Combinations of Chronic Conditions, Functional Limitations, and Geriatric Syndromes that Predict Health Outcomes

Siran M. Koroukian, Ph.D.[1], Nicholas Schiltz, Ph.D.[1], David F. Warner, Ph.D.[2], Jiayang Sun, Ph.D.[1], Paul M. Bakaki, M.D., Ph.D.[1], Kathleen A. Smyth, Ph.D.[1], Kurt C. Stange, M.D., Ph.D.[3], and Charles W. Given, Ph.D.[4]

[1]Department of Epidemiology and Biostatistics, School of Medicine, Case Western Reserve University, Cleveland, OH, USA; [2]Department of Sociology, University of Nebraska–Lincoln, Lincoln, NE, USA; [3]Department of Family Medicine and Community Health, School of Medicine, Epidemiology and Biostatistics, Sociology, Case Comprehensive Cancer Center, Cleveland Clinical & Translational Science Collaborative, Case Western Reserve University, Cleveland, OH, USA; [4]Department of Family Medicine, Michigan State University, East Lansing, MI, USA.

BACKGROUND: The strategic framework on multiple chronic conditions released by the US Department of Health and Human Services calls for identifying homogeneous subgroups of older adults to effectively target interventions aimed at improving their health.

OBJECTIVE: We aimed to identify combinations of chronic conditions, functional limitations, and geriatric syndromes that predict poor health outcomes.

DESIGN, SETTING AND PARTICIPANTS

Data from the 2010–2012 Health and Retirement Study provided a representative sample of U.S. adults 50 years of age or older (*n*=16,640).

MAIN MEASURES: Outcomes were: Self-reported fair/poor health, self-rated worse health at 2 years, and 2-year mortality. The main independent variables included self-reported chronic conditions, functional limitations, and geriatric syndromes. We conducted tree-based classification and regression analysis to identify the most salient combinations of variables to predict outcomes.

KEY RESULTS: Twenty-nine percent and 23 % of respondents reported fair/poor health and self-rated worse health at 2 years, respectively, and 5 % died in 2 years. The top combinations of conditions identified through our tree analysis for the three different outcome measures (and percent respondents with the outcome) were: a) for fair/poor health status: difficulty walking several blocks, depressive symptoms, and severe pain (> 80 %); b) for self-rated worse health at 2 years: 68.5 years of age or older, difficulty walking several blocks and being in fair/poor health (60 %); and c) for 2-year mortality: 80.5 years of age or older, and presenting with limitations in both ADLs and IADLs (> 40 %).

CONCLUSIONS: Rather than chronic conditions, functional limitations and/or geriatric syndromes were the most prominent conditions in predicting health outcomes. These findings imply that accounting for chronic conditions alone may be less informative than also accounting for the co-occurrence of functional limitations and geriatric syndromes, as the latter conditions appear to drive health outcomes in older individuals.

INTRODUCTION

More than two-thirds of Medicare beneficiaries have two or more chronic conditions, and more than 15 % present with six or more.[1,2] As stated in the 2010 US Department of Health and Human Services "Strategic Framework on Multiple Chronic Conditions",[3] the aging of the U.S. population warrants a detailed investigation into specific combinations of conditions associated with adverse health outcomes to identify the most vulnerable subgroups of older individuals and their health care needs. Such an inquiry would also advance patient-centered outcomes research in important ways,[4] by moving away from a paradigm based on the identification of an index disease.[5,6]

The study of multimorbidity faces numerous challenges, including the lack of a uniform definition.[5,7–9] Given the ubiquitous use of administrative data, however, multimorbidity has often been equated with multiple chronic conditions, especially when these conditions are identified from claims data.[10–12]

In a previous study,[13] we argued that multimorbidity should be redefined to include the co-occurrence of chronic conditions, functional limitations, and geriatric syndromes. We demonstrated the strong cumulative effects of their co-occurrence on adverse health outcomes, including self-report of fair/poor health, self-rated worse health at 2 years, and 2-year mortality.

In this study, we expand on the above-referenced study, by identifying specific combinations of conditions within and across the broad rubrics of chronic conditions, functional limitations, and geriatric syndromes that are particularly relevant in predicting poor health outcomes. We address important limitations noted in previous studies on multimorbidity:

First, rather than relying on administrative claims data, which are not optimal when gauging functional limitations and geriatric syndromes, we use data from the U.S. Health and Retirement Study (HRS). The HRS includes a rich array of longitudinal data that can be used to characterize, in detail, both multimorbidity and health outcomes.

Second, given the inadequacies presented by traditional statistical methods to study combinations of numerous conditions, we rely on a tree-based analytic approach—the Classification and Regression Tree analysis (CART), a modern version of the original CART by Breiman et al.,[14] described in detail below. CART allows for an exhaustive search of all possibilities by identifying splitting variables, and groups individuals in empirically emerging homogeneous, but mutually exclusive categories relative to a given outcome of interest.[15]

We hypothesize that the combinations of conditions that are relevant in predicting health outcomes include functional limitations and/or geriatric syndromes, rather than chronic conditions only.

METHODS

Overview

This study uses the 2010–2012 HRS public use files to identify specific combinations of chronic conditions, functional limitations, and geriatric syndromes that are most prominent in predicting the following health outcomes: current fair/poor health status, self-rated worse health at 2 years, and 2-year mortality.

This study was approved by the Case Western Reserve University Institutional Review Board.

