Predicting Disability among Community-Dwelling Medicare Beneficiaries Using Claims-Based Indicators

Yonatan Ben-Shalom and David C. Stapleton

Objectives. To assess the feasibility of using existing claims-based algorithms to identify community-dwelling Medicare beneficiaries with disability based solely on the conditions for which they are being treated, and improving on these algorithms by combining them in predictive models.

Data Source. Data on 12,415 community-dwelling fee-for-service Medicare beneficiaries who first responded to the Medicare Current Beneficiary Survey (MCBS) in 2003–2006.

Study Design. Logistic regression models in which six claims-based disability indicators are used to predict self-reported disability. Receiver operating characteristic (ROC) curves were used to assess the performance of the predictive models.

Principal Findings. The predictive performance of the regression-based models is better than that of the individual claims-based indicators. At a predicted probability threshold chosen to maximize the sum of sensitivity and specificity, sensitivity is 0.72 for beneficiaries age 65 or older and specificity is 0.65. For those under 65, sensitivity is 0.54 and specificity is 0.67. The findings also suggest ways to improve predictive performance for specific disability populations of interest to researchers.

Conclusions. Predictive models that incorporate multiple claims-based indicators provide an improved tool for researchers seeking to identify people with disabilities in claims data.

Key Words. Self-reported disability, Medicare, claims-based algorithms, receiver operating characteristic (ROC) curves

Claims data are a potentially powerful tool for research on the long-term service and support (LTSS) needs of people with disabilities. Identification of people with disabilities in claims data—particularly those who live in the community and might benefit from home- and community-based service (HCBS) interventions—is helpful in assessments of interventions designed to improve

outcomes for the broad class of individuals who need such services. Although claims data could potentially be used to identify individuals with disabilities, current claims-based tools have limited ability to identify people with disabilities and to distinguish among important groups within this population.

Research on disability services must start by defining what is meant by disability. We use the International Classification of Functioning, Disability, and Health (ICF) as a conceptual framework for defining disability, and we adopt ICF definitions for disability-related terms. Under the ICF, a disability is a functional limitation resulting from the interaction between the individual's health, personal characteristics, and environment (Jette 2009). Disability consists of decreased functionality of a body function or structure (an *impairment*), decreased ability to perform an activity (an *activity limitation*), or decreased ability to participate in basic social roles (a *participation restriction*). Although we focus on identifying those with activity limitations, our methodology could be used to support research on other disability populations.

A major challenge in identifying disability in claims data is that not all disability-related conditions generate disability-related claims. Further, the disability-related information in claims data is far from what is required to determine whether a person has a disability (however defined), the severity of the disability, and whether it is a temporary or long-term disability (Iezzoni 2002, 2013).

Past efforts to identify people with disabilities using claims have been much less successful than those that use more extensive medical data. For instance, Bogardus et al. (2001) reviewed entire medical records from 519 hospitalizations for comments regarding limitations in 14 activities of daily living and instrumental activities of daily living (ADLs and IADLs, respectively) and compared them to interviews in which patients were asked specifically about these limitations. They were able to correctly classify 71 percent of those who reported a limitation in any of the 14 activities and 81 percent of those who reported no limitations (sensitivity = 0.71, specificity = 0.81). In contrast, Palsbo et al. (2008) used claims data to develop the Access Risk Classification System (ARCS) algorithm, which classifies people with disabilities according to their ability to access routine care. When they classified individuals into two groups of those needing proactive care coordination due to "one or more chronic conditions that cause major functional limitations," and all others,

Address correspondence to Yonatan Ben-Shalom, Ph.D., Mathematica Policy Research, 1100 1st Street NE, 12th Floor, Washington, DC 20002; e-mail: yben-shalom@mathematica-mpr.com. David C. Stapleton, Ph.D., is with Mathematica Policy Research, Washington, DC.

they correctly classified 83 percent of those who needed care coordination but only 30 percent of those who did not (sensitivity = 0.83, specificity = 0.30).

Another body of literature has focused on comparison of survey self-reports of inability to work due to a medically determinable condition to the eligibility decisions made by the Social Security Administration (SSA) for Social Security Disability Insurance (SSDI) applicants, reflecting the definition of disability used by that program (see, for instance, Benitez-Silva et al. 2004). Although the SSDI population is germane to our analysis, it is important to recognize that the concept of disability used here is quite different than the concept used by SSA; it reflects research interest in LTSS rather than in SSDI eligibility.

We explore the feasibility of improving existing claims-based algorithms and other criteria that identify disability by combining them in predictive models tailored to the disability measure and subgroup of interest. We estimate these models and assess their predictive performance using indicators of self-reported disability for Medicare Current Beneficiary Survey (MCBS) respondents. Using receiver operating characteristic (ROC) curves, we compare the sensitivity and specificity of the regression-based predictive models to those of the individual indicators (see the Methods section for further details on ROC analysis and Lasko et al. [2005] for a useful methodological review). We also compare the characteristics and service use of those for which the model prediction is congruent with self-reported disability and those for which it is incongruent.

