., 2018, to empirically select the variables that best predicted functional impairment. Given the advantages and disadvantages of the data used for algorithm development, we decided it was reasonable to compare the relative performance of the algorithms in a data set representative of all Medicare FFS beneficiaries and predict a new set of claims-based outcomes. Overall, we found that the Kim model, which included the Kim frailty score in deciles, combined with age and sex variables was the best at predicting hospitalizations and days at home in the following year. The RAND and Kim models were more comparable at predicting nursing facility stays in the following year, and the RAND model performed better among PAC patients.

The ultimate goal of this work was to develop or identify one or more measures of frailty or functional impairment that could be added to the CCW for use by CMS, researchers, and other stakeholders. Given our project findings, we recommend that CMS make the following resources available:

1. **Kim CFI scores in the CCW**. We recommend that yearly and mid-year CFI scores be made available for all Medicare FFS beneficiaries with any enrollment data for the period of interest. The yearly and midyear format is consistent with other condition variables in the CCW and will help facilitate analyses using multiple CCW variables. These scores represent the predicted proportion (between zero and one) of abnormalities present out of a total of 56 possible self-reported symptoms, diagnoses, and functional limitations. A separate variable will indicate whether beneficiaries had continuous enrollment and paid claims during the reference period, facilitating interpretation of the scores.
2. **Guidance on using and interpreting Kim CFI scores**. This guidance will consist of detailed information on the calculation of the CFI scores, how they might be used in research and risk adjustment, recommended cutoffs for using the scores to predict risk of hospitalization, and considerations for applying scores to different populations.

We also recommend **exploring the value and feasibility of adding PAC assessment outcomes to the CCW** as a potential future line of work. In our research, we found that using PAC assessment outcomes to develop claims-based algorithms predicting functional impairment suffered from limited generalizability to the overall Medicare population. However, several members of the PATF suggested that adding select PAC assessment outcomes for applicable Medicare beneficiaries to the CCW could be an important future resource for researchers who want to further investigate the relationships between claims-based diagnoses and assessments of functional status.

Appendix. Additional Tables and Figure

These tables and the figure provide additional details on the development and validation of the RAND algorithms. Table A.1 provides information on members of the PATF. Tables A.2 through A.4 provide details on specifications of conditions and health services included in the pool of candidate predictors used to develop the algorithms. Tables A.5 through A.7 show the frequency of these candidate predictors in the study population. Table A.8 gives information on the outcomes used to validate the RAND algorithms. Table A.9 provides additional validation analyses for the RAND activity and mobility limitations algorithm. Tables A.10 through A.12 provide the full set of included predictors and coefficients for the RAND algorithms. Figure A.1 shows the distribution of factor scores (an alternate specification of the activity and mobility limitations outcome) by PAC status.

Table A.1. Members of the Project Advisory Task Force

Name	Affiliation
Emmanuelle Belanger, Ph.D.	Assistant Professor of Health Services, Policy and Practice, Center for Gerontology and Healthcare Research, Brown University
Yonatan Ben-Shalom, Ph.D.	Senior Researcher, Mathematica
Cynthia Boyd, M.D., M.P.H.	Professor of Medicine, Johns Hopkins University School of Medicine, Bloomberg School of Public Health
Rebecca Brown, M.D., M.P.H.	Assistant Professor of Medicine, Perelman School of Medicine, University of Pennsylvania
Anne Deutsch, Ph.D.	Senior Research Public Health Analyst, RTI International, Northwestern University Feinberg School of Medicine
Carlos Jackson, Ph.D.	Chief Data and Analytics Officer, Community Care of North Carolina
Dae Kim, M.D., M.P.H., Sc.D.	Geriatrician and Epidemiologist, Brigham and Women‘s Hospital, Assistant Professor of Medicine, Harvard University Medical School
Bruce Kinosian, M.D.	Associate Professor of Medicine, Hospital of the University of Pennsylvania
Riccardo Miotto, Ph.D.	Director of Machine Learning, Tempus Labs
Wayne Saltsman, M.D., Ph.D.	Medical Director, Senior Care Options, Commonwealth Care Alliance

Table A.2. Value Sets of Predictors from Faurot Algorithm

Variable	Associated ICD-9,[a] CPT, or HCPCS Codes
Home hospital bed	E0250, E0251, E0255, E0256, E0260, E0261, E0265, E0266, E0270, E0290, E0291–297, E0301–304, E0316
Wheelchair	E1050, E1060, E1070, E1083–1093, E1100, E1110, E1120, E1140, E1150, E1160, E1161, E1170, K0001–9
Home oxygen	E1390–1392, E0431, E0433–435, E0439, E0441–443
Ambulance/life support	A0426, A0427, A0428, A0429, A0999
Paralysis	342., 438.2, 438.3, 438.4, 438.5, 344., 781.4
Dementia	290., 294., 331., 333.90, 333.92, 333.99, 780.93, 438.0, 797
Cancer screening	V76.
Heart failure	428., 425., 429.0, 429.1, 429.3, 429.4
Lipid abnormality	272
Psychiatric	29., 311., 300.00, 310.
Vertigo	386., 780.4
Difficult walking	719.7, 781.2, 781.3, 438.85, v46.3
Parkinson's disease	332
Podiatric care	700., 703., 681.1
Rehabilitation services	V57.1, v57.21, v57.3, v57.89, v57.9[b]
Arthritis	719.0, 719.1, 719.4, 719.5, 719.9, 711., 715., 716.5, 716.6, 716.8, 716.9, 718., 725., 710., 712., 714.
Sepsis	01., 036. 038., 040.0, 041., 032.0, 032.1, 681.,682., 730., 031.0, 031.2, 790.7, 032.82, 032.83, 053.0, 053.13, 054.5, 136.3, 320.0, 785.4, 112.83, 112.81, 112.5
Stroke/brain injury	348., 430., 431., 432., 852., 853., 854., 349.82, 433.01, 433.11, 433.21, 433.31, 433.91, 434.01, 434.11, 434.91
Weakness	728.2, 728.87, 799.3, 728.2, 728.3, v49.84
Diabetes mellitus complications	250.4, 250.6, 250.7, 250.9

[a] CMS ICD-9 to ICD-10 crosswalk used (National Bureau of Economic Research, 2016).
[b] In place of the ICD-9 codes from Faurot et al., 2015, listed here, we used a current procedure terminology (CPT) and ICD-9 procedure code–based definition to address changing ways of using the given rehabilitation service codes from the ICD-9 to ICD-10 transition. We used CPT codes 97010–97039, 97110–97546, 97161–97164, 97165–97168, 97169–97172, 97750–97755, 97760–97763, 97799, 31579, 92507–92508, 92511–92512, 92520–92524, 92526, 92597, 92607–92617, 92626–92627, 96105, 96112, 92613, 96125, 97533, 97535, G0451, G0515, and ICD-9 procedure codes 93.0X–93.4X, 93.74–93.75, 93.83, and 93.89. CMS general equivalence mappings were used to obtain ICD-10 procedure codes (National Bureau of Economic Research, 2016).

Table A.3. Value Sets of Predictors from Kim Algorithm

Description of Claims-Based Variables	ICD-9, CPT, or HCPCS Codes[a]
Hospital beds and associated supplies	E0250–E0373
Wheelchairs, components, and accessories	K0001–K0462, K0669
Organic psychotic conditions	290–294
Hereditary and degenerative diseases of the central nervous system	330–338
Walking aids and attachments	E0100–E0159
Accessories for oxygen delivery devices	E1353–E1406
Other supplies, including diabetes supplies and contraceptives	A4244–A4290
Diabetic footwear	A5500–A5513
Other psychoses	295–299
Other forms of heart disease	420–429
Open wound of lower limb	890–897
Ischemic heart disease	410–414
Hypertensive disease	401–405
Cerebrovascular disease	430–438
Neurotic disorders, personality disorders, and other nonpsychotic mental disorders	300–316
Arthropathies and related disorders	710–719
Nursing facility care—subsequent	99308
Chronic obstructive pulmonary disease and allied conditions	490–496
Other bacterial diseases	030–041
Diseases of veins and lymphatics and other diseases of circulatory system	451–459
Pneumonia and influenza	480–487
Diseases of other endocrine glands	250–259
Other diseases of urinary system	590–599
Ill-defined and unknown causes of morbidity and mortality	797–799
Contusion with intact skin surface	920–924
Nephritis, nephrotic syndrome, and nephrosis	580–589
Transportation services, including ambulance	A0021–A0999

NOTE: CPT = current procedure terminology.
[a] ICD-9, CPT, and HCPCS codes are from Kim et al., 2018, and ICD-10 equivalents are used as described in Gautam et al., 2021.

Table A.4. Chronic Conditions Data Warehouse Categories

Chronic Conditions Data Warehouse Other Chronic or Potentially Disabling Conditions
ADHD/conduct disorders, hyperkinetic syndrome
Alcohol use disorders
Anxiety disorders
Autism spectrum disorders
Bipolar disorder
Cerebral palsy
Cystic fibrosis and other metabolic developmental disorders
Depressive disorders
Drug use disorders
Epilepsy
Fibromyalgia, chronic pain, and fatigue
HIV/AIDS
Intellectual disabilities and related conditions
Learning disabilities
Leukemias and lymphomas
Liver disease, cirrhosis, and other liver conditions
Migraine and chronic headache
Mobility impairments
Multiple sclerosis and transverse myelitis
Muscular dystrophy
Obesity
Other developmental delays
Peripheral vascular disease
Personality disorders
Posttraumatic stress disorder
Pressure and chronic ulcers
Schizophrenia
Schizophrenia and other psychotic disorders
Sensory: blindness and visual impairment
Sensory: deafness and hearing impairment
Spina bifida and other congenital anomalies of the nervous system
Spinal cord injury
Tobacco use
Traumatic brain injury and nonpsychotic mental disorders due to brain damage
Viral hepatitis

SOURCE: CCW, undated.
NOTE: ADHD = attention-deficit/hyperactivity disorder; HIV/AIDS = human immunodeficiency virus and acquired immunodeficiency syndrome. A list of ICD-9 and ICD-10 codes for each condition is available at CCW, undated.