Data Source and Study Population

Sponsored by the National Institute on Aging, the HRS is an ongoing, nationally representative longitudinal survey of approximately 30,000 older adults. The HRS collects biennial data on a broad array of measures, including demographics; chronic conditions; self-reported health status; functional status (e.g., upper/lower body limitations, activities of daily living (ADL) limitations); and geriatric syndromes (e.g., depressive symptoms, and incontinence).

The 2010 wave of the HRS contains data on 19,924 living respondents 50 years of age or older, including data on 1382 proxy respondents (typically the spouse) who provided answers for the focal respondents. In the 2012 wave, 2279 respondents were lost to follow-up (not interviewed and not deceased). We excluded 555 respondents with missing information on relevant variables. Those excluded were sicker than those who remained in the study population, as measured by the percentage of individuals who rated their health as fair/poor (38.8 % vs. 28.8 %, $p<0.01$), for example.

The study population included 16,640 respondents.

Variables of Interest

Outcomes. We created the following three binary outcomes:

Self-rated fair or poor health status in 2010 (vs. excellent, very good or good).

Self-rated worse health, coded as '1' if the respondent rated his/her health as "worse" (vs. "better" or "about the same") in 2012 compared to when s/he was interviewed in 2010.

Deceased status in 2012, usually reported by a spouse or other relative or as identified through a probabilistic National Death Index match.

Primary Predictors. Chronic Conditions. Chronic conditions were identified by whether the respondent was ever told by a physician that s(he) had hypertension, heart disease (including subtype of the disease—e.g., myocardial infarction or congestive heart failure), lung disease, diabetes, stroke, arthritis, cancer, or psychiatric conditions. Additional questions assessed whether the respondent was being treated for a given condition (e.g., medication for hypertension, or oxygen for lung disease), and/or that a condition caused limitations in usual activities. For each condition, we categorized respondents as no disease (=0), self-reported diagnosis as mild (=1), and active treatment/limitations as severe (=2). We also identified whether a respondent was ever told by a physician that she or he had Alzheimer's disease or other dementia.

For functional performance, the respondent was asked to assess whether s(he) has difficulty with a specific task, using validated questions.[16] *Functional limitations* were grouped in the following categories: a) Strength limitations (e.g., pulling/pushing a large object; lifting ten pounds); b) upper body limitations (e.g., picking up a dime; reaching overhead); c) lower body limitations (e.g., walking one block; climbing stairs); d) Activities of Daily Living (ADL: e.g., bathing; dressing); and Instrumental Activities of Daily Living (IADL: e.g., preparing meals, managing money). For parsimony, some of our analyses incorporated broader, rather than more detailed, categories constituting functional limitations (e.g., IADL, instead of each of the items contained in that rubric).

Geriatric Syndromes. Geriatric syndromes are conditions commonly experienced by older individuals,[17] including: a) visual impairment (self-rated as fair or poor when wearing corrective lenses as usual, or legally blind); b) hearing impairment (self-rated as fair or poor when wearing hearing aid as usual); c) moderate or severe depressive symptoms (≥ 4 symptoms on the eight-item Center for Epidemiologic Studies Depression Scale [CES-D]); d) urinary incontinence; e) low cognitive performance (< 11, based on the 35-point Telephone Interview Cognitive Status scale),[18] or a proxy reporting that the respondent's cognitive performance was fair or poor; f) persistent dizziness or lightheadedness; and g) severe pain ("often troubled by"). Conditions for which we observed relatively

high percentages of missing values (e.g., depressive symptoms), we created a category of 'missing.'

Other Variables of Interest. Sociodemographic variables: Age was grouped in 5-year increments. All respondents older than 85 years of age were grouped in the ≥ 85 category to protect privacy. Race/Ethnicity included four categories: White non-Hispanic, Black non-Hispanic, Hispanic, and Other. Marital status was identified as Married, Divorced, Widowed, or Never Married. Years of education were grouped in six categories: < 9, 9–11, 12, 13–15, 16, and ≥ 17. Income adjusted for household size was expressed as the ratio of household income to the federal poverty level, as follows: < 100, 100–199, 200–299, and ≥ 300 %. Smoking status was grouped in three categories (never smoked, former smoker, and current smoker), as was alcohol use (none, moderate, heavy and missing). We characterized body mass index (BMI, measured as kg/m^2) as underweight (BMI ≤ 18.5), normal/overweight (BMI of 18.5–30), obese (BMI ≥ 30) or missing (2 % of respondents). Vigorous exercise was measured as a dichotomous variable for engagement in vigorous sports or activities more than once a week. Self-rated fair or poor health status in 2010 was also used as predictor for prospective health outcomes.

Statistical Analyses

We used modern generalizations of the classification and regression tree (CART) analysis to identify combinations of predictors associated with the outcomes of interest. CART is a nonparametric, machine-learning method that recursively partitions the values of each of the predictor variables into two sets so that the values of the outcome variable are as homogeneous as possible in each set.[14] Each predictor variable is considered as a potential split, including every grouped value of a numerical predictor. The optimal split is the one with the largest reduction in an impurity index, which measures the extent of misclassification at a given node. We used the following stopping criteria (based on model-tuning described below): a maximum tree depth of five splits, a minimum node size of 50 subjects (25 for mortality outcome), requiring a split to increase the complexity parameter by a minimum of 0.001, and using the information impurity index to determine node splits.