The predictive models lead to disability classifications that are substantially more congruent with MCBS self-reports than the individual claims-based disability indicators. Congruence is much higher for Medicare beneficiaries age 65 or older than for those under age 65. Even so, for each age group, the predictions of the preferred model differ from self-reported disability for many respondents. Most notably, for those over age 65, the model predicts “no disability” relatively frequently when the individual self-reports both a disability and obesity, while for those ages 18–64, the model predicts “no disability” relatively frequently when the individual reports a disability and is also eligible for Medicaid (hereafter “dual-eligible” beneficiaries). In both age groups, the model frequently predicts “disability” when the claims data show relatively high service use during the observation period but survey responses indicate both no disability and relatively good health.

The findings suggest ways to make predictions from claims data more congruent with survey-based self-reports of disability when there is a specific research objective to guide the modeling. This includes the incorporation of

data on service use (such as home health care and skilled nursing facilities), prescriptions, and long-term support services provided to dual-eligible beneficiaries (from Medicaid claims). Researchers interested in disability populations different from those examined in this study could use the same methodology and data to develop a model for their population of interest.

METHODS

Data Source

In each year, the MCBS sample consists of four overlapping panels of Medicare beneficiaries (Murgolo 2010). Each panel is a nationally representative sample of Medicare beneficiaries with disproportionately large samples for those age 85 and older and those younger than 65. Individuals in the latter group are eligible for Medicare because they have been determined to have a work disability under the definition used for SSDI or have end-stage renal disease (ESRD). Respondents are interviewed in person three times a year for 4 years. Information on respondents in the community is collected through a community interview, at their home. Information on respondents in institutions is collected through an institutional interview of knowledgeable staff. For both groups, a baseline questionnaire is administered when the respondent first enters the survey. The Access to Care MCBS files include the always-enrolled Medicare population—those enrolled in Medicare at the beginning of the year and still alive for the MCBS fall interview. The fall survey collects information about beneficiaries' health status and functioning, access to medical providers, and satisfaction with their health care. Additional information on medical diagnoses and procedures is available from matched Medicare claims.

Study Population

We pooled four consecutive MCBS panels, including respondents who completed the community baseline questionnaire in 2003–2006 and either the community or facility questionnaire in their second MCBS year. We exclude from the sample: (1) respondents who were institutionalized at baseline, because of interest in LTSS for those residing in the community; (2) respondents who died before their second MCBS year, because we require disability data from the second-year survey (see below); and (3) beneficiaries who were enrolled in Medicare Advantage in any month during their first and second

MCBS calendar year, because they have no claims data. The study samples include 10,057 Medicare beneficiaries age 65 or older and 2,357 ages 18–64.

Self-Reported Disability Indicators

We construct three indicators of self-reported disability based on survey responses in the fall of the second MCBS year regarding difficulties in performing ADLs or IADLs (Katz et al. 1963). The ADLs include bathing/showing, dressing, eating, transferring (getting in/out of bed/chair), walking, and toileting. The IADLs include using the telephone, light housework, heavy housework, meal preparation, shopping, or managing money. For each ADL and IADL, we first create a binary indicator for difficulty in performing the activity or inability to perform the activity due to health. The disability indicator used in the results presented here equals 1 if the individual reports having one or more ADL limitations. We have also assessed the sensitivity of the results to more restrictive (at least three ADLs) and expansive (at least one ADL or IADL) definitions.

Claims-Based Disability Indicators

We construct six claims-based disability indicators using information in Medicare claims from the first and second MCBS years. None of these indicators was specifically designed to identify those with disabilities according to the ADL/IADL definitions we have adopted for this study, but each identifies groups that seem very likely to have ADL/IADL limitations. All of the groups overlap with each other.

The first indicator is the Chronic Illness and Disability Payment System (CDPS), developed by Kronick et al. (2000) as a diagnostic classification system to support risk-adjusted Medicaid payments to health maintenance organizations. The CDPS was designed to predict health expenditures but has also been used to identify people with disabilities (Irvin and Johnson 2006; Pavetti et al. 2010). The CDPS classifies individuals into 19 diagnostic categories and further disaggregates them by cost (extra high, very high, high, medium, low, very low, and super low). Our CDPS indicator equals 1 if the algorithm classifies the individual as medium cost or above in any diagnostic category, and 0 otherwise.

Palsbo et al. (2008) developed ARCS to identify and classify people with disabilities according to their ability to access routine care, based on claims data. The algorithm uses ICD-9 codes; Healthcare Common Proce-

dure Coding System (HCPCS) codes; Current Procedural Terminology, 4th edition (CPT-4) codes; and the number of prescriptions. Our ARCS indicator equals 1 for individuals in the highest-risk class (level 3), those with “multiple chronic conditions and complex medical needs that severely impair a person’s independence.”¹

SSA’s Health Information Technology (SSA-HIT) business rules (Social Security Administration 2012) are being developed to identify diagnoses and procedure codes or combinations of codes that are sufficient to meet or equal SSA’s Listing of Impairments—a primary component of the medical criteria that SSA uses to determine eligibility for SSDI, as well as Supplemental Security Income for those under age 65. These rules cover only those impairments most likely to be identified in claims data and are preliminary. Our SSA-HIT indicator equals 1 for all individuals identified as having one of these impairments.²

Our three additional indicators identify those with diagnoses of (1) psychiatric disorders (denoted PSYCH); (2) cognitive disorders (dementia and Alzheimer’s, or DEMALZ); and (3) intellectual disability (ID). PSYCH equals 1 if an ICD-9 code in a claim begins with 295 (schizophrenia) or 296 (affective disorder), or any one of the ICD-9 codes indicating psychotic, neurotic, personality, and other nonpsychotic mental disorders³; the DEMALZ indicator equals 1 for the ICD-9 codes 290.2, 290.3, 290.4, 290.8, and 331.0; the ID indicator equals 1 for ICD-9 codes 317 (mild intellectual disability), 318.0 (moderate intellectual disability), 318.1 (severe intellectual disability), 318.2 (profound intellectual disability), and 319 (unspecified intellectual disability).