Table A.5. Frequency of Faurot Candidate Predictors by Post-Acute Care Group

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
N =	6,421,516	2,082,220	31,942,637
Proximal: At least one inpatient claim with diagnosis within last 14 days of reference period			
Paralysis	60,257 (0.94%)	64,846 (3.11%)	3,231 (0.01%)
Dementia	131,657 (2.05%)	204,068 (9.80%)	6,025 (0.02%)
Heart failure	446,950 (6.96%)	314,223 (15.09%)	26,392 (0.08%)
Lipid abnormality	325,243 (5.06%)	116,351 (5.59%)	16,565 (0.05%)
Psychiatric	155,696 (2.42%)	110,194 (5.29%)	12,378 (0.04%)
Vertigo	13,021 (0.20%)	6,591 (0.32%)	695 (0.00%)
Difficulty walking	14,172 (0.22%)	13,199 (0.63%)	409 (0.00%)
Parkinson's disease	24,880 (0.39%)	29,811 (1.43%)	831 (0.00%)
Podiatric care	5,957 (0.09%)	2,396 (0.12%)	2,396 (0.12%)
Arthritis	595,881 (9.28%)	207,157 (9.95%)	10,678 (0.03%)
Sepsis	324,277 (5.05%)	250,933 (12.05%)	24,801 (0.08%)
Stroke/brain injury	208,987 (3.25%)	269,545 (12.95%)	18,651 (0.06%)
Weakness	5,993 (0.09%)	7,187 (0.35%)	208 (0.00%)
Diabetes mellitus complications	184,649 (2.88%)	106,885 (5.13%)	8,723 (0.03%)
Multiple prior: At least two claims on separate dates (any type) with diagnosis at any point during the reference period, excluding the last 14 days			
Paralysis	225,992 (3.52%)	91,606 (4.40%)	84,967 (0.27%)
Dementia	742,231 (11.56%)	326,212 (15.67%)	656,681 (2.06%)
Cancer screening	136,372 (2.12%)	30,920 (1.48%)	934,807 (2.93%)
Heart failure	1,097,348 (17.09%)	448,405 (21.53%)	757,790 (2.37%)
Lipid abnormality	1,890,370 (29.44%)	567,769 (27.27%)	5,194,787 (16.26%)
Psychiatric	960,694 (14.96%)	362,296 (17.40%)	1,824,621 (5.63%)
Vertigo	243,020 (3.78%)	77,707 (3.73%)	387,560 (1.21%)
Difficulty walking	1,072,412 (16.70%)	403,118 (19.36%)	572,780 (1.79%)
Parkinson's disease	142,401 (2.22%)	55,275 (2.65%)	117,227 (0.37%)

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
Podiatric care	323,770 (5.04%)	119,540 (5.74%)	528,208 (1.65%)
Arthritis	2,165,696 (33.73%)	649,170 (31.18%)	3,248,300 (10.17%)
Sepsis	649,386 (10.11%)	287,308 (13.80%)	365,837 (1.15%)
Stroke/brain injury	502,377 (7.82%)	234,468 (11.26%)	214,862 (0.67%)
Weakness	951,777 (14.82%)	381,880 (18.34%)	461,839 (1.45%)
Diabetes mellitus complications	896,046 (13.95%)	305,710 (14.68%)	1,302,816 (4.08%)
At least one claim (any type) with DME *and* other services *code at any point during the reference period*			
Home hospital bed	185,040 (2.88%)	44,944 (2.16%)	24,138 (0.08%)
Wheelchair	309,702 (4.82%)	77,136 (3.70%)	53,842 (0.17%)
Home oxygen	621,871 (9.68%)	177,079 (8.50%)	472,251 (1.48%)
Ambulance/life support	2,544,050 (39.62%)	1,529,894 (73.47%)	1,018,669 (3.19%)

[a] Predictor codes are listed in Table A.2.
[b] Use a randomly selected stay out of those that have an eligible reference period.
[c] One randomly selected six-month reference period of continuous enrollment from beneficiaries with no HHA, SNF, or other type of PAC stay.

Table A.6. Frequency of Kim Candidate Predictors by Post-Acute Care Group

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
N =	6,421,516	2,082,220	31,942,637
Proximal: At least one inpatient claim with diagnosis within last 14 days of reference period			
Organic psychotic conditions	67,259 (1.05%)	111,426 (5.53%)	5,166 (0.02%)
Hereditary and degenerative diseases of the central nervous system	93,907 (1.46%)	101,817 (4.89%)	5,015 (0.02%)
Other psychoses	31,778 (0.49%)	26,181 (1.26%)	5,095 (0.02%)
Other forms of heart disease	777,957 (12.11%)	522,959 (25.12%)	45,943 (0.14%)
Open wound of lower limb	2,570 (0.04%)	1,525 (0.07%)	56 (0.00%)
Ischemic heart disease	306,720 (4.78%)	153,213 (7.36%)	22,957 (0.07%)
Hypertensive disease	977,206 (15.22%)	464,008 (22.28%)	52,042 (0.16%)
Cerebrovascular disease	132,707 (2.07%)	121,709 (5.85%)	10,207 (0.03%)
Neurotic disorders, personality disorders, and other nonpsychotic mental disorders	43,102 (0.67%)	20,755 (1.00%)	3,784 (0.01%)
Arthropathies and related disorders	600,867 (9.36%)	210,093 (10.09%)	10,878 (0.03%)
Chronic obstructive pulmonary disease and allied conditions	359,206 (5.59%)	173,185 (8.32%)	21,544 (0.07%)
Other bacterial diseases	55,464 (0.86%)	50,890 (2.44%)	11,623 (0.04%)
Diseases of veins and lymphatics, and other diseases of circulatory system	161,864 (2.52%)	107,458 (5.16%)	9,615 (0.03%)
Pneumonia and influenza	231,797 (3.61%)	170,797 (8.20%)	21,647 (0.07%)
Diseases of other endocrine glands	478,409 (7.45%)	261,902 (12.58%)	24,678 (0.08%)
Other diseases of urinary system	217,365 (3.38%)	232,446 (11.06%)	12,211 (0.04%)
Ill-defined and unknown causes of morbidity and mortality	27,290 (0.42%)	29,128 (1.40%)	2,694 (0.01%)
Contusion with intact skin surface	10,114 (0.16%)	12,143 (0.58%)	358 (0.00%)
Nephritis, nephrotic syndrome, and nephrosis	529,371 (8.24%)	425,217 (20.42%)	40,604 (0.13%)

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
Multiple prior: At least two claims on separate dates (any type) with diagnosis at any point during the reference period, excluding the last 14 days			
Organic psychotic conditions	355,328 (5.53%)	175,667 (8.44%)	351,318 (1.1%)
Hereditary and degenerative diseases of the central nervous system	647,681 (10.09%)	247,026 (11.16%)	685,706 (2.15%)
Other psychoses	448,683 (6.99%)	176,053 (8.46%)	1,124,765 (3.52%)
Other forms of heart disease	2,070,302 (32.24%)	787,095 (37.8%)	2,411,860 (7.55%)
Open wound of lower limb	84,484 (1.32%)	36,363 (1.88%)	39,686 (0.12%)
Ischemic heart disease	1,292,515 (20.13%)	452,112 (21.71%)	1,726,675 (5.41%)
Hypertensive disease	3,749,031 (58.38%)	1,223,145 (58.74%)	7,729,786 (24.20%)
Cerebrovascular disease	615,228 (9.58%)	244,309 (11.73%)	530,297 (1.66%)
Neurotic disorders, personality disorders, and other nonpsychotic mental disorders	364,772 (5.68%)	138,879 (6.67%)	660,382 (2.07%)
Arthropathies and related disorders	2,324,162 (36.19%)	709,517 (34.08%)	3,488,383 (10.92%)
Chronic obstructive pulmonary disease and allied conditions	1,235,718 (19.24%)	429,183 (20.61%)	1,658,315 (5.19%)
Other bacterial diseases	148,691 (2.32%)	73,613 (3.54%)	63,771 (0.20%)
Diseases of veins and lymphatics, and other diseases of circulatory system	663,147 (10.33%)	265,663 (12.76%)	550,975 (1.72%)
Pneumonia and influenza	461,743 (7.19%)	213,154 (10.24%)	258,001 (0.81%)
Diseases of other endocrine glands	1,944,948 (30.29%)	645,073 (30.98%)	4,340,561 (13.59%)
Other diseases of urinary system	949,172 (14.78%)	372,231 (17.88%)	987,537 (3.09%)
Ill-defined and unknown causes of morbidity and mortality	416,582 (6.49%)	199,803 (9.60%)	210,068 (0.66%)
Contusion with intact skin surface	144,579 (2.25%)	57,187 (2.75%)	114,157 (0.36%)
Nephritis, nephrotic syndrome, and nephrosis	1,160,628 (18.07%)	463,449 (22.26%)	1,186,582 (3.71%)
At least one claim (any type) with DME *and* other services *code at any point during the reference period*			
Hospital beds and associated supplies	185,040 (2.98%)	44,944 (2.16%)	24,128 (0.08%)
Wheelchairs, components, and accessories	333,428 (5.19%)	83,015 (3.99%)	71,197 (0.22%)

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
Walking aids and attachments	726,840 (11.32%)	95,015 (4.56%)	131,853 (0.41%)
Accessories for oxygen delivery devices	597,072 (9.30%)	167,483 (8.04%)	452,244 (1.42%)
Other supplies, including diabetes supplies and contraceptives	779,210 (12.13%)	231,277 (11.11%)	1,629,562 (5.10%)
Diabetic footwear	94,616 (1.47%)	28,490 (1.37%)	146,290 (0.46%)
Transportation services, including ambulance	2,640,341 (41.12%)	1,550,608 (74.47%)	1,044,639 (3.27%)

[a] Predictor codes are listed in Table A.3.
[b] Use a randomly selected stay out of those that have an eligible reference period.
[c] Six-month reference periods from beneficiaries with no HHA, SNF, or other type of PAC stay; randomly selected and matched on time only to a visit; no more than one reference period per beneficiary.