To build our model, we partitioned the study data into training data sets (70 % of the data) and validation data sets (remaining 30 %) using random sampling within each class of the outcome variables. We used ten-fold cross-validation repeated three times on the training data set testing to build the CART models. Since mortality is highly unbalanced, we weighted the "cost" of a false negative to be higher than a false positive in order to improve sensitivity and produce a more meaningful model. We then tested the accuracy of our models on the validation data set. We chose our final model for each outcome based on accuracy and interpretability.

Additionally, we compared our final models with those from a random forest model to see if they were in agreement on variables that are the most important predictors (data not shown). Random forest is a bootstrap aggregation method that creates multiple decision trees using random variable selection. Breiman et al. provide a detailed description of random forest.[19] We used SAS for data management; for our statistical analyses, we used R version 3.1.3 and the 'rpart' (CART), partykit (tree graphics), 'randomForest' (random forest), and 'caret' (model tuning and cross-validation) packages.

RESULTS

Table 1 presents the distribution of sociodemographic characteristics and behavioral factors for the total study population and by self-reported health, self-rated worse health at 2 years, and 2-year mortality. Twenty-nine percent of respondents reported fair/poor health, 23 % reported 2-year decline, and 5 % died by the end of the 2-year follow-up period. These proportions varied considerably across subgroups of the population. Self-reported fair/poor health status was highest among the oldest respondents (39 %); those of Hispanic origin (48 %); persons with low incomes or educational attainment (53 and 57 %, respectively); current smokers (40 %); those who are underweight (48 %); and among those with proxy respondents (54 %). Self-rated worse health at 2 years followed a pattern similar to that of self-reported fair/poor status, and 2-year mortality was markedly higher among those who were older, widows, underweight; and those with proxy respondents.

Table 2 presents the percent of respondents with chronic conditions, functional limitations, and geriatric syndromes. The percent of respondents presenting with our outcomes of interest varied markedly by the presence of these conditions. We observed the highest levels of fair/poor self-reported health among respondents with severe lung disease (75 %); limitations in ADLs or IADLs (76 and 65 %, respectively); severe pain (77 %), depressive symptoms (62 %), and visual impairment (61 %).

The percent of respondents with self-rated worse health at 2 years was highest among those with severe lung disease (42 %); limitations in ADLs and IADLs (40 and 37 %); severe pain (44 %) and depressive symptoms (40 %).

Two-year mortality was highest among respondents with Alzheimer's disease (35 %), limitations in ADLs and IADLs (24 and 16 %, respectively), and poor cognitive functioning (27 %).

Figures 1, 2, and 3 show the trees obtained from our Classification and Regression Tree (CART) analysis for each of the three outcomes: fair/poor health status, self-rated worse health at 2 years, and 2-year mortality. See the Appendix for the more detailed analysis.

Figure 1 shows that the combination of specific conditions associated with the highest percentage of fair/poor health

Table 1 Sociodemographic Characteristics and the Prevalence of the Outcomes of Interest Across Subgroups of the Study Population

	Total N (column %)	Fair/poor self-rated health N (row %)	Self-rated worse health at 2 years N (row %)	2-year mortality N (row %)
No. with outcome	16,640	4797	3825	832
Age categories				
50–54	2969 (18 %)	829 (28 %)	518 (17 %)	35 (1 %)
55–59	3067 (18 %)	930 (30 %)	596 (19 %)	53 (2 %)
60–64	2293 (14 %)	643 (28 %)	482 (21 %)	36 (2 %)
65–69	1828 (11 %)	409 (22 %)	387 (21 %)	53 (3 %)
70–74	2224 (13 %)	588 (26 %)	561 (25 %)	111 (5 %)
75–79	1823 (11 %)	548 (30 %)	542 (30 %)	117 (6 %)
80–84	1247 (7 %)	382 (31 %)	397 (32 %)	150 (12 %)
≥ 85	1189 (7 %)	468 (39 %)	342 (29 %)	277 (23 %)
Sex				
Male	7063 (42 %)	1953 (28 %)	1545 (22 %)	400 (6 %)
Female	9577 (58 %)	2844 (30 %)	2280 (24 %)	432 (5 %)
Race/ethnicity				
White non-Hispanic	11,020 (66 %)	2490 (23 %)	2659 (24 %)	607 (6 %)
Black non-Hispanic	3105 (19 %)	1191 (38 %)	590 (19 %)	164 (5 %)
Hispanic	2038 (12 %)	981 (48 %)	461 (23 %)	53 (3 %)
Other	477 (3 %)	135 (28 %)	115 (24 %)	8 (2 %)
Marital status				
Married	10,275 (62 %)	2553 (25 %)	2196 (21 %)	394 (4 %)
Divorced	2516 (15 %)	899 (36 %)	626 (25 %)	96 (4 %)
Widowed	2847 (17 %)	996 (35 %)	797 (28 %)	308 (11 %)
Never married	1002 (6 %)	349 (35 %)	206 (21 %)	34 (3 %)
Education, years				
< 9	1537 (9 %)	880 (57 %)	420 (27 %)	121 (8 %)
9–11	1917 (12 %)	860 (45 %)	515 (27 %)	159 (8 %)
12	5258 (32 %)	1516 (29 %)	1271 (24 %)	273 (5 %)
13–15	3950 (24 %)	970 (25 %)	868 (22 %)	159 (4 %)
16	2060 (12 %)	328 (16 %)	392 (19 %)	69 (3 %)
≥ 17	1918 (12 %)	243 (13 %)	359 (19 %)	51 (3 %)
Income as % of federal poverty level				
< 100 %	2076 (12 %)	1104 (53 %)	577 (28 %)	147 (7 %)
100–199 %	3219 (19 %)	1353 (42 %)	929 (29 %)	244 (8 %)
200–299 %	2755 (17 %)	834 (30 %)	709 (26 %)	160 (6 %)
≥ 300 %	8590 (52 %)	1506 (18 %)	1610 (19 %)	281 (3 %)
Smoking status				
Never smoked	7282 (44 %)	1796 (25 %)	1510 (21 %)	277 (4 %)
Former smoker	6903 (41 %)	2015 (29 %)	1634 (24 %)	442 (6 %)
Current smoker	2455 (15 %)	986 (40 %)	681 (28 %)	113 (5 %)
Alcohol use				
None	10,440 (63 %)	3574 (34 %)	2648 (25 %)	656 (6 %)
Moderate	4593 (28 %)	767 (17 %)	864 (19 %)	133 (3 %)
Heavy	1575 (9 %)	*	*	*
Missing	32 (< 0.1%)	*	*	*
Body mass index				
Missing	291 (2 %)	128 (44 %)	58 (20 %)	21 (7 %)
Underweight	222 (1 %)	106 (48 %)	55 (25 %)	51 (23 %)
Normal/overweight	10,424 (63 %)	2558 (25 %)	2319 (22 %)	571 (5 %)
Obese	5703 (34 %)	2005 (35 %)	1393 (24 %)	189 (3 %)
Vigorous exercise				
No	12,610 (76 %)	4279 (34 %)	3189 (25 %)	767 (6 %)
Yes	4030 (24 %)	518 (13 %)	636 (16 %)	65 (2 %)
Proxy respondent				
Yes	783 (5 %)	419 (54 %)	219 (28 %)	194 (25 %)
No	15,857 (95 %)	4378 (28 %)	3606 (23 %)	638 (4 %)