Analytical Methods

Measuring the Performance of Claims-Based Disability Indicators in Predicting Self-Reported Disability. For each age group, we estimate a logistic model in which the disability indicator is the dependent variable and the main predictors are the six claims-based disability indicators:

$$P(\text{disability} = 1) = \text{logit}(\beta_0 + \beta_1 \text{CDPS} + \beta_2 \text{ARCS} + \beta_3 \text{SSAHIT} + \beta_4 \text{PSYCH} + \beta_5 \text{DEMALZ} + \beta_{ID} + \gamma X + \varepsilon)$$

For both age groups, X includes a small number of variables from Medicare administrative records—age, sex, and an indicator for dual-eligibility (defined as any Medicaid participation during the year). For those age 65 or older, additional predictors include an indicator for Medicare entitlement

before age 65 (indicative of SSDI or ESRD prior to age 65) and interactions between this variable and the dual-eligibility indicator and each of the six disability indicators.⁴

For each estimated model, we construct an ROC curve that depicts all possible sensitivity–specificity combinations for that model. ROC analysis evaluates the predictive performance of a model by capturing the model’s trade-off between sensitivity and specificity. “Sensitivity” is the proportion of cases with a condition that the model correctly classifies as having the condition, and “specificity” is the proportion of cases not having the condition that are correctly classified as not having the condition. Here, the condition is our MCBS disability indicator. Each point on a model’s ROC curve corresponds to a threshold that could be used along with the model’s predicted value to classify observations for the presence of the condition: If the predicted value is above the threshold, the observation is labeled positive; if not, it is labeled negative.

The ROC curve is a plot that shows how sensitivity and specificity change as the threshold varies from a value so high that all cases are classified as negative to one that is so low that all cases are classified as positive, with sensitivity measured on the vertical axis and one minus specificity (or the false-positive rate) measured on the horizontal axis. The ROC curve starts at the (0, 0) point and slopes upward until it reaches the (1,1) point. The most commonly used index for a model’s accuracy is the area under the ROC curve (AUC), which is also the probability that the predicted value for a randomly chosen positive is greater than the predicted value for a randomly chosen negative (Bamber 1975). The AUC for a perfect test is 1.0, while the AUC for random chance is 0.5. The latter corresponds to an ROC curve that follows the diagonal from (0,0) to (1,1).

In the context of this study, it is important to recognize that the terms “true” and “false” do not imply that the self-report of disability is correct and the prediction is either correct or incorrect. Instead, “true” means that the prediction matches the survey response and “false” means that it does not. The adopted terminology is the technical standard for the literature and reflects the goal of using claims data to predict survey responses.

For each age group, we determine the point (and respective predicted probability threshold) on the ROC curve that maximizes the sum of sensitivity and specificity, called Youden’s Index (Youden 1950). We then compare the sensitivity and specificity of the model’s prediction at Youden’s Index to the sensitivity and specificity of the individual claims-based indicators.

Analysis of Congruence. For the model that produces predictions most congruent with self-reported disability, we conduct additional analysis to better understand the reasons why some cases are congruent and others are not. We first use the estimated-logit model and its respective Youden's Index to classify the MCBS respondents in our sample into one of the four groups: (1) those who self-report a disability and are also classified as having a disability according to the model (true positives), (2) those who self-report a disability but are classified as not having a disability (false negatives), (3) those who do not self-report a disability but are classified as having a disability (false positives), and (4) those who do not self-report a disability and are classified as not having a disability (true negatives). We then compare descriptive statistics for numerous sociodemographic characteristics, self-reported health variables, and measures of service use across these four groups. These descriptive statistics are based on data for the second MCBS year.

RESULTS

Self-Reported and Claims-Based Disability

Among individuals age 65 or older, 26.5 percent reported a limitation in at least one ADL—the definition of disability used in the reported analyses (Table 1). Prevalence under the two definitions used in the unreported sensitivity analysis is considerably different: 7.9 percent reported limitations in three or more ADL, and 40.3 percent reported a limitation in at least one ADL or IADL. The respective percentages are considerably higher among the subgroup that first qualified for Medicare before age 65 and are much closer to those for the younger age group.