Table A.7. Frequency of Chronic Conditions Data Warehouse Candidate Predictors by Post-Acute Care Group

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
N =	6,421,516	2,082,220	31,942,637
Proximal: At least one inpatient claim with diagnosis within last 14 days of reference period			
ADHD/conduct disorders, hyperkinetic syndrome	1,152 (0.02%)	394 (0.02%)	220 (0.00%)
Alcohol use disorders	21,712 (0.34%)	24,258 (1.17%)	3,787 (0.01%)
Anxiety disorders	59,893 (0.93%)	23,364 (1.12%)	4,150 (0.01%)
Autism spectrum disorders	1,044 (0.02%)	484 (0.02%)	197 (0.00%)
Bipolar disorder	12,100 (0.19%)	7,109 (0.34%)	1,987 (0.01%)
Cerebral palsy	2,194 (0.03%)	1,373 (0.07%)	311 (0.00%)
Cystic fibrosis and other metabolic developmental disorders	3,483 (0.05%)	2,213 (0.11%)	317 (0.00%)
Depressive disorders	71,954 (1.12%)	36,345 (1.75%)	4,632 (0.01%)
Drug use disorders	21,260 (0.33%)	15,696 (0.75%)	3,523 (0.01%)
Epilepsy	27,502 (0.43%)	20,700 (0.99%)	2,254 (0.01%)
Fibromyalgia, chronic pain, and fatigue	40,148 (0.63%)	16,783 (0.81%)	2,769 (0.01%)
HIV/AIDS	5,628 (0.09%)	2,699 (0.13%)	693 (0.00%)
Intellectual disabilities and related conditions	4,074 (0.06%)	2,774 (0.13%)	905 (0.00%)
Learning disabilities	148 (0.00%)	87 (0.00%)	9 (0.00%)
Leukemias and lymphomas	34,158 (0.53%)	19,072 (0.92%)	3,620 (0.01%)
Liver disease, cirrhosis, and other liver conditions	62,182 (0.97%)	37,127 (1.78%)	8,492 (0.03%)
Migraine and chronic headache	8,417 (0.13%)	2,107 (0.10%)	776 (0.00%)
Mobility impairments	60,267 (0.94%)	64,909 (3.12%)	3,237 (0.01%)
Multiple sclerosis and transverse myelitis	8,686 (0.14%)	5,374 (0.26%)	388 (0.00%)
Muscular dystrophy	1,063 (0.02%)	666 (0.03%)	73 (0.00%)
Obesity	178,539 (2.78%)	78,069 (3.75%)	8,688 (0.03%)
Opioid disorder	4,829 (0.08%)	3,339 (0.16%)	911 (0.00%)

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
Other developmental delays	198 (0.00%)	139 (0.01%)	33 (0.00%)
Peripheral vascular disease	50,463 (0.79%)	26,725 (1.28%)	2,352 (0.01%)
Personality disorders	607 (0.01%)	403 (0.02%)	274 (0.00%)
Posttraumatic stress disorder	2,178 (0.03%)	626 (0.03%)	476 (0.00%)
Pressure and chronic ulcers	46,074 (0.72%)	60,277 (2.89%)	1,916 (0.01%)
Schizophrenia	6,485 (0.10%)	7,143 (0.34%)	1,674 (0.01%)
Schizophrenia and other psychotic disorders	8,936 (0.14%)	10,320 (0.50%)	2,054 (0.01%)
Sensory: blindness and visual impairment	1,523 (0.02%)	1,212 (0.06%)	82 (0.00%)
Sensory: deafness and hearing impairment	10,894 (0.17%)	5,326 (0.26%)	440 (0.00%)
Sickle cell anemia	1,143 (0.02%)	235 (0.01%)	612 (0.00%)
Spina bifida and other congenital anomalies of the nervous system	1,100 (0.02%)	476 (0.02%)	93 (0.00%)
Spinal cord injury	10,210 (0.16%)	17,264 (0.83%)	200 (0.00%)
Tobacco use	47,541 (0.74%)	14,863 (0.71%)	5,049 (0.02%)
Traumatic brain injury and nonpsychotic mental disorders due to brain damage	1,212 (0.02%)	1,137 (0.05%)	70 (0.00%)
Viral hepatitis	8,214 (0.13%)	3,644 (0.18%)	1,010 (0.00%)
Multiple prior: At least two claims on separate dates (any type) with diagnosis at any point during the reference period, excluding the last 14 days			
ADHD/conduct disorders, hyperkinetic syndrome	17,806 (0.28%)	8,261 (0.40%)	99,958 (0.31%)
Alcohol use disorders	87,823 (1.37%)	42,593 (2.05%)	147,798 (0.46%)
Anxiety disorders	457,649 (7.13%)	157,033 (7.54%)	953,474 (2.98%)
Autism spectrum disorders	3,414 (0.05%)	802 (0.04%)	28,128 (0.09%)
Bipolar disorder	125,059 (1.95%)	49,494 (2.38%)	344,352 (1.08%)
Cerebral palsy	16,193 (0.25%)	3,834 (0.18%)	35,648 (0.11%)
Cystic fibrosis and other metabolic developmental disorders	16,350 (0.25%)	5,969 (0.29%)	26,977 (0.08%)

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
Depressive disorders	557,864 (8.69%)	206,251 (9.91%)	952,947 (2.98%)
Drug use disorders	82,627 (1.29%)	28,447 (1.37%)	190,728 (0.60%)
Epilepsy	113,334 (1.76%)	46,270 (2.22%)	155,412 (0.49%)
Fibromyalgia, chronic pain, and fatigue	501,237 (7.81%)	153,695 (7.38%)	964,565 (3.02%)
HIV/AIDS	19,036 (0.30%)	5,819 (0.28%)	67,300 (0.21%)
Intellectual disabilities and related conditions	27,310 (0.43%)	9,580 (0.46%)	101,292 (0.32%)
Learning disabilities	4,459 (0.07%)	1,730 (0.08%)	6,072 (0.02%)
Leukemias and lymphomas	95,191 (1.48%)	31,826 (1.53%)	181,841 (0.57%)
Liver disease, cirrhosis, and other liver conditions	162,694 (2.53%)	61,346 (2.95%)	234,801 (0.74%)
Migraine and chronic headache	50,061 (0.78%)	11,623 (0.56%)	168,176 (0.53%)
Mobility impairments	226,047 (3.52%)	91,567 (4.40%)	85,387 (0.26%)
Multiple sclerosis and transverse myelitis	42,316 (0.66%)	11,218 (0.54%)	65,526 (0.21%)
Muscular dystrophy	3,709 (0.06%)	938 (0.05%)	4,188 (0.01%)
Obesity	375,622 (5.85%)	109,598 (5.26%)	696,160 (2.18%)
Opioid disorder	20,926 (0.33%)	7,642 (0.37%)	42,187 (0.13%)
Other developmental delays	2,475 (0.04%)	727 (0.035%)	5,497 (0.02%)
Peripheral vascular disease	607,955 (9.47%)	237,337 (11.40%)	691,037 (2.16%)
Personality disorders	29,919 (0.47%)	11,505 (0.55%)	90,733 (0.28%)
Posttraumatic stress disorder	25,707 (0.40%)	6,966 (0.33%)	112,961 (0.35%)
Pressure and chronic ulcers	358,840 (5.59%)	165,894 (7.97%)	157,221 (0.49%)
Schizophrenia	73,034 (1.14%)	33,861 (1.63%)	257,541 (0.81%)
Schizophrenia and other psychotic disorders	139,064 (2.17%)	71,013 (3.41%)	342,221 (1.07%)
Sensory: blindness and visual impairment	27,836 (0.43%)	9,431 (0.45%)	10,407 (0.03%)
Sensory: deafness and hearing impairment	68,909 (1.07%)	22,455 (1.08%)	174,515 (0.55%)
Sickle cell anemia	2,755 (0.04%)	512 (0.02%)	6,500 (0.02%)

Predictor,[a] *N* (%)	Home Health Agency[b]	Skilled Nursing Facility[b]	Non–Post-Acute Care[c]
Spina bifida and other congenital anomalies of the nervous system	6,777 (0.11%)	2,051 (0.10%)	8,539 (0.03%)
Spinal cord injury	41,811 (0.65%)	14,636 (0.70%)	11,278 (0.04%)
Tobacco use	182,617 (2.84%)	56,493 (2.71%)	373,723 (1.17%)
Traumatic brain injury and nonpsychotic mental disorders due to brain damage	17,530 (0.27%)	7,153 (0.34%)	16,381 (0.05%)
Viral hepatitis	49,405 (0.77%)	17,667 (0.85%)	108,915 (0.34%)

NOTE: ADHD = attention-deficit/hyperactivity disorder; HIV/AIDS = human immunodeficiency virus and acquired immunodeficiency syndrome.
[a] Predictor codes available at CCW, undated.
[b] Use a randomly selected stay out of those that have an eligible reference period.
[c] Six-month reference periods from beneficiaries with no HHA, SNF, or other type of PAC stay; randomly selected and matched on time only to a visit; no more than one reference period per beneficiary.