* denotes cells with counts < 11. In accordance with privacy rules, these cells, as well as cells in corresponding columns were masked

consists of difficulty walking several blocks, depressive symptoms, and severe pain (nodes 1, 22, and 28, leading to the second bar from the right). Indeed, over 80 % of respondents with this combination of conditions reported fair/poor health. Conversely, the lowest prevalence of fair/poor health was observed among respondents with no difficulty walking several blocks, no impaired vision, but reporting incomes at or exceeding 300 % of the FPL (nodes 1, 2, and 3, leading to a prevalence of fair/poor health of less than 10 %).

Figure 2 shows that the highest percentage of self-rated worse health at 2 years (about 60 %) was

observed among respondents who had difficulty walking several blocks, who self-rated their health as fair/poor, and who were 68.5 years of age or older (nodes 1, 23, and 31). When difficulty walking several blocks co-occurred with reporting fair/poor health in individuals who are younger than 68.5 years of age, it was the presence of depressive symptoms that placed them at a high risk of reporting self-rated worse health, at nearly 50 % (nodes 1, 23, 31, and 32). In contrast, the lowest percentage of self-rated worse health (less than 10 %) was observed among those who had no difficulty

Table 2 Prevalence of the Outcomes of Interest in the Presence of Chronic Conditions, Functional Limitations, and Geriatric Syndromes

	Total *N* (column %)	Fair/poor self-rated health *N* (row %)	Self-rated worse health at 2 years *N* (row %)	2-year mortality *N* (row %)
No. with outcome	16,640	4797	3825	832
Chronic conditions				
Hypertension (mild)	1447 (9 %)	519 (36 %)	377 (26 %)	129 (9 %)
Hypertension (severe)	8762 (53 %)	3131 (36 %)	2276 (26 %)	526 (6 %)
Arthritis (mild)	4771 (29 %)	983 (21 %)	1076 (23 %)	262 (5 %)
Arthritis (Severe)	4450 (27 %)	2278 (51 %)	1627 (37 %)	304 (7 %)
Heart Disease (mild)	3328 (20 %)	1331 (40 %)	1049 (32 %)	364 (11 %)
Heart Disease (severe)	1004 (6 %)	567 (56 %)	357 (36 %)	87 (9 %)
Heart disease subconditions:				
Angina	724 (4 %)	410 (57 %)	274 (38 %)	77 (11 %)
Cardiac arrhythmia	1777 (11 %)	851 (48 %)	614 (35 %)	187 (11 %)
Congestive heart failure	827 (5 %)	559 (68 %)	310 (37 %)	154 (19 %)
Myocardial infarction	1338 (8 %)	701 (52 %)	440 (33 %)	157 (12 %)
Lung disease (mild)	1262 (8 %)	610 (48 %)	495 (39 %)	91 (7 %)
Lung disease (severe)	335 (2 %)	251 (75 %)	140 (42 %)	88 (26 %)
Stroke (mild)	1111 (7 %)	517 (47 %)	342 (31 %)	181 (16 %)
Stroke (severe)	564 (3 %)	364 (65 %)	202 (36 %)	83 (15 %)
Cancer (mild)	1845 (11 %)	628 (34 %)	524 (28 %)	166 (9 %)
Cancer (severe)	751 (5 %)	350 (47 %)	200 (27 %)	111 (15 %)
Diabetes (mild)	3022 (18 %)	1252 (41 %)	832 (28 %)	181 (6 %)
Diabetes (severe)	873 (5 %)	560 (64 %)	291 (33 %)	103 (12 %)
Psychiatric conditions (mild)	1035 (6 %)	445 (43 %)	336 (32 %)	63 (6 %)
Psychiatric conditions (severe)	2346 (14 %)	1191 (51 %)	851 (36 %)	165 (7 %)
Alzheimer's	231 (1 %)	144 (62 %)	76 (33 %)	81 (35 %)
Other dementia	299 (2 %)	202 (68 %)	114 (38 %)	74 (25 %)
Functional limitations				
Strength limitations	9472 (57 %)	3943 (42 %)	2843 (30 %)	689 (7 %)
Upper body mobility limitations	5825 (35 %)	3085 (53 %)	2036 (35 %)	581 (10 %)
Lower body mobility limitations	9800 (59 %)	4102 (42 %)	2977 (30 %)	743 (8 %)
Limitations in ADLs	1069 (6 %)	817 (76 %)	425 (40 %)	257 (24 %)
Limitations in IADLs	2795 (17 %)	1828 (65 %)	1026 (37 %)	444 (16 %)
Geriatric syndromes				
Poor cognitive functioning	888 (5 %)	519 (58 %)	271 (31 %)	244 (27 %)
Depressive symptoms	2409 (14 %)	1503 (62 %)	966 (40 %)	141 (6 %)
Incontinence	3985 (24 %)	1581 (40 %)	1218 (31 %)	338 (8 %)
Severe pain	1142 (7 %)	877 (77 %)	502 (44 %)	97 (8 %)
Visual impairment	3935 (24 %)	2400 (61 %)	1248 (32 %)	310 (8 %)
Hearing impairment	3431 (21 %)	1711 (50 %)	1067 (31 %)	303 (9 %)