Among individuals ages 18–64, 16.5 percent reported limitations in three or more ADLs, 44.0 percent reported a limitation in at least one ADL, and 70.3 percent reported a limitation in at least one ADL or IADL. Notably, close to 30 percent in this age group reported no ADL or IADL, reflecting the difference between the SSDI definition of disability and our own. The CDPS, ARCS, and SSA-HIT indicators pick up roughly similar percentages of each age group; the PSYCH and ID indicators pick up notably larger percentages of those ages 18–64, and the DEMALZ indicator picks up a notably larger percentage of individuals age 65 or older. In each age group, the percentage of beneficiaries classified as having a disability by each claims-based indicator is higher for those who self-report a limitation in at least one ADL than for others (Table 2).

Table 1: Percentage of Medicare Beneficiaries with Disability, Self-Reported, and Claims-Based

	<i>All</i>		<i>Age 65 or Older</i>				<i>Age 18–64</i>	
			<i>All</i>		<i>Medicare before Age 65</i>			
	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>	<i>Mean</i>	<i>SE</i>
Self-reported disability								
Three or more ADL limitations	9.4	0.4	7.9	0.3	17.3	1.5	16.5	1.4
One or more ADL limitations	29.4	0.7	26.5	0.6	47.5	2.1	44.0	1.6
One or more ADL or IADL limitations	45.4	0.7	40.3	0.7	65.6	2.1	70.3	1.7
Claims-based flag								
CDPS (“medium cost” and above)	50.3	0.5	49.1	0.6	66.3	1.8	56.3	1.3
ARCS (“level 3”)	72.8	0.5	73.2	0.6	81.1	1.4	70.5	1.5
SSA-HIT	26.5	0.5	26.9	0.6	40.2	1.8	24.7	1.1
Psychiatric disorder	18.0	0.4	14.5	0.4	22.2	1.7	35.1	1.4
Dementia/Alzheimer’s	3.1	0.2	3.6	0.2	4.5	0.6	0.6	0.2
Intellectual disability	0.4	0.1	0.1	0.0	0.2	0.1	2.1	0.3
Number of observations								
Unweighted	12,415		10,957		840		2,357	
Weighted (millions)	35.3		29.3		2.2		6.0	

Notes. Based on analysis of the Medicare Current Beneficiary Survey (MCBS). The study population includes respondents who completed the community baseline questionnaire in 2003–2006 and either the community or facility questionnaire in their second MCBS year. The claims-based indicators are based on Medicare claims from each beneficiary’s first and second MCBS years. Beneficiaries enrolled in Medicare Advantage in any month during their first and second MCBS calendar year are excluded, as are those who died before their second MCBS year.

Logistic Model Estimates

The logistic regression results for the models that predict having at least one ADL limitation appear in Table 3. For comparison, we include results for reference models that include only sex and age as predictors; they indicate the predictive performance of our approach using only the most basic variables.

In the older age group (age 65 or older), being female and of older age are both positively associated with the disability indicator (self-reported limitation in at least one ADL), as are dual-eligibility and Medicare entitlement before age 65. Of the six claims-based indicators, only the ID indicator is not a significant predictor, likely reflecting the very small number of MCBS respondents flagged by the latter as well as the extent to which those with ID are identified by the other indicators. The estimated odds ratios for many of the

Table 2: Cross-Tabulation of Self-Reported Limitation in at Least One ADL (Compared to No Limitation) and Claims-Based Disability

	Age 65 or Older						Age 18–64			
	All			All			Medicare before Age 65			
	1+ ADL Limitation	No ADL Limitation		1+ ADL Limitation	No ADL Limitation		1+ ADL Limitation	No ADL Limitation	1+ ADL Limitation	No ADL Limitation
Claims-based flag										
CDPS (medium cost and above)	67.0*	43.3		69.3*	41.8		73.6*	59.7	60.2*	53.4
ARCS (level 3)	82.8*	68.6		85.2*	68.9		87.4*	75.3	75.6*	66.5
SSA-HIT	35.3*	22.9		37.7*	23.0		43.2	37.4	28.2*	21.9
Psychiatric disorder	26.3*	14.6		23.0*	11.5		25.0	19.7	36.1*	34.2
Dementia/Alzheimer's	6.7*	1.6		8.6*	1.8		6.9*	2.3	1.0*	0.2
Intellectual disability	0.8*	0.3		0.2	0.0		0.2	0.2	2.4	2.0
Number of observations	3,899	8,515		2,953	7,104		394	446	946	1,411
Unweighted										
Weighted (millions)	10.4	24.9		7.8	21.6		1.0	1.1	2.6	3.3

Notes. Based on analysis of the Medicare Current Beneficiary Survey (MCBS). The study population includes respondents who completed the community baseline questionnaire in 2003–2006 and either the community or facility questionnaire in their second MCBS year. The claims-based indicators are based on Medicare claims from each beneficiary's first and second MCBS years. Beneficiaries enrolled in Medicare Advantage in any month during their first and second MCBS calendar year are excluded, as are those who died before their second MCBS year.

*Difference from percentage for "No ADL limitation" group is statistically significant, $p < .01$ in all cases.