Table A.8. Preliminary Harmonized Items and Corresponding MDS and OASIS Items and Response Levels

Proposed Harmonized Item and Response Levels	OASIS (Home Health Agency) Response Levels (Item Numbers)	MDS (Skilled Nursing Facility) Response Levels (Item Numbers)
Functional limitations		
1. Grooming or personal hygiene	**M1800. Grooming**: Current ability to tend safely to personal hygiene needs (specifically: washing face and hands, hair care, shaving or make up, teeth or denture care, or fingernail care).	**G0110J. Personal hygiene**: how resident maintains personal hygiene, including combing hair, brushing teeth, shaving, applying makeup, washing/drying face and hands (excludes baths and showers).
Independent	0. Able to groom self unaided, with or without the use of assistive devices or adapted methods.	0. Independent: No help or staff oversight
Partial dependence	1. Grooming utensils must be placed within reach before able to complete grooming activities. **OR** 2. Someone must assist the patient to groom self.	1. Supervision: oversight, encouragement or cueing. **OR** 2. Limited assistance: resident highly involved in activity; staff provide guided maneuvering of limbs or other non–weight-bearing assistance. **OR** 3. Extensive assistance: resident involved in activity, staff provide weight-bearing support.
Total dependence	3. Patient depends entirely upon someone else for grooming needs.	4. Total dependence: full staff performance every time during entire 7-day period.
2. Dressing	**M1810. Current Ability to Dress Upper Body** safely (with or without dressing aids), including undergarments, pullovers, front-opening shirts and blouses, managing zippers, buttons, and snaps. **M1820. Current Ability to Dress Lower Body** safely (with or without dressing aids), including undergarments, slacks, socks or nylons, and shoes.	**G0110G. Dressing**: how resident puts on, fastens, and takes off all items of clothing, including donning/removing a prosthesis or TED hose. Dressing includes putting on and changing pajamas and housedresses.
Independent	0. Able to get clothes out of closets and drawers, put them on, and remove them from the upper body without assistance. (M1810) **AND** 0. Able to obtain, put on, and remove clothing and shoes without assistance. (M1820)	0. Independent: no help or staff oversight
Partial dependence	1. Able to dress upper body without assistance if clothing is laid out or handed to the patient. (M1810) **OR** 2. Someone must help the patient put on upper body clothing (M1810). **OR** 1. Able to dress lower body without assistance if clothing and	1. Supervision: oversight, encouragement or cueing. **OR** 2. Limited assistance: resident highly involved in activity; staff provide guided maneuvering of limbs or other non–weight-bearing assistance. **OR**

Proposed Harmonized Item and Response Levels	OASIS (Home Health Agency) Response Levels (Item Numbers)	MDS (Skilled Nursing Facility) Response Levels (Item Numbers)
	shoes are laid out or handed to the patient. (M1820) **OR** 2. Someone must help the patient put on undergarments, slacks, socks or nylons, and shoes. (M1820)	3. Extensive assistance: resident involved in activity, staff provide weight-bearing support.
Total dependence	3. Patient depends entirely upon another person to dress the upper body. (M1810) **AND** 3. Patient depends entirely upon another person to dress lower body. (M1820)	4. Total dependence: full staff performance every time during entire 7-day period.
3. Toilet use	**M1840. Toilet Transferring**: Current ability to get to and from the toilet or bedside commode safely and transfer on and off toilet/commode. **M1845. Toileting Hygiene**: Current ability to maintain perineal hygiene safely, adjust clothes and/or incontinence pads before and after using toilet, commode, bedpan, urinal; if managing ostomy, includes cleaning area around stoma, but not managing equipment.	**G0110I. Toilet use**: how resident uses the toilet room, commode, bedpan, or urinal; transfers on/off toilet; cleanses self after elimination; changes pad; manages ostomy or catheter; and adjusts clothes. Do not include emptying of bedpan, urinal, bedside commode, catheter bag or ostomy bag.
Independent	0. Able to get to and from the toilet and transfer independently with or without a device. (M1840) **AND** 0. Able to manage toileting hygiene and clothing management without assistance. (M1845)	0. Independent: no help or staff oversight
Partial dependence	1. When reminded, assisted, or supervised by another person, able to get to and from the toilet and transfer. (M1840) **OR** 2. Unable to get to and from the toilet but is able to use a bedside commode (with or without assistance). (M1840) **OR** 3. Unable to get to and from the toilet or bedside commode but is able to use a bedpan/urinal independently (M1840) **OR** 1. Able to manage toileting hygiene and clothing management without assistance if supplies/implements are laid out for the patient. (M1845) **OR** 2. Someone must help the patient to maintain toileting hygiene and/or adjust clothing. (M1850)	1. Supervision: oversight, encouragement or cueing. **OR** 2. Limited assistance: resident highly involved in activity; staff provide guided maneuvering of limbs or other non–weight-bearing assistance. **OR** 3. Extensive assistance: resident involved in activity, staff provide weight-bearing support.
Total dependence	4. Is totally dependent in toileting. (M1840) **AND** 3. Patient depends entirely on another person to maintain toileting hygiene. (M1845)	4. Total dependence: full staff performance every time during entire 7-day period.
4. Transferring	**M1850. Transferring**: Current ability to move safely from bed to chair, or ability to turn and position self in bed if patient is bedfast.	**G0110B. Transfer**: how resident moves between surfaces including to or from: bed, chair, wheelchair, standing position (**excludes** to/from bath/toilet).
Independent	0. Able to independently transfer.	0. Independent: no help or staff oversight

Proposed Harmonized Item and Response Levels	OASIS (Home Health Agency) Response Levels (Item Numbers)	MDS (Skilled Nursing Facility) Response Levels (Item Numbers)
Partial dependence	1. Able to transfer with minimal human assistance or with use of an assistive device. **OR** 2. Able to bear weight and pivot during the transfer process but unable to transfer self.	1. Supervision: oversight, encouragement or cueing. **OR** 2. Limited assistance: resident highly involved in activity; staff provide guided maneuvering of limbs or other non–weight-bearing assistance. **OR** 3. Extensive assistance: resident involved in activity, staff provide weight-bearing support.
Total dependence	3. Unable to transfer self and is unable to bear weight or pivot when transferred by another person. **OR** 4. Bedfast, unable to transfer but is able to turn and position self in bed. **OR** 5. Bedfast, unable to transfer and is unable to turn and position self.	4. Total dependence: full staff performance every time during entire 7-day period.
5. Feeding or eating	**M1870. Feeding or Eating**: Current ability to feed self meals and snacks safely. Note: This refers only to the process of eating, chewing, and swallowing, not preparing the food to be eaten.	**G0110H. Eating**: how resident eats and drinks, regardless of skill. Do not include eating/drinking during medication pass. Includes intake of nourishment by other means (e.g., tube feeding, total parenteral nutrition, IV fluids administered for nutrition or hydration).
Independent	0. Able to independently feed self.	0. Independent: no help or staff oversight
Partial dependence	1. Able to feed self independently but requires: (a) meal set-up; **OR** (b) intermittent assistance or supervision from another person; **OR** (c) a liquid, pureed or ground meat diet.	1. Supervision: oversight, encouragement or cueing. **OR** 2. Limited assistance: resident highly involved in activity; staff provide guided maneuvering of limbs or other non–weight-bearing assistance. **OR** 3. Extensive assistance: resident involved in activity, staff provide weight-bearing support.
Total dependence	2. Unable to feed self and must be assisted or supervised throughout the meal/snack. **OR** 3. Able to take in nutrients orally and receives supplemental nutrients through a nasogastric tube or gastrostomy. **OR** 4. Unable to take in nutrients orally and is fed nutrients through a nasogastric tube or gastrostomy. **OR** 5. Unable to take in nutrients orally or by tube feeding.	4. Total dependence: full staff performance every time during entire 7-day period.
Mobility limitations		
6. Mobility	**M1860. Ambulation/Locomotion**: Current ability to walk safely, once in a standing position, or use a wheelchair, once in a seated position, on a variety of surfaces.	**G0110E. Locomotion on unit**: how resident moves between locations in his/her room and adjacent corridor on same floor. If in wheelchair, self-sufficiency once in chair. **G0110F. Locomotion off unit**: how resident moves to and returns from off-unit locations (e.g., areas set aside for dining,

Proposed Harmonized Item and Response Levels	OASIS (Home Health Agency) Response Levels (Item Numbers)	MDS (Skilled Nursing Facility) Response Levels (Item Numbers)
		activities, or treatments). **If facility has only one floor**, how resident moves to and from distant areas on the floor. If in wheelchair, self-sufficiency once in chair.
Walk without assistance or walk with assistance or device or can wheel self independently once in wheelchair	0. Able to independently walk on even and uneven surfaces and negotiate stairs with or without railings (specifically: needs no human assistance or assistive device). **OR** 1. With the use of a one-handed device (for example, cane, single crutch, hemi-walker), able to independently walk on even and uneven surfaces and negotiate stairs with or without railings. **OR** 2. Requires use of a two-handed device (for example, walker or crutches) to walk alone on a level surface and/or requires human supervision or assistance to negotiate stairs or steps or uneven surfaces. **OR** 3. Able to walk only with the supervision or assistance of another person at all times. **OR** 4. Chairfast, unable to ambulate but is able to wheel self independently.	0. Independent: no help or staff oversight **OR** 1. Supervision: oversight, encouragement, or cueing. **OR** 2. Limited assistance: resident highly involved in activity; staff provide guided maneuvering of limbs or other non–weight-bearing assistance. **OR** 3. Extensive assistance: resident involved in activity, staff provide weight-bearing support. **[For either G0110E or G0110F]**
Unable to walk with assistance and unable to wheel self in chair	5. Chairfast, unable to ambulate and is unable to wheel self. **OR** 6. Bedfast, unable to ambulate or be up in a chair.	4. Total dependence: full staff performance every time during entire 7-day period. **[For both G0110E and G0110F]**
Cognitive limitations		
7. Memory and recall	**M1740.** Cognitive, behavioral, and psychiatric symptoms that are demonstrated at least once a week (Reported or Observed): (Mark all that apply.) 1. Memory deficit: failure to recognize familiar persons/places, inability to recall events of past 24 hours, significant memory loss so that supervision is required	**C0500.** BIMS Summary Score **C0700.** Short-term Memory OK **C0800.** Long-term Memory OK
No memory deficits	Item not checked	BIMS Score ≥ 13 or 0. Memory OK. (For both items)
Some degree memory deficit	Item checked	BIMS Score < 13 or 1. Memory problem. (For at least one item)

SOURCE: Assessment items and response levels are from CMS, 2019a, and CMS, 2019b.
NOTE: BIMS = brief interview for mental status; OK = okay; TED = thromboembolism-deterrent.