walking several blocks, rated their health as Excellent, Very Good, or Good, and had no arthritis (nodes 1, 2, and 3, leading to the first bar at the left hand side).

Figure 3 shows that the most important element in predicting mortality is age, which is split at 80.5 years. Thus, the highest percentage of 2-year mortality (slightly over 40 %) was observed among those who were 80.5 years of age or older, and who had limitations in both ADLs and IADLs (nodes 1, 19, and 27). The second highest percentage (at slightly lower than 40 %) was observed among respondents who had limitations in IADLs, difficulty pulling/pushing a large object, and were 84.5 years of age or older (nodes 1, 19, 28, and 32, leading to the second bar from the right).

Both Figs. 1 and 2 identified difficulty walking several blocks as the most important variable predicting both self-reported fair/poor health and 2-year worse health. The models seldom identified chronic conditions as splitting variables at critical junctures when predicting any of our three outcomes.

DISCUSSION

In this study, we analyzed health outcomes (self-reported fair/poor health, self-rated worse health at 2 years, and 2-year mortality) in a U.S. representative sample of older adults, and identified combinations of conditions that are most relevant in predicting these outcomes. The most remarkable finding was that, rather than chronic conditions, the combination of functional limitations and geriatric syndromes was most important in predicting health outcomes. Indeed, when all of these conditions were accounted for, and when CART conducted an exhaustive search of all the possibilities,[15] the model determined the most important "splitting" variables to be age, functional limitations, and geriatric syndromes, rather than chronic conditions. This is akin to the disablement process described by Verbrugge and Jette.[20]

These findings have important implications for research and clinical practice. Methodologically, these findings imply that accounting for chronic conditions alone may not be as informative as accounting for the co-occurrence of functional lim-
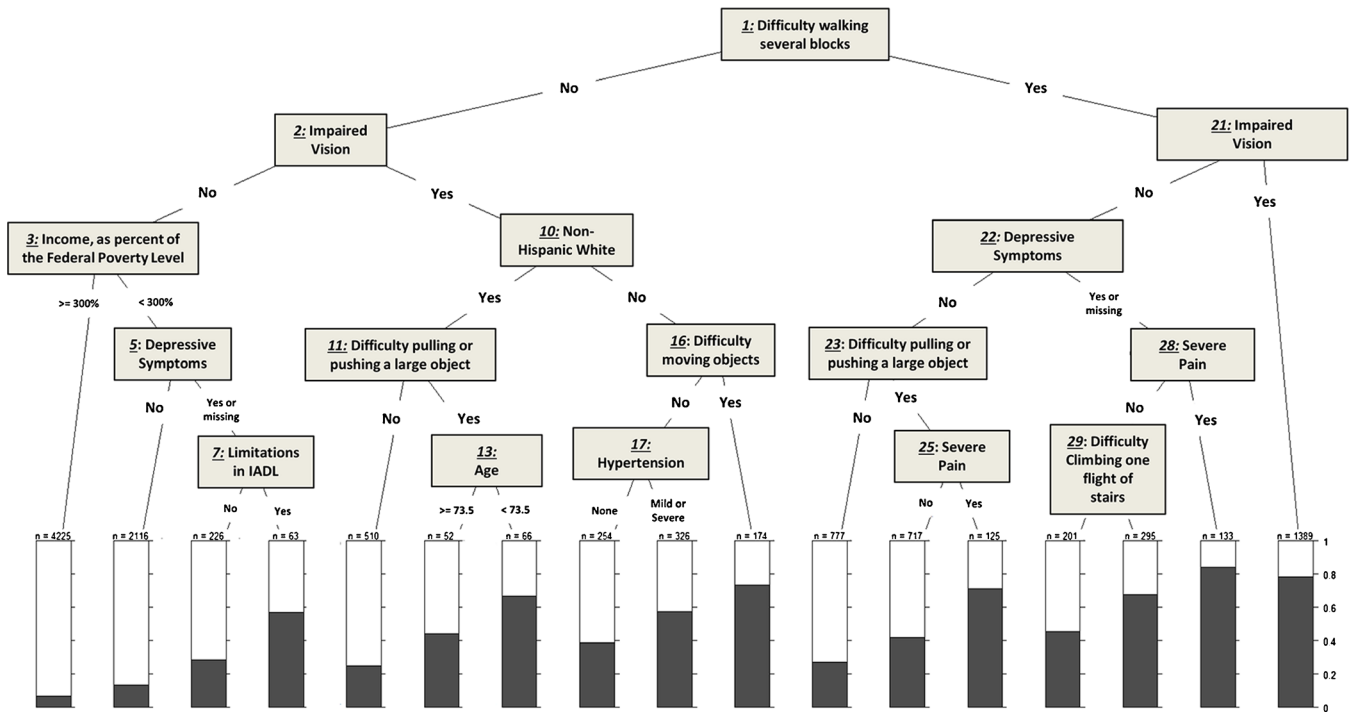
Figure 1 Classification and regression tree analysis for self-reported fair/poor health. IADL: Instrumental activities of daily living.