Table 3: Logistic Regression Results for Models Predicting a Limitation in at Least One ADL

Variable	Age 65 or Older		Age 18–64	
	Reference Model (OR)	Full Model (OR)	Reference Model (OR)	Full Model (OR)
Female	1.32*	1.31*	1.11	1.09
Age	1.08*	1.07*	1.03*	1.02*
Disability flags				
CDPS (“medium cost” and above)		2.12*		1.04
ARCS (“level 3”)		1.18*		1.51*
SSA-HIT		1.26*		1.19
Psychiatric disorder		1.49*		1.04
Dementia and Alzheimer’s		2.47*		4.62*
Intellectual disability		4.83		1.52
Dual-eligible†		2.43*		0.63*
Medicare entitlement before age 65		4.17*		
Interaction with Medicare entitlement before age 65				
CDPS (“medium cost” and above)		0.66*		
ARCS (“level 3”)		1.54		
SSA-HIT		0.73		
Psychiatric disorder		0.67		
Dementia and Alzheimer’s		0.74		
Intellectual disability		0.16		
Dual-eligible†		0.55*		
Area under ROC curve	0.65	0.75	0.58	0.62

*Odds ratio is statistically significantly different from 1.00, $p < .05$.
†We identified MCBS respondents as dually eligible if they had any Medicaid eligibility during their second MCBS year.

interaction terms are smaller than 1.0, indicating that that the claims-based indicators are more strongly associated with our disability indicator for those who aged into Medicare than for those who entered Medicare before age 65.

In the younger age group (ages 18–64), being female is not significantly associated with the disability indicator, older age is positively associated with it, and dual-eligibility is negatively associated with it. Of the six claims-based indicators, only the ARCS and DEMALZ indicators are significant predictors.

ROC Curves and Predicted Probability Thresholds

In Figure 1, we show the ROC curves corresponding to the full model for Medicare beneficiaries age 65 or older (top panel) and ages 18–64 (bottom panel). Each panel also includes: an ROC curve corresponding to the age–sex

reference model; four points corresponding to the sensitivity and specificity of the individual CDPS, ARCS, SSA-HIT, and PSYCH indicators; and Youden's Index.⁵

For those age 65 or older, the full logistic model outperforms each of the individual claims-based indicators, allowing for both higher sensitivity and specificity in each case. That model's AUC is 0.75, above the 0.70 value that is commonly deemed acceptable in the ROC literature (see Hosmer and Lemeshow 2004). The point that maximizes the sum of sensitivity and specificity (Youden's Index) corresponds to a predicted probability threshold of 0.24. Therefore, we classify Medicare beneficiaries age 65 or older as having disability if their predicted value is greater than 0.24.

The full logistic model for Medicare beneficiaries ages 18–64 also outperforms each of the individual claims-based indicators. For this age group, none of the individual claims-based indicators performs as well as the reference model. That is not remarkable for the SSA-HIT indicators, as these indicators focus on identification of subpopulations with specific conditions, but it is remarkable for CDPS and ARCS, which are designed to capture broader groups. Although the full model outperforms the other options considered, its ROC curve is well below that for the older age group and has an AUC of 0.62, well below the 0.70 value often considered to be acceptable in the literature. The point that maximizes the sum of sensitivity and specificity (Youden's Index) corresponds to a predicted probability threshold of 0.45. We therefore classify Medicare beneficiaries ages 18–64 as having disability if their predicted value is greater than 0.45.

Congruence Analysis

Demographic and health statistics for Medicare beneficiaries who are age 65 or older and those who are ages 18–64 appear in Table 4. Using the Youden's Index as the critical value for classifying cases, our model prediction is congruent with self-reported disability for 5.6 million of the 7.8 million beneficiaries age 65 or older with a limitation in at least one ADL (sensitivity = 0.72). The prediction is also congruent with self-reported disability for 14.1 million of the 21.6 million without any ADL limitations (specificity = 0.65). Those who do not self-report a disability but are classified as having a disability (false positives) are similar to those who self-report a disability and are also classified as having a disability (true positives) in many ways—not surprisingly, given the method used to classify them and the fact that we classified those in both groups as positive. The characteristics reported in the MCBS are not necessar-

Figure 1: ROC Curve versus Claims-Based Disability Flags, Limitation in at Least One ADL: (a) Age 65 or older; (b) Age 18-64

Notes. The CDPS flag includes all diagnostic subcategories identified as medium cost and above, and the ARCS flag includes only the algorithm's highest-risk class (level 3). Beneficiaries age 65 or older are identified as having disability if their predicted probability is above 0.24. Beneficiaries ages 18–64 are identified as having disability if their predicted probability is above 0.45.

ily captured in the model's predictors, however, so it is also not surprising that we find some differences between the false and true positives. The false positives are in better health (according to their self-report) than the true positives and have lower service use and fewer prescriptions filled during their second MCBS year.

Similarly, Medicare beneficiaries age 65 or older who self-report disability but are classified as not having a disability (false negatives) resemble those who do not self-report disability and are also classified as not having a disability (true negatives) in many respects. However, the false negatives report worse health than the true negatives and have higher service use and more prescriptions filled. Compared with true positives, false negatives (identified as having a disability by the survey but classified as not having a disability by the model) are especially characterized by younger ages, higher education, higher prevalence of obesity, and lower service use.