Table A.9. Predicted and Actual Factor Score Decile, Weighted Frequency and Row Percentage

Predicted Factor Score Decile	Actual Factor Score Decile									
	1	2	3	4	5	6	7	8	9	10
1	143,265 (15.68%)	124,526 (13.63%)	108,457 (11.87%)	107,705 (11.79%)	95,447 (10.45%)	94,144 (10.31%)	79,537 (8.71%)	77,432 (8.48%)	50,676 (5.55%)	32,287 (3.53%)
2	117,772 (12.89%)	112,846 (12.35%)	101,968 (11.16%)	105,807 (11.58%)	95,000 (10.40%)	96,945 (10.61%)	87,893 (9.62%)	87,128 (9.54%)	59,281 (6.49%)	48,857 (5.35%)
3	111,756 (12.23%)	110,467 (12.09%)	98,526 (10.79%)	104,808 (11.47%)	98,529 (10.79%)	91,785 (10.05%)	84,584 (9.26%)	84,007 (9.20%)	71,583 (7.84%)	57,444 (6.29%)
4	107,533 (11.77%)	107,408 (11.76%)	98,833 (10.82%)	103,878 (11.37%)	96,628 (10.58%)	94,203 (10.31%)	89,989 (9.85%)	86,978 (9.52%)	64,813 (7.10%)	63,223 (6.92%)
5	101,803 (11.14%)	102,365 (11.21%)	98,658 (10.80%)	97,935 (10.72%)	96,780 (10.59%)	94,416 (10.34%)	87,553 (9.58%)	90,721 (9.93%)	73,726 (8.07%)	69,512 (7.61%)
6	92,984 (10.18%)	97,635 (10.69%)	94,925 (10.39%)	95,056 (10.41%)	95,725 (10.48%)	94,919 (10.39%)	92,486 (10.12%)	98,818 (10.82%)	78,753 (8.62%)	72,196 (7.90%)
7	83,500 (9.14%)	87,741 (9.60%)	95,390 (10.44%)	91,642 (10.03%)	93,652 (10.25%)	94,481 (10.34%)	96,675 (10.58%)	103,831 (11.37%)	84,342 (9.23%)	82,240 (9.00%)
8	72,504 (7.94%)	74,374 (8.14%)	92,348 (10.11%)	85,397 (9.35%)	93,060 (10.19%)	91,252 (9.99%)	99,097 (10.85%)	98,138 (10.74%)	111,717 (12.23%)	95,598 (10.47%)
9	55,557 (6.08%)	61,968 (6.78%)	78,161 (8.56%)	73,701 (8.07%)	86,531 (9.47%)	89,506 (9.80%)	102,998 (11.28%)	92,604 (10.14%)	145,416 (15.92%)	127,048 (13.91%)
10	26,812 (2.94%)	34,142 (3.74%)	46,231 (5.06%)	47,547 (5.21%)	62,152 (6.80%)	71,799 (7.86%)	92,712 (10.15%)	93,831 (10.27%)	173,180 (18.96%)	265,084 (29.02%)

NOTE: Gray shading indicates weighted number of beneficiaries with accurately predicted number of limitations.

Table A.10. Memory Limitation Model Specifications, Twelve-Month Reference Periods

Parameter	Estimate	Parameter	Estimate
Intercept	–1.104697	Viral hepatitis, multiple prior	–0.002948
Age category 70–75	–1.717512	Diabetes mellitus complications, proximal	–0.002400
Age category 60–65	–1.626781	Fibromyalgia, multiple prior pain and fatigue, proximal	–0.001360
Age category 50–55	–1.583455	Arthritis (CCW), proximal	–0.001166
Age category 75–80	–1.571454	Opioid disorder, multiple prior	–0.001080
Age category 55–60	–1.520645	Wheelchairs, components, and accessories	–0.000804
Age category 80–85	–1.463534	Vertigo, proximal	–0.000690
Age category 85–90	–1.456278	Pneumonia and influenza, multiple prior	–0.000509
Age category 45–50	–1.442771	Leukemias and lymphomas, proximal	–0.000201
Age category 0–25	–1.326818	Hereditary and degenerative diseases of the central nervous system, proximal	–0.000032
Age category 65–70	–1.309189	Posttraumatic stress disorder, multiple prior	–0.000011
Age category 35–40	–1.132438	Heart failure (CCW), multiple prior	–0.000009
Age category 40–45	–0.803344	Muscular dystrophy, multiple prior	–0.000005
Age category 25–30	–0.496284	Alcohol use disorders, proximal	0.000003
Walking aids and attachments	–0.372360	Schizophrenia, proximal	0.000003
Arthropathies and related disorders, proximal	–0.301112	Contusion with intact skin surface, multiple prior	0.000022
Cancer screening, multiple prior	–0.298597	Epilepsy, proximal	0.000319
Arthritis, proximal	–0.296398	Other psychoses, multiple prior	0.000466
Obesity, multiple prior	–0.292370	Schizophrenia and other psychotic disorders, proximal	0.001543
Age category 30–35	–0.267544	Learning disabilities, multiple prior	0.002409
Fibromyalgia, multiple prior pain and fatigue, multiple prior	–0.228473	Age in years	0.003854
Rehabilitative services, multiple prior	–0.218942	Autism spectrum disorders, multiple prior	0.005393
Arthritis, multiple prior	–0.212264	Alcohol use disorders, multiple prior	0.005814
Lipid abnormality, proximal	–0.205698	Traumatic brain injury and nonpsychotic mental disorders due to brain damage, proximal	0.006449
Drug use disorders, multiple prior	–0.185352	Other developmental delays, multiple prior	0.006890
Vertigo, multiple prior	–0.176112	Intellectual disabilities and related conditions, proximal	0.008714
Chronic obstructive pulmonary disease and allied conditions, proximal	–0.165379	Difficulty walking, multiple prior	0.009372
Lipid abnormality, multiple prior	–0.162865	Diabetes mellitus complications, multiple prior	0.016008
Arthritis (CCW), multiple prior	–0.152518	Pneumonia and influenza, proximal	0.017754
Ischemic heart disease, multiple prior	–0.148943	Peripheral vascular disease, multiple prior	0.018188

Parameter	Estimate	Parameter	Estimate
Arthropathies and related disorders, multiple prior	–0.146025	Transportation services, including ambulance	0.018428
Other forms of heart disease, multiple prior	–0.145463	Paralysis, proximal	0.019440
Hypertension, proximal	–0.143903	Mobility impairments, proximal	0.019837
Diseases of veins and lymphatics and other diseases of circulatory system, multiple prior	–0.136153	Mobility impairments, multiple prior	0.023020
Leukemias and lymphomas, multiple prior	–0.128964	Parkinson's disease, multiple prior	0.025675
Chronic obstructive pulmonary disease and allied conditions, multiple prior	–0.124641	Paralysis, multiple prior	0.026364
Hypertensive disease, proximal	–0.121418	Diseases of other endocrine glands, multiple prior	0.028473
Hypertensive disease, multiple prior	–0.118109	Wheelchair	0.028500
Sensory: deafness and hearing impairment, multiple prior	–0.116603	Chronic kidney disease, proximal	0.033264
Ischemic heart disease, proximal	–0.116355	Traumatic brain injury and nonpsychotic mental disorders due to brain damage, multiple prior	0.035018
Other forms of heart disease, proximal	–0.100636	Heart failure, proximal	0.039712
Sepsis, multiple prior	–0.096968	Ill-defined and unknown causes of morbidity and mortality, proximal	0.040735
Anxiety disorders, multiple prior	–0.096090	ADHD/conduct disorders, hyperkinetic syndrome, multiple prior	0.042022
Viral hepatitis, proximal	–0.094473	Heart failure (CCW), proximal	0.043012
Hypertension, multiple prior	–0.088645	Home hospital bed	0.046599
Asthma, multiple prior	–0.087089	Low income subsidy	0.047710
Obesity, proximal	–0.085678	Heart failure, multiple prior	0.048550
Chronic kidney disease, multiple prior	–0.079266	Home hospital bed	0.053938
Home oxygen	–0.072275	Other psychoses, proximal	0.055919
Other supplies, including diabetes supplies and contraceptives	–0.071371	Schizophrenia, multiple prior	0.059081
Diseases of other endocrine glands, proximal	–0.069445	Nephritis, nephrotic syndrome, and nephrosis, proximal	0.071454
Depression, multiple prior	–0.060871	Neurotic disorders, personality disorders, and other nonpsychotic mental disorders, proximal	0.076355
Migraine and chronic headache, multiple prior	–0.055319	Pressure and chronic ulcers, proximal	0.077975
Accessories for oxygen delivery devices	–0.048745	Cerebrovascular disease, multiple prior	0.080626
Sepsis, proximal	–0.042085	Podiatric care, multiple prior	0.092340
Diseases of veins and lymphatics, and other diseases of circulatory system, proximal	–0.038570	Cerebrovascular disease, proximal	0.093751