itations and geriatric syndromes, which appear to drive health outcomes in older individuals. However, research in this realm has relied on International Coding of Diseases 9th Clinical Modification (ICD-9-CM) diagnosis codes documented in administrative data, which present a challenge for identifying disability[21] and geriatric syndromes. As a result, multimorbidity has often been equated to *multiple chronic conditions*, and the importance of functional limitations and

geriatric syndromes in predicting health outcomes has seldom been examined in relation to health outcomes.

With regard to clinical practice, the findings call for the use of such instruments as the Comprehensive Geriatric Assessment to evaluate the health status and clinical presentation of older individuals in a way that makes it possible to identify subgroups of older adults that are most vulnerable to experience adverse outcomes,[22–25] especially when making clinical
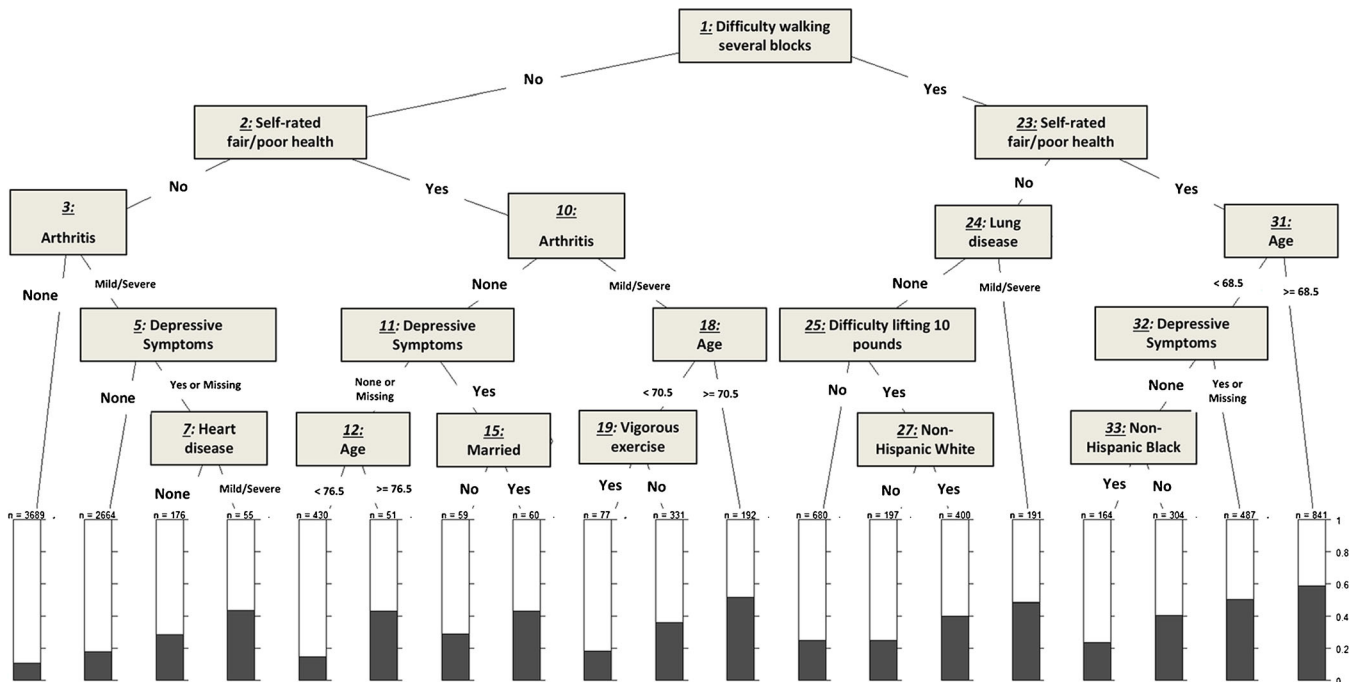