Our model for the 18–64 age group congruently classifies 1.4 million of the 2.6 million beneficiaries with at least one ADL limitation (sensitivity = 0.54), and also congruently classifies 2.2 million of the 3.3 million without any ADL limitations (specificity = 0.67). As with the older age group, false positives resemble true positives in many ways, but the false positives report better health and lower service use. Also as with the older age group, the false negatives are similar to the true negatives in many ways, but the false negatives report worse health and higher service use. Compared with true positives, false negatives are especially characterized by younger ages, dual-eligibility, fewer reported illnesses (data not shown), and lower service use.

DISCUSSION

Our objective was to improve upon the ability of claims-based indicators to identify community-dwelling Medicare beneficiaries with self-reported

Table 4: Characteristics for Medicare Beneficiaries Classified Congruently and Incongruently by the Predictive Model, by Self-Reported Disability Status*

Variable (Percent Unless Otherwise Noted)	1+ ADL Limitation				No ADL Limitation			
	Congruent (True Positive)		Incongruent (False Negative)		Incongruent (False Positive)		Congruent (True Negative)	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Age 65 or older								
Age at Medicare enrollment								
Less than 65	17.9	0.9	1.6	0.5 [†]	14.3	0.6 ^{†,‡}	0.6	0.1
65 or above	82.1	0.9	98.4	0.5 [†]	85.7	0.6 ^{†,‡}	99.4	0.1
Age								
65–74	24.3	1.1	68.5	1.7 ^{†,‡}	28.0	1.0 [‡]	73.9	0.5
75–84	47.6	1.1	28.8	1.5 [†]	51.3	0.9 ^{†,‡}	24.8	0.5
85 or above	28.1	0.8	2.7	0.6 [†]	20.7	0.8 ^{†,‡}	1.3	0.1
Female	65.9	1.0	53.9	2.0 [†]	62.4	1.0 [‡]	48.5	0.8
Dual-eligible	25.4	1.3	3.9	0.8 [†]	18.0	1.1 ^{†,‡}	2.1	0.3
Married	39.2	0.9	62.5	1.8 [†]	45.9	1.0 ^{†,‡}	67.1	0.8
At least some college	26.1	1.4	39.4	2.0 [†]	30.9	1.4	43.5	1.3
Self-reported health								
Health status is good/ very good/excellent	51.9	1.3	69.2	1.9 ^{†,‡}	80.1	0.8 ^{†,‡}	93.0	0.4
Health limits social activity	38.0	1.2	17.5	1.7 ^{†,‡}	8.8	0.5 ^{†,‡}	2.7	0.3
Obese	27.2	1.0	39.0	2.0 ^{†,‡}	19.8	1.0 ^{†,‡}	20.7	0.7
Medical events (number of)								
Home health visits	60.6	4.6	12.0	2.9 ^{†,‡}	12.2	2.0 ^{†,‡}	1.3	0.3
Outpatient visits	7.2	0.3	4.6	0.4 ^{†,‡}	6.1	0.2 ^{†,‡}	3.1	0.1
Medical provider visits	49.4	1.2	28.5	1.5 ^{†,‡}	34.5	0.7 ^{†,‡}	18.0	0.4
Prescriptions	51.0	1.2	35.0	1.4 ^{†,‡}	40.0	0.8 ^{†,‡}	22.7	0.4
Number of observations								
Unweighted	2,289		664		2,915		4,189	
Weighted (millions)	5.6		2.2		7.5		14.1	
Age 18–64								
Age								
Less than 50	13.1	1.6	54.7	3.0 [†]	20.1	2.7 [‡]	59.5	2.1
50–64	86.9	1.6	45.3	3.0 [†]	79.9	2.7 [‡]	40.5	2.1
Female	55.4	3.3	44.5	2.6 [†]	58.9	3.5 [‡]	41.2	2.1
Dual-eligible	14.2	2.1	57.5	3.0 [†]	13.7	2.3 [‡]	62.5	2.2
Married	52.2	3.1	37.4	3.1 [†]	45.5	3.2 [‡]	28.5	2.0
At least some college	33.6	3.1	31.0	2.4	34.7	3.4	27.5	2.0
Self-reported health								
Health status is good/ very good/excellent	23.5	2.7	25.2	2.3 [‡]	47.8	3.6 [†]	55.0	2.1
Health limits social activity	45.6	3.4	47.7	3.0 [‡]	23.6	3.3 [†]	18.6	2.1
Obese	50.6	3.6	46.2	2.8 [‡]	42.1	4.0	35.0	2.1

continued

Table 4. *Continued*

Variable (Percent Unless Otherwise Noted)	1+ ADL Limitation				No ADL Limitation			
	Congruent (True Positive)		Incongruent (False Negative)		Incongruent (False Positive)		Congruent (True Negative)	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Medical events (number of)								
Home health visits	35.0	9.0	30.7	5.3 [‡]	8.8	2.9 [†]	7.1	1.4
Outpatient visits	9.3	0.9	7.0	0.7	6.8	0.7	5.6	0.4
Medical provider visits	44.4	3.1	27.6	1.9 ^{†,‡}	31.9	2.5 ^{†,‡}	20.2	1.6
Prescriptions	64.7	3.0	55.0	3.4 [‡]	49.0	2.7 [†]	40.1	2.6
Number of observations								
Unweighted	381		565		306		1,105	
Weighted (millions)	1.4		1.2		1.1		2.2	

*MCBS respondents are true positive if they have a limitation in at least one ADL and are predicted to have disability by the model, false negative if they have a limitation in at least one ADL and are not predicted to have disability by the model, false positive if they do not have a limitation in at least one ADL and are predicted to have disability by the model, and true negative if they do not have a limitation in at least one ADL and are not predicted to have disability by the model. Beneficiaries age 65 or older (18–64) are identified as having disability if their predicted probability is above 0.24 (0.45). Self-reported health variables and measures of service use are based on data for the second MCBS year.