Parameter	Estimate	Parameter	Estimate
Open wound of lower limb, multiple prior	–0.036774	Psychiatric, multiple prior	0.149410
HIV/AIDS, multiple prior	–0.034571	Other diseases of urinary system, proximal	0.159562
Diabetic footwear	–0.033197	Cerebral palsy, multiple prior	0.166498
Sex	–0.033044	Ill-defined and unknown causes of morbidity and mortality, multiple prior	0.193895
Liver disease, cirrhosis, and other Liver conditions, multiple prior	–0.032511	Hereditary and degenerative diseases of the central nervous system, multiple prior	0.209622
Weakness, multiple prior	–0.028613	Organic psychotic conditions, proximal	0.248108
Spinal cord injury, multiple prior	–0.026618	Alzheimer's, proximal	0.248472
Other diseases of urinary system, multiple prior	–0.016297	Ambulance/life support	0.256184
Multiple sclerosis and transverse myelitis	–0.015746	Dually eligible beneficiary	0.263448
Anxiety disorders, proximal	–0.014591	Schizophrenia and other psychotic disorders, multiple prior	0.287693
Personality disorders, multiple prior	–0.011976	Stroke/brain injury, multiple prior	0.318878
Psychiatric, proximal	–0.011948	Epilepsy, multiple prior	0.411572
Tobacco use, multiple prior	–0.010387	Organic psychotic conditions, multiple prior	0.459877
Depressive disorders, proximal	–0.008678	Alzheimer's, multiple prior	0.519832
Depression, proximal	–0.008437	Stroke/brain injury, proximal	0.754288
Pressure and chronic ulcers, multiple prior	–0.007633	Dementia, proximal	0.864878
Asthma, proximal	–0.006166	Intellectual disabilities and related conditions, multiple prior	1.169462
Peripheral vascular disease, proximal	–0.006144	Dementia, multiple prior	1.379501
Nephritis, nephrotic syndrome, and nephrosis, multiple prior	–0.003262		

NOTE: ADHD = attention-deficit/hyperactivity disorder; HIV/AIDS = human immunodeficiency virus and acquired immunodeficiency syndrome.

Table A.11. Activity and Mobility Limitations Model Specifications, Twelve-Month Reference Periods

Parameter	Estimate	Parameter	Estimate
Intercept	4.544280	Spina bifida and other congenital anomalies of the nervous system, multiple prior	0.000010
Age category 70–75	–0.745284	Diseases of veins and lymphatics, and other diseases of circulatory system, proximal	0.000020
Age category 75–80	–0.735389	Spinal cord injury, multiple prior	0.000850
Age category 60–65	–0.657020	Ill-defined and unknown causes of morbidity and mortality, proximal	0.001010
Age category 65–70	–0.641752	Autism spectrum disorders, multiple prior	0.001490
Age category 55–60	–0.560551	Fibromyalgia, multiple prior pain and fatigue, multiple prior	0.001550
Age category 50–55	–0.498982	Muscular dystrophy, multiple prior	0.001620
Age category 80–85	–0.442639	Obesity, proximal	0.001730
Age category 85–90	–0.379139	Walking aids and attachments	0.002690
Age category 45–50	–0.322112	Female sex (reference = unknown)	0.004528
Age category 0–25	–0.311723	Psychiatric, multiple prior	0.005030
Age category 35–40	–0.256533	Hypertension, proximal	0.006230
Age category 40–45	–0.208447	Chronic kidney disease, multiple prior	0.007400
Neurotic disorders, personality disorders, and other nonpsychotic mental disorders, proximal	–0.177700	Ischemic heart disease, multiple prior	0.010070
Schizophrenia, multiple prior	–0.174470	Diseases of other endocrine glands, multiple prior	0.010230
Age category 25–30	–0.166189	Organic psychotic conditions, proximal	0.010440
Age category 30–35	–0.125304	Nephritis, nephrotic syndrome, and nephrosis, proximal	0.010490
Bipolar disorder, multiple prior	–0.085610	Accessories for oxygen delivery devices	0.010810
Tobacco use, multiple prior	–0.085210	Wheelchair	0.011710
Drug use disorders, multiple prior	–0.072760	Sensory: blindness and visual impairment, multiple prior	0.012460
Cancer screening, multiple prior	–0.067570	Pneumonia and influenza, proximal	0.012610
Schizophrenia and other psychotic disorders, multiple prior	–0.059870	Other diseases of urinary system, multiple prior	0.012880
Other supplies, including diabetes supplies and contraceptives	–0.059440	Asthma, multiple prior	0.012900
Arthritis (CCW), multiple prior	–0.053950	Arthritis, multiple prior	0.014120
Lipid abnormality, multiple prior	–0.051270	Depressive disorders, multiple prior	0.016100
Other psychoses, proximal	–0.042930	Hereditary and degenerative diseases of the central nervous system, proximal	0.017670

Parameter	Estimate	Parameter	Estimate
Posttraumatic stress disorder, multiple prior	–0.041910	Chronic kidney disease, proximal	0.018590
Vertigo, multiple prior	–0.041140	Low income subsidy	0.020750
Alcohol use disorders, multiple prior	–0.039360	Arthropathies and related disorders, multiple prior	0.022350
Rehabilitative services, multiple prior	–0.034200	Cerebrovascular disease, multiple prior	0.022850
Other forms of heart disease, multiple prior	–0.031100	Pressure and chronic ulcers, proximal	0.026350
Sepsis, proximal	–0.029790	Peripheral vascular disease, multiple prior	0.026390
Other bacterial diseases, multiple prior	–0.024680	Alzheimer's, proximal	0.030940
Sensory: deafness and hearing impairment, multiple prior	–0.024240	Mobility impairments, proximal	0.034460
Hypertensive disease, multiple prior	–0.023830	Paralysis, proximal	0.034520
Leukemias and lymphomas, multiple prior	–0.022550	Pneumonia and influenza, multiple prior	0.037520
Sepsis, multiple prior	–0.021160	Other diseases of urinary system, proximal	0.039430
Open wound of lower limb, multiple prior	–0.020980	Ambulance/life support	0.039510
Migraine and chronic headache, multiple prior	–0.016480	Arthritis, proximal	0.048370
Other bacterial diseases, proximal	–0.015060	Arthropathies and related disorders, proximal	0.053030
Cystic fibrosis and other metabolic developmental disorders, proximal	–0.013710	Diabetes mellitus complications, multiple prior	0.054830
Diabetic footwear	–0.012650	Organic psychotic conditions, multiple prior	0.057000
Male sex (reference = unknown)	–0.010822	Heart failure (CCW), multiple prior	0.060140
Viral hepatitis, multiple prior	–0.010500	Difficulty walking, multiple prior	0.062430
Hypertension, multiple prior	–0.009650	Pressure and chronic ulcers, multiple prior	0.067220
Contusion with intact skin surface, multiple prior	–0.008180	Stroke/brain injury, multiple prior	0.072760
Heart failure, proximal	–0.006780	Multiple sclerosis and transverse myelitis	0.078860
Personality disorders, multiple prior	–0.006430	Dually eligible beneficiary	0.082240
Diseases of veins and lymphatics, and other diseases of circulatory system, multiple prior	–0.006060	Weakness, multiple prior	0.084600
Other forms of heart disease, proximal	–0.005670	Hereditary and degenerative diseases of the central nervous system, multiple prior	0.085100
Lipid abnormality, proximal	–0.004290	Transportation services, including ambulance	0.087630

Parameter	Estimate	Parameter	Estimate
Chronic obstructive pulmonary disease and allied conditions, multiple prior	–0.004020	Dementia, proximal	0.087750
Anxiety disorders, multiple prior	–0.003800	Dementia, multiple prior	0.106050
Ischemic heart disease, proximal	–0.003440	Parkinson's disease, multiple prior	0.121350
HIV/AIDS, multiple prior	–0.002720	Home hospital bed	0.129720
Age in years	–0.000820	Stroke/brain injury, proximal	0.130380
Nephritis, nephrotic syndrome, and nephrosis, multiple prior	–0.000240	Epilepsy, multiple prior	0.136990
Chronic obstructive pulmonary disease and allied conditions, proximal	–0.000010	Alzheimer's, multiple prior	0.155050
Neurotic disorders, personality disorders, and other nonpsychotic mental disorders, multiple prior	–0.000010	Paralysis, multiple prior	0.173690
Arthritis (CCW), proximal	0.000000	Home hospital bed	0.174240
Obesity, multiple prior	0.000000	Mobility impairments, multiple prior	0.179440
Other developmental delays, multiple prior	0.000000	Wheelchairs, components, and accessories	0.212120
Psychiatric, proximal	0.000000	Intellectual disabilities and related conditions, multiple prior	0.264450
Spinal cord injury, proximal	0.000000	Cerebral palsy, multiple prior	0.287320
Podiatric care, multiple prior	0.000010		

NOTE: HIV/AIDS = human immunodeficiency virus and acquired immunodeficiency syndrome.