Figure 2 Classification and regression tree analysis for self-rated worse health at 2 years.
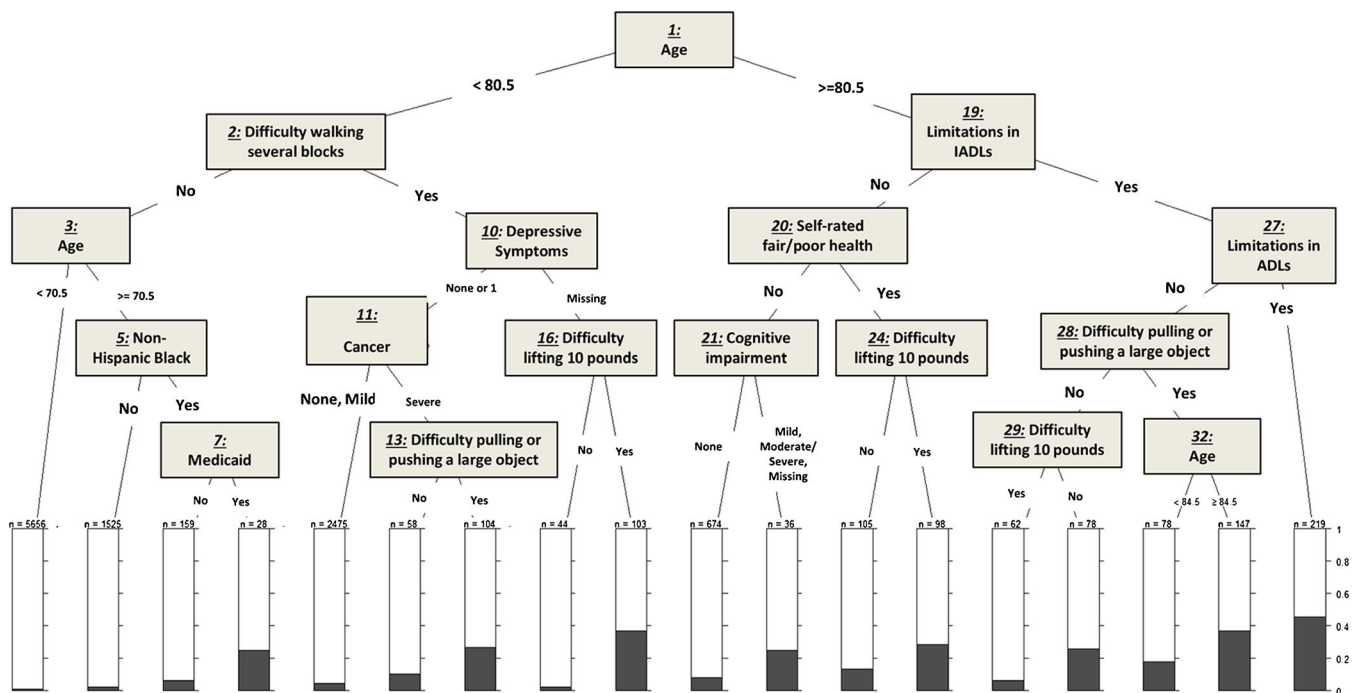
Figure 3 Classification and regression tree analysis for 2-year mortality. IADL: Instrumental activities of daily living.

decisions. For example, our findings document the importance of visual impairment in predicting health outcomes, and yet this condition has been mostly ignored in the literature. As such, this analysis constitutes an important step towards a more personalized approach when characterizing patients with multimorbidity.

To our knowledge, this is the first study to identify combinations of specific conditions that are associated with adverse health outcomes using modern CART. The main advantage of using this analytic approach lies in our ability to learn about *empirically emerging combinations of conditions* instead of having our queries guided solely by a priori hypotheses. In addition, an important strength of this study consists of characterizing multimorbidity based on a rich array of nationally representative data from the HRS, including sociodemographic variables, behavioral factors, and a number of clinically pertinent measures, which we were able to characterize by level of severity. Finally, we note that rather than just including those 65 years of age and older, our study population included those 50–64 years of age, given that a sizable percentage (18.2 %) present with co-occurring chronic conditions, functional limitations, and geriatric syndromes.[13] This compares to 33.3 % in their older counterparts.[13] In addition, our age-stratified analysis yielded CART trees that also underscored the importance of functional limitations and geriatric syndromes (rather than chronic conditions) driving health outcomes.

This study has several limitations: First, all of the conditions accounted for are self- or proxy-reported, and the chronic conditions examined are limited to the ten conditions reported in the HRS. However, our access to the linked HRS-Medicare data allowed us to repeat our analysis using claims-derived chronic conditions as per the Chronic Conditions Data Warehouse (CCW) that are flagged in the Medicare Beneficiary Summary File. The CART trees obtained using the CCW conditions instead of the ten self-reported chronic conditions did not change markedly, even though this change in analytic strategy also required that we limit our study population to Medicare fee-for-service beneficiaries. Second, while the CART uses the entire dataset, it produces a single tree. This is in contrast with the random forest approach, which uses subsets of the data, as well as bootstrapping, to identify the most important predictors. On the other hand, it is not possible to use the random forest approach to identify combinations of conditions, which was the primary objective of this study. Third, because of the randomness in splitting data into particular training and testing subsets involved with building trees, our figures may look somewhat different if they are regenerated by others. However, we conducted thorough sensitivity analysis; and the trees presented herein are the ones that produced the best statistics relative to their accuracy and predictive abilities (see Online supplemental data). Finally, we note that our study population was exclusive of 555 respondents who differed significantly from others (e.g., in self-reported health status). However, given that they represented only 3 % of the population, it is unlikely that their exclusion affected our results, except possibly by yielding more conservative estimates.