[†]Difference from true positive is statistically significant, $p < .01$.

[‡]Difference from true negative is statistically significant, $p < .01$.

disabilities by using multiple indicators along with age, sex, and dual-eligibility variables in predictive models. The findings indicate that the predictive models outperform the individual indicators in that they offer an improved sensitivity-specificity trade-off. We also find that the sensitivity-specificity trade-off for the models is considerably better for those who are age 65 or older than for those ages 18–64.

The relatively low congruence of the model's predictions for the younger age group might be due to the high percentage of dual-eligibles in that group, for whom we may be missing key service use because some services for this group are covered through Medicaid and not Medicare. Another possible explanation is related to errors in the self-reporting of ADL and IADL limitations in the MCBS. Measurement error probably accounts for a larger share of variation in the self-reported disability indicators for the under-65 population than for the 65-or-older population—only 44 percent of the younger group report one or more ADL limitations even though they all qualified for Medicare due to SSDI eligibility and not age.

In many ways, beneficiaries in both age groups who do not self-report disability but are classified as having a disability by the model (false

positives) are similar to those who self-report disability and are also classified as having a disability (true positives), but they report being in better health, use fewer services, and fill fewer prescriptions. Similarly, those who report disability but are classified as not having a disability (false negatives) resemble those who do not report disability and are also classified as not having a disability (true negatives), but they report worse health, use more services, and fill more prescriptions. These findings suggest that data on such service use as home health care and skilled nursing facilities as well as the number of prescriptions filled may further help differentiate between those with and without disabilities.

Beneficiaries who are age 65 or older and who self-report disability but are classified as not having a disability are characterized by high obesity rates. Use of obesity-related information from claims data, which are currently not included in our set of predictors, might therefore improve the identification of those with disability in this age group—at least as disability is defined in this study. A large share of beneficiaries ages 18–64 who self-report disability but are classified by the model as not having a disability are eligible for Medicaid as well as Medicare, which suggests that information on use of long-term support services, captured in Medicaid data, might also improve prediction.

Researchers interested in disability populations different from that examined in this study could use the same methodology and data to develop a model for their population of interest. Our analysis focused on limitations in ADLs and IADLs, but disability is a broad term that encompasses many dimensions and varies in severity. This study did not exhaust all of the disability-related questions included in the MCBS, and other disability populations can be defined. Predictive models can also be tailored to various other subgroups of Medicare beneficiaries. Other matches of administrative and survey data may facilitate targeting of additional populations.

Within the context of MCBS analysis, several options might increase the congruence of claims-based predictions with self-reported disability indicators. Adding data from Medicaid claims for the dually eligible MCBS respondents would likely improve congruence. Using a larger set of claims-based indicators, with separate indicators for different categories from the CDPS and SSA-HIT rules, could also result in improved congruence. Other claims-based algorithms, even if not specifically designed to identify people with disability, might also improve performance. Relevant examples include the Center for Medicare and Medicaid Services (CMS) Hierarchical Condition Category risk adjuster (Pope et al. 2004) and the

chronic conditions included in CMS's Chronic Conditions Data Warehouse (2013).

One practical application of using the approach presented here is in the targeting of health and disability service interventions. Frail Medicare beneficiaries who are age 65 or older, live in the community, and are more likely than others to benefit from certain HCBS interventions are a case in point. In the past, several interventions have been deemed too costly because many who received the services did not actually need them. For example, the Medicare Coordinated Care Demonstration evaluation (Brown et al. 2012) found that while only 1 of the 11 demonstration sites that operated for 6 years significantly reduced hospitalizations overall, 4 of the programs significantly reduced hospitalizations for a subset of high-need patients. For these four programs, overall effects were due almost entirely to the large effects on the high-need subgroup. Other programs noted that one reason they were unsuccessful in reducing hospitalizations was their inability to focus resources on only the subset of high-risk need enrollees. Improved targeting of the relevant population would improve the cost-effectiveness of the intervention because a larger proportion of those who need the services will be included. More broadly, and consistent with the goals of comparative effectiveness research, this approach could be used to define alternative target populations for the purpose of comparing the effectiveness of various interventions across those populations.

This approach might also have practical applications in program administration. For instance, if the absence of certain ADL or IADL limitations is highly predictive of ability to work, as defined for the SSDI program, then this approach could be used to identify beneficiaries whose eligibility should be reviewed. Such models could also be applied to the medical claims data of SSDI applicants, to group applications into those that could be allowed or denied with little additional scrutiny (high or low predicted probabilities), and those that require substantial additional development—in essence an extension of the SSA-HIT rules.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The authors thank Chelsea Swete, Barton Baker, and Mei-Ling Mason for programming support, and Craig Thornton and Randall Brown for useful comments on preliminary drafts. This study was conducted under contract HHSP23320095642WC between the U.S. Department of Health and Human Services (HHS), Office of Disability,

Aging, and Long-Term Care Policy (DALTCP), and Mathematica Policy Research. The opinions and views expressed in this paper are those of the authors. They do not necessarily reflect the views of the Department of Health and Human Services or Mathematica Policy Research.