Table A.12. Factor Score Model Specifications, Twelve-Month Reference Periods

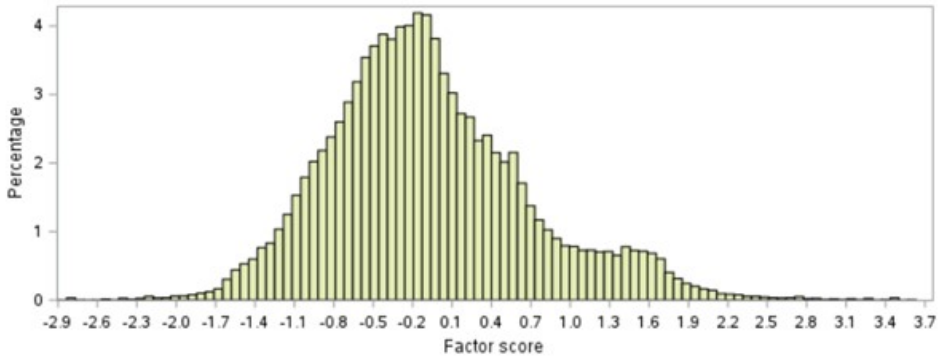
Parameter	Estimate	Parameter	Estimate
Age category 75–80	–0.568699	Podiatric care, multiple prior	–0.000040
Age category 60–65	–0.520204	ADHD/conduct disorders, hyperkinetic syndrome, multiple prior	0.000000
Age category 70–75	–0.493274	Arthritis (CCW), proximal	0.000000
Age category 55–60	–0.492615	Chronic kidney disease, multiple prior	0.000000
Age category 50–55	–0.472488	Intellectual disabilities and related conditions, proximal	0.000000
Age category 80–85	–0.440926	Ill-defined and unknown causes of morbidity and mortality, multiple prior	0.000010
Age category 85–90	–0.402375	Traumatic brain injury and nonpsychotic mental disorders because of brain damage, multiple prior	0.000120
Age category 65–70	–0.380961	Other developmental delays, multiple prior	0.000220
Age category 45–50	–0.379313	Heart failure, multiple prior	0.000340
Age category 0–25	–0.348123	Psychiatric, multiple prior	0.000440
Age category 35–40	–0.316787	Parkinson's disease, proximal	0.000460
Age category 40–45	–0.267942	Asthma, multiple prior	0.001020
Age category 25–30	–0.216096	Diseases of veins and lymphatics, and other diseases of circulatory system, proximal	0.002310
Schizophrenia, multiple prior	–0.189970	Spina bifida and other congenital anomalies of the nervous system, multiple prior	0.003130
Age category 30–35	–0.161848	Hypertension, proximal	0.004060
Neurotic disorders, personality disorders, and other nonpsychotic mental disorders, proximal	–0.153430	Spinal cord injury, proximal	0.004220
Tobacco use, multiple prior	–0.102040	Diseases of other endocrine glands, multiple prior	0.005030
Vertigo, multiple prior	–0.101740	Obesity, multiple prior	0.005360
Drug use disorders, multiple prior	–0.097860	Ill-defined and unknown causes of morbidity and mortality, proximal	0.005710
Bipolar disorder, multiple prior	–0.086380	Autism spectrum disorders, multiple prior	0.006240
Lipid abnormality, multiple prior	–0.073830	Pneumonia and influenza, proximal	0.007000
Arthritis (CCW), multiple prior	–0.072210	Arthritis, multiple prior	0.007970
Alcohol use disorders, multiple prior	–0.069000	Female sex (reference = unknown)	0.008921
Cancer screening, multiple prior	–0.065250	Muscular dystrophy, multiple prior	0.009140
Hypertensive disease, multiple prior	–0.063500	Spinal cord injury, multiple prior	0.010930
Schizophrenia and other psychotic disorders, multiple prior	–0.054660	Sensory: blindness and visual impairment, multiple prior	0.011990

Parameter	Estimate	Parameter	Estimate
Sensory: deafness and hearing impairment, multiple prior	–0.047680	Obesity, proximal	0.019100
Migraine and chronic headache, multiple prior	–0.045260	Cerebrovascular disease, multiple prior	0.020710
Other supplies, including diabetes supplies and contraceptives	–0.043800	Peripheral vascular disease, multiple prior	0.021660
Other psychoses, proximal	–0.041090	Nephritis, nephrotic syndrome, and nephrosis, proximal	0.024270
Walking aids and attachments	–0.039190	Organic psychotic conditions, proximal	0.026270
Other forms of heart disease, multiple prior	–0.037040	Accessories for oxygen delivery devices	0.029180
Viral hepatitis, multiple prior	–0.036410	Chronic kidney disease, proximal	0.030560
Diabetic footwear	–0.036070	Diabetes mellitus complications, multiple prior	0.034370
Leukemias and lymphomas, multiple prior	–0.035630	Difficulty walking, multiple prior	0.034530
Posttraumatic stress disorder, multiple prior	–0.029740	Other diseases of urinary system, multiple prior	0.034960
Open wound of lower limb, multiple prior	–0.027660	Alzheimer's, proximal	0.046450
Male sex (reference = unknown)	–0.023860	Hereditary and degenerative diseases of the central nervous system, proximal	0.047550
Anxiety disorders, multiple prior	–0.019650	Arthritis, proximal	0.048600
Chronic obstructive pulmonary disease and allied conditions, multiple prior	–0.019300	Pneumonia and influenza, multiple prior	0.049680
Rehabilitative services, multiple prior	–0.019220	Dually eligible beneficiary	0.052360
Chronic obstructive pulmonary disease and allied conditions, proximal	–0.017840	Arthropathies and related disorders, proximal	0.052890
HIV/AIDS, multiple prior	–0.016560	Heart failure (CCW), multiple prior	0.057720
Sepsis, proximal	–0.012240	Paralysis, proximal	0.078190
Cystic fibrosis and other metabolic developmental disorders, proximal	–0.011800	Mobility impairments, proximal	0.078350
Diseases of veins and lymphatics, and other diseases of circulatory system, multiple prior	–0.010890	Stroke/brain injury, multiple prior	0.080580
Fibromyalgia, multiple prior pain and fatigue, multiple prior	–0.010860	Weakness, multiple prior	0.085640
Hypertension, multiple prior	–0.010070	Transportation services including ambulance	0.087910
Lipid abnormality, proximal	–0.009680	Pressure and chronic ulcers, proximal	0.095950
Other bacterial diseases, multiple prior	–0.008410	Other diseases of urinary system, proximal	0.099720
Ischemic heart disease, proximal	–0.007610	Organic psychotic conditions, multiple prior	0.102670
Low income subsidy	–0.007300	Ambulance/life support	0.109420

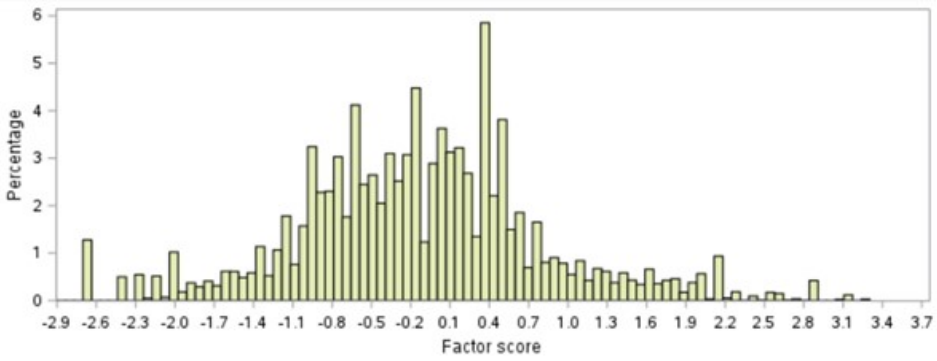
NOTE: ADHD = attention-deficit/hyperactivity disorder; HIV/AIDS = human immunodeficiency virus and acquired immunodeficiency syndrome.

Figure A.1. Factor Score by Post-Acute Care Assessment Group

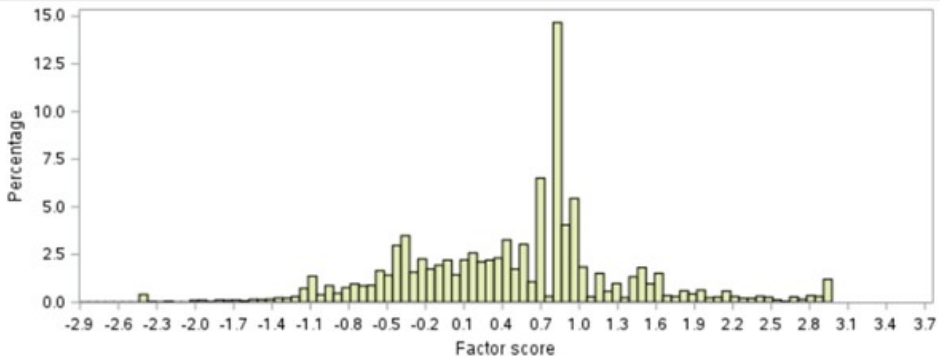
Beneficiaries with both assessments (*N* = 610,332 beneficiaries)



Home Health Agency (*N* = 6,378,403)



Skilled Nursing Facility (*N* = 1,833,730)



Abbreviations

ADI	area deprivation index
ADL	activities of daily living
ASPE	Assistant Secretary for Planning and Evaluation
AUC	area under the curve
CCW	Chronic Conditions Data Warehouse
CFI	claims-based frailty index
CMS	Centers for Medicare and Medicaid Services
DME	durable medical equipment
FFS	fee-for-service
HCPCS	Healthcare Common Procedure Coding System
HHA	home health agency
ICD-9	International Classification of Diseases, version 9
ICD-10	International Classification of Diseases, version 10
MCBS	Medicare Current Beneficiary Survey
MDS	Minimum Data Set
OASIS	Outcome and Assessment Information Set
PAC	post-acute care
PATF	project advisory task force
PPV	positive predictive value
RMSE	root mean squared error
RMSEA	root mean squared error of approximation
SNF	skilled nursing facility
TLI	Tucker-Lewis index

References

CCW—*See* Chronic Conditions Data Warehouse.