In conclusion, this study highlighted the importance of characterizing multimorbidity in broader terms, rather than

limiting its definition to the co-occurrence of chronic conditions alone. Future studies should analyze additional outcomes, including health care utilization and costs.

Corresponding Author: *Siran M. Koroukian, Ph.D.; Department of Epidemiology and Biostatistics, School of MedicineCase Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-4945, USA (e-mail: skoroukian@case.edu).*

Compliance with Ethical Standards:

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*

REFERENCES

1. **Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK.** Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. Prev Chronic Dis. 2013;10:E66.
2. **Lochner KA, Shoff CM.** County-level variation in prevalence of multiple chronic conditions among Medicare beneficiaries, 2012. Prev Chronic Dis. 2015;12:E07.
3. U.S. Department of Health and Human Services. Multiple chronic conditions—a strategic framework: optimum health and quality of life for individuals with multiple chronic conditions. Washington, DC. December 2010. http://www.hhs.gov/ash/initiatives/mcc/mcc_framework.pdf, Accessed December 15, 2015.
4. **Parekh AK, Kronick R, Tavenner M.** Optimizing health for persons with multiple chronic conditions. JAMA. 2014;312(12):1199–200.
5. **Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M.** Defining comorbidity: implications for understanding health and health services. Ann Fam Med. 2009;7(4):357–63.
6. **Bayliss EA, Edwards AE, Steiner JF, Main DS.** Processes of care desired by elderly patients with multimorbidities. Fam Pract. 2008;25(4):287–93.
7. **Fortin M, Soubhi H, Hudon C, Bayliss EA, van den Akker M.** Multimorbidity's many challenges. BMJ. 2007;334(7602):1016–7.
8. **Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H.** A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. Ann Fam Med. 2012;10(2):142–51.
9. **Fortin M, Hudon C, Haggerty J, Akker M, Almirall J.** Prevalence estimates of multimorbidity: a comparative study of two sources. BMC Health Serv Res. 2010;10:111.
10. **Kone Pefoyo AJ, Bronskill SE, Gruneir A, et al.** The increasing burden and complexity of multimorbidity. BMC Public Health. 2015;15(1):415.
11. **Melis R, Marengoni A, Angleman S, Fratiglioni L.** Incidence and predictors of multimorbidity in the elderly: a population-based longitudinal study. PLoS One. 2014;9(7). e103120.
12. **Violan C, Foguet-Boreu Q, Roso-Llorach A, et al.** Burden of multimorbidity, socioeconomic status and use of health services across stages of life in urban areas: a cross-sectional study. BMC Public Health. 2014;14:530.
13. **Koroukian SM, Warner DF, Owusu C, Given CW.** Multimorbidity redefined: prospective health outcomes and the cumulative effect of co-occurring conditions. Prev Chronic Dis. 2015;12:E55.
14. **Breiman L.** Classification and Regression Trees. Belmont: Wadsworth International Group; 1984.
15. **Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W.** Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Ann Behav Med. 2003;26(3):172–81.
16. **Fonda A, Herzog AR.** HRS/AHEAD Documentation Report. Documentation of Physical Functioning Measured in the Health and Retirement Study and the Asset and Health Dynamics Among the Oldest Old Study. Ann Arbor, MI: Survey Research Center, University of Michigan; 2004. **Accessed October 31, 2015**.
17. **Inouye SK, Studenski S, Tinetti ME, Kuchel GA.** Geriatric syndromes: clinical, research, and policy implications of a core geriatric concept. J Am Geriatr Soc. 2007;55(5):780–91.
18. **Brandt J, Folstein SE, Folstein MF.** Differential cognitive impairment in Alzheimer's disease and Huntington's disease. Ann Neurol. 1988;23(6):555–61.
19. **Breiman LCA, Liuw A, et al.** Breiman and Cutler's random forests for classification and regression. http://cran.r-project.org/web/packages/randomForest/index.html, Accessed December 15, 2015.
20. **Verbrugge LM, Jette AM.** The disablement process. Soc Sci Med. 1994;38(1):1–14.
21. **Iezzoni LI.** Using administrative data to study persons with disabilities. Milbank Q. 2002;80(2):347–79.
22. **Meneses K, Benz R, Azuero A, Jablonski-Jaudon R, McNees P.** Multimorbidity and breast cancer. Semin Oncol Nurs. 2015;31(2):163–9.
23. **Deschodt M, Devriendt E, Sabbe M, et al.** Characteristics of older adults admitted to the emergency department (ED) and their risk factors for ED readmission based on comprehensive geriatric assessment: a prospective cohort study. BMC Geriatr. 2015;15(1):54.
24. **Stotter A, Reed MW, Gray LJ, Moore N, Robinson TG.** Comprehensive geriatric assessment and predicted 3-year survival in treatment planning for frail patients with early breast cancer. Br J Surg. 2015;102(5):525–33. **discussion 533**.
25. **Carlson C, Merel SE, Yukawa M.** Geriatric syndromes and geriatric assessment for the generalist. Med Clin N Am. 2015;99(2):263–79.