Disclosures: None.

Disclaimers: None.

NOTES

1. We do not include information on prescriptions in our ARCS indicator.
2. Our implementation of the SSA-HIT indicator omits the few rules that include a time dimension, reflecting the complexity of the longitudinal analysis required to implement them.
3. The following ICD-9 codes were included in addition to 295 (schizophrenia) and 296 (affective disorder): 297–301, 306–7, 309–10, and 314–315.
4. We estimated simplified specifications that included either a single claims-based indicator or a single flag indicating identification by at least one of the claims-based indicators. For both age groups, the full models described in the text are at least as accurate as these alternative specifications. Results from all analyses performed are available upon request from the authors.
5. Not shown here are the indicators for intellectual disability and dementia/Alzheimer's, which pick up only a tiny fraction of each age group.

REFERENCES

- Bamber, D. 1975. "The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graph." *Journal of Mathematical Psychology* 12 (4): 387–415.
- Benitez-Silva, H., M. Buchinsky, H. M. Chan, S. Cheidvasser, and J. Rust. 2004. "How Large Is the Error in Self-Reported Disability?" *Journal of Applied Econometrics* 19 (6): 649–70.
- Bogardus, S. T., V. Towle, C. S. Williams, M. M. Desai, and S. Inouye. 2001. "What Does the Medical Record Reveal about Functional Status?" *Journal of General Internal Medicine* 16 (11): 728–36.
- Brown, R. S., D. Peikes, G. Peterson, J. Schore, and C. M. Razafindrakoto. 2012. "Six Features of Medicare Coordinated Care Demonstration Programs That Cut Hospital Admissions of High-Risk Patients." *Health Affairs* 31 (6): 1156–66.
- Chronic Conditions Data Warehouse. 2013. [accessed on March 29, 2013]. Available at <http://www.ccwdata.org/chronic-conditions/index.htm>
- Hosmer, D. W., and S. Lemeshow. 2004. *Applied Logistic Regression*. San Francisco, CA: John Wiley & Sons.

- Iezzoni, L. I. 2002. "Using Administrative Data to Study Persons with Disabilities." *Milbank Quarterly* 80 (2): 347–79.
- . 2013. "Imperatives for HSR Addressing: Individuals with Disabilities—The Canaries in Health Care's Coal Mine." *Medical Care* 51 (2): 133–6.
- Irvin, C. V., and C. Johnson. 2006. *Medicaid Populations with Chronic and Disabling Conditions: A Compilation of Data on Their Characteristics, Health Conditions, Service Use, and Medicaid Payments*. Washington, DC: Mathematica Policy Research.
- Jette, A. M. 2009. "Toward a Common Language of Disablement." *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 64 (11): 1165–8.
- Katz, S., A. B. Ford, R. W. Moskowitz, B. A. Jackson, and M. W. Jaffe. 1963. "Studies of Illness in the Aged." *Journal of the American Medical Association* 185 (12): 914–9.
- Kronick, R., T. Gilmer, T. Dreyfus, and L. Lee. 2000. "Improving Health-Based Payment for Medicaid Beneficiaries: CDPS." *Health Care Financing Review* 21 (3): 29–64.
- Lasko, T. A., J. G. Bhagwat, K. H. Zou, and L. Ohno-Machado. 2005. "The Use of Receiver Operating Characteristic Curves in Biomedical Informatics." *Journal of Biomedical Informatics* 38 (5): 404–15.
- Murgolo, M. 2010. "The Characteristics and Perceptions of the Medicare Population: Data from the 2010 Medicare Current Beneficiary Survey (Appendix A: Technical Documentation for the Medicare Current Beneficiary Survey)" [accessed on March 26, 2013]. Available at http://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Downloads/2010_C_and_P.zip
- Palsbo, S. E., C. D. Sutton, M. F. Mastal, S. Johnson, and A. Cohen. 2008. "Identifying and Classifying People with Disabilities Using Claims Data: Further Development of the Access Risk Classification System (ARCS) Algorithm." *Disability and Health Journal* 1 (4): 215–23.
- Pavetti, L. D., M. Derr, J. Kauff, and A. Barrett. 2010. "Mental Disorders and Service Use among Welfare and Disability Program Participants in Fee-for-Service Medicaid." *Psychiatric Services* 61 (5): 495–9.
- Pope, G. C., J. Kautter, R. P. Ellis, A. S. Ash, J. Z. Ayanian, M. J. Ingber, J. M. Levy, and J. Robst. 2004. "Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model." *Health Care Financing Review* 25 (4): 119–41.
- Social Security Administration. 2012. Social Security & Health IT. SSA Publication No. 02-007. Washington, DC.
- Youden, W. J. 1950. "Index for Rating Diagnostic Tests." *Cancer* 3 (1): 32–5.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.