Centers for Medicare and Medicaid Services, *OASIS-D—All Item Set: Home Health Patient Tracking Sheet*, assessment, Baltimore, Md., January 1, 2019a. As of February 23, 2022: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Downloads/OASIS-D_All-Items_final.pdf

———, *Minimum Data Set (MDS) 3.0 for Nursing Homes and Swing Bed Providers*, assessment, Baltimore, Md., October 1, 2019b. As of February 23, 2022: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Other-Content-Types/MDS30_Final_Item_Sets_v1171_for-Oct_1_2019.zip

Chalmers, Phil, Joshua Pritikin, Alexander Robitzsch, and Mateusz Zoltak, *Package "Mirt,"* Comprehensive R Archive Network, Version 1.35.1, 2021. As of March 9, 2021: https://cran.r-project.org/web/packages/mirt/mirt.pdf

Chronic Conditions Data Warehouse, "Condition Categories," webpage, undated. As of March 9, 2021: https://www2.ccwdata.org/web/guest/condition-categories

CMS—*See* Centers for Medicare and Medicaid Services.

Ensrud, Kristine E., Susan K. Ewing, Peggy M. Cawthon, Howard A. Fink, Brent C. Taylor, Jane A. Cauley, Thuy-Tien Dam, Lynn M. Marshall, Eric S. Orwoll, and Steven R. Cummings, "A Comparison of Frailty Indexes for the Prediction of Falls, Disability, Fractures, and Mortality in Older Men," *Journal of the American Geriatrics Society*, Vol. 57, No. 3, March 2009, pp. 492–498.

Faurot, Keturah R., Michele Jonsson Funk, Virginia Pate, M. Alan Brookhart, Amanda Patrick, Laura C. Hanson, Wendy Camelo Castillo, and Til Stürmer, "Using Claims Data to Predict Dependency in Activities of Daily Living as a Proxy for Frailty," *Pharmacoepidemiology and Drug Safety*, Vol. 24, No. 1, 2015, pp. 59–66.

Fried, Linda P., Catherine M. Tangen, Jeremy Walston, Anne B. Newman, Calvin Hirsch, John Gottdiener, Teresa Seeman, Russell Tracy, Willem J. Kop, Gregory Burke, and Mary Ann McBurnie, "Frailty in Older Adults: Evidence for a Phenotype," *Journals of Gerontology Series A*, Vol. 56, No. 3, March 2001, pp. M146–M157.

Gautam, Nileesa, Lily Bessette, Ajinkya Pawar, Raisa Levin, and Dae Hyun Kim, "Updating International Classification of Diseases 9th Revision to 10th Revision of a Claims-Based Frailty Index," *The Journals of Gerontology Series A*, Vol. 76, No. 7, July 2021, pp. 1316–1317.

Goldfarb, Michael, Melissa Bendayan, Lawrence G. Rudski, Jean-Francois Morin, Yves Langlois, Felix Ma, Kevin Lachapelle, Renzo Cecere, Benoit DeVarennes, and Christo I. Tchervenkov, James M. Brophy, and Jonathan Afilalo, "Cost of Cardiac Surgery in Frail Compared with Nonfrail Older Adults," *Canadian Journal of Cardiology*, Vol. 33, No. 8, August 2017, pp. 1020–1026.

Graham, James E., Soham Al Snih, Ivonne M. Berges, Laura A. Ray, Kyriakos S. Markides, and Kenneth J. Ottenbacher, "Frailty and 10-Year Mortality in Community-Living Mexican American Older Adults," *Gerontology*, Vol. 55, No. 6, 2009, pp. 644–651.

Hu, Li-tze, and Peter M. Bentler, "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 6, No. 1, 1999, pp. 1–55.

Kan, Hong J., Hadi Kharrazi, Bruce Leff, Cynthia Boyd, Ashwini Davison, Hsien-Yen Chang, Joe Kimura, Shannon Wu, Laura Anzaldi, Tom Richards, Elyse C. Lasser, and Jonathan P. Weiner, "Defining and Assessing Geriatric Risk Factors and Associated Health Care Utilization Among Older Adults Using Claims and Electronic Health Records," *Medical Care*, Vol. 56, No. 3, March 2018, pp. 233–239.

Keeler, Emmett, Jack M. Guralnik, Haijun Tian, Robert B. Wallace, and David B. Reuben, "The Impact of Functional Status on Life Expectancy in Older Persons," *Journals of Gerontology Series A*, Vol. 65, No. 7, July 2010, pp. 727–733.

Kim, Dae Hyun, Elisabetta Patorno, Ajinkya Pawar, Hemin Lee, Sebastian Schneeweiss, and Robert J. Glynn, "Measuring Frailty in Administrative Claims Data: Comparative Performance of Four Claims-Based Frailty Measures in the U.S. Medicare Data," *Journals of Gerontology Series A*, Vol. 75, No. 6, May 2020, pp. 1120–1125.

Kim, Dae Hyun, Sebastian Schneeweiss, Robert J. Glynn, Lewis A. Lipsitz, Kenneth Rockwood, and Jerry Avorn, "Measuring Frailty in Medicare Data: Development and Validation of a Claims-Based Frailty Index," *Journals of Gerontology Series A*, Vol. 73, No. 7, June 2018, pp. 980–987.

Kind, Amy J. H., Steve Jencks, Jane Brock, Menggang Yu, Christie Bartels, William Ehlenbach, Caprice Greenberg, and Maureen Smith, "Neighborhood Socioeconomic Disadvantage and 30-Day Rehospitalization: A Retrospective Cohort Study," *Annals of Internal Medicine*, Vol. 161, No. 11, December 2014, pp. 765–774.

Kuhn, Randall, Omar Rahman, and Jane Menken, "Survey Measures of Health: How Well Do Self-Reported and Observed Indicators Measure Health and Predict Mortality?" in Barney Cohen and Jane Menken, eds., *Aging in Sub-Saharan Africa: Recommendations for Furthering Research*, Washington, D.C.: National Academies Press, 2006, pp. 314–342.

Liu, Chiung-Ju, and Nancy K. Latham, "Progressive Resistance Strength Training for Improving Physical Function in Older Adults," *Cochrane Database of Systematic Reviews*, No. 3, July 2009.

McIsaac, D. I., P. E. Beaulé, G. L. Bryson, and C. Van Walraven, "The Impact of Frailty on Outcomes and Healthcare Resource Usage After Total Joint Arthroplasty: A Population-Based Cohort Study," *Bone & Joint Journal*, Vol. 98-B, No. 6, June 2016, pp. 799–805.

McNallan, Sheila M., Mandeep Singh, Alanna M. Chamberlain, Robert L. Kane, Shannon M. Dunlay, Margaret M. Redfield, Susan A. Weston, and Véronique L. Roger, "Frailty and Healthcare Utilization Among Patients with Heart Failure in the Community," *JACC: Heart Failure*, Vol. 1, No. 2, April 2013, pp. 135–141.

Mitnitski, Arnold B., Alexander J. Mogilner, and Kenneth Rockwood, "Accumulation of Deficits as a Proxy Measure of Aging," *Scientific World Journal*, Vol. 1, No. 321027, 2001, pp. 323–336.

National Bureau of Economic Research, "CMS' ICD-9-CM to and from ICD-10-CM and ICD-10-PCS Crosswalk or General Equivalence Mappings," webpage, last updated May 11, 2016. As of March 30, 2022: https://data.nber.org/data/icd9-icd-10-cm-and-pcs-crosswalk-general-equivalence-mapping.html

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan, "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," *Science*, Vol. 366, No. 6464, October 2019, pp. 447–453.

Office of the Assistant Secretary for Planning and Evaluation, *Report to Congress: Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program: The Second of Two Reports Required by the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014*, Washington, D.C.: U.S. Department of Health & Human Services, March 2020.

Rockwood, Kenneth, "Frailty and Its Definition: A Worthy Challenge," *Journal of the American Geriatrics Society*, Vol. 53, No. 6, June 2005, pp. 1069–1070.

Sager, Mark A., Nancy Cross Dunham, Amy Schwantes, Laurie Mecum, Kaia Halverson, and Diane Harlowe, "Measurement of Activities of Daily Living in Hospitalized Elderly: A Comparison of Self-Report and Performance-Based Methods," *Journal of the American Geriatrics Society*, Vol. 40, No. 5, May 1992, pp. 457–462.

Sternberg, Shelley A., Andrea Wershof Schwartz, Sathya Karunananthan, Howard Bergman, and A. Mark Clarfield, "The Identification of Frailty: A Systematic Literature Review," *Journal of the American Geriatrics Society*, Vol. 59, No. 11, November 2011, pp. 2129–2138.

Tibshirani, Robert, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58, No. 1, 1996, pp. 267–288.

Üstün, Bedirhan, and Cille Kennedy, "What Is 'Functional Impairment'? Disentangling Disability from Clinical Significance," *World Psychiatry*, Vol. 8, No. 2, June 2009, pp. 82–85.

Walter, Louise C., Richard J. Brand, Steven R. Counsell, Robert M. Palmer, C. Seth Landefeld, Richard H. Fortinsky, and Kenneth E. Covinsky, "Development and Validation of a Prognostic Index for 1-Year Mortality in Older Adults After Hospitalization," *JAMA*, Vol. 285, No. 23, 2001, pp. 2987–2994.

Xue, Qian-Li, "The Frailty Syndrome: Definition and Natural History," *Clinics in Geriatric Medicine*, Vol. 27, No. 1, February 2011, pp. 1–15.