


METHODS PAPER

Development of a specialty intensity score to estimate a patient's need for care coordination across physician specialties

Ashley Hodgson  | Thomas Bernardin | Benjamin Westermeyer | Ella Hagopian | Tyler Radtke | Ahmed Noman

St. Olaf College, Northfield, Minnesota

Correspondence

Ashley Hodgson, St. Olaf College, 1520 St. Olaf Ave, Northfield, MN 55306, USA.
Email: hodgsona@stolaf.edu

Funding information

Collaborative Undergraduate Research and Inquiry program at St. Olaf College.; Frank Gary Endowment of St. Olaf College; St. Olaf College; Centers for Medicare and Medicaid Services

Abstract

Backgrounds and aims: This article develops a Specialty Intensity Score, which uses patient diagnosis codes to estimate the number of specialist physicians a patient will need to access. Conceptually, the score can serve as a proxy for a patient's need for care coordination across doctors. Such a measure may be valuable to researchers studying care coordination practices for complex patients. In contrast with previous comorbidity scores, which focus primarily on mortality and utilization, this comorbidity score approximates the complexity of a patient's the interaction with the health care system.

Methods: We use 2015 inpatient claims data from the Centers for Medicare and Medicaid Services to model the relationship between a patient's diagnoses and physician specialty usage. We estimate usage of specialist doctors by using a least absolute shrinkage and selection operator Poisson model. The Specialty Intensity Score is then constructed using this predicted specialty usage. To validate our score, we test its power to predict the occurrence of patient safety incidents and compare that with the predictive power of the Charlson comorbidity index.

Results: Our model uses 127 of the 279 International Classification of Disease, 10th Revision, Clinical Modification (ICD-10-CM) diagnosis subchapters to predict specialty usage, thus creating the Specialty Intensity Score. This score has significantly greater power in predicting patient safety complications than the widely used Charlson comorbidity index.

Conclusion: The Specialty Intensity Score developed in this article can be used by health services researchers and administrators to approximate a patient's need for care coordination across multiple specialist doctors. It, therefore, can help with evaluation of care coordination practices by allowing researchers to restrict their analysis of outcomes to the patients most impacted by those practices.

KEYWORDS

care coordination, comorbidity score, inpatients, medical specialization, patient safety incidents

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Health Science Reports* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Health services researchers and policy-makers have increasingly acknowledged the importance of improving care coordination for patients with multiple chronic conditions.^{1,2} Multiple chronic condition patients spend 65% of all health care dollars and 95% of Medicare dollars.³ Yet, these patients receive substandard care,⁴ often due to complications of the coordination process.⁵ One definition of care coordination is “the deliberate organization of patient care activities between two or more participants (including the patient) involved in a patient's care to facilitate the appropriate delivery of health care services.”⁶ The complexity of these patient's interaction with the health care system may put the patient's health at greater risk, particularly if different doctors treating them provide conflicting care plans or fail to communicate well with each other or with the patient.

For these reasons, care coordination has become a focus among policy-makers and health care administrators alike. The majority hospitals employ a variety of care management tools and techniques such as case managers, predictive analytic tools, checklists, visit summaries, conversations prompts in the medical records, and more.⁷ Insurers incentivize good care coordination practices through bundled payment and pay-for-performance systems, although evidence is mixed about the effectiveness of these at improving care coordination.⁸⁻¹¹ The Affordable Care Act included incentives for the development of Accountable Care Organizations in part as a way of promoting innovation and accountability for care coordination practices.¹² And not just the United States, but countries across the globe have highlighted care coordination as a goal.¹³⁻¹⁶

Recognizing the importance of developing measurements of care coordination, the Agency for Healthcare Research and Quality undertook a project called the Care Coordination Measures Atlas, which summarized the available literature and identified 64 existing instruments for measuring the quality of care coordination.¹⁷ This project was most recently updated in June 2014, when it introduced a section on care coordination measures that can be constructed from electronic health records (EHR). Measurements that rely on EHR have numerous advantages, including the limited data collection burden and ease of aggregating across broader populations.¹⁸ While the project uncovered 26 measures of care coordination developed from EHR data, all of them required more specific information beyond simply the International Classification of Disease, 10th Revision, Clinical Modification (ICD-10-CM) codes that make up the bulk of many widely available data sets. Since the release of the Care Coordination Measures Atlas, a 2018 Veterans Affairs conference focused on care coordination, and published a report identifying gaps in the available measures of care coordination.¹⁹ This report highlighted a continued need to identify which patients could benefit most from care coordination practices.

Our project seeks to fill that gap identified by Agency for Healthcare Research and Quality and the 2018 Veterans Affairs coordination conference by providing a measure of a patient's need for care coordination that can be used by researchers with access to widely available datasets with ICD-10-CM patient diagnostic codes. Specialty Intensity Score aims to allow researchers to identify which patients are most likely to require care from multiple medical specialists. The score serves as an estimate for a

patient's need for care coordination and can be combined with coordination-sensitive outcome measures to assess the quality of care coordination within and across health care facilities. By separating out patients likely to need the most care coordination, researchers can study coordination-sensitive outcomes with greater precision. Coordination-sensitive outcomes may include death rate among patients with low-mortality diagnoses, hospital-acquired infections, and hospital readmissions.

To check for advancements in the field since the 2014 Agency for Healthcare Research and Quality report on care coordination instruments, we conducted a structured literature review in search of more recent measures of care coordination, described in the appendix. The process uncovered a number of process-oriented and survey-based care coordination measures.²⁰⁻²² However, of these new measures, only one aims to estimate a patient's need for care coordination, the Care Coordination Tier Assessment Tool.^{23,24} However, this instrument requires individual assessment by a person reviewing a patient's case, and this person must be able to account for duration of their conditions and the care team available to them. This would not be possible for researchers who only have access to large databases of hospital discharges or insurer claims. Our Specialty Intensity Score is intended for use by health services researchers who work with widely available datasets, like the Agency for Healthcare Research and Quality National Inpatient Sample database or state-level hospital discharge datasets.

Because the Specialty Intensity Score is a version of a comorbidity index, we wanted to compare it to existing such measures, particularly those constructed using similar data. We therefore conducted a structured literature review in search of existing comorbidity instruments. Our findings appear in Table 1, and the process for identifying these measures is described in an appendix. All measures in the table are instruments that aggregate patient comorbidities into scores that predict patient outcome or usage patterns. None of the existing comorbidity scores was proxies for the amount of coordination patients would need, and none used number of unique physician specialties as the response variable of interest. Most of the scores were constructed using mortality or prognosis as the main response variable, with various patient bases and techniques for score design. The most commonly cited such score is the Charlson Comorbidity Index.²⁷ A number of the scores were utilization based, but measured utilization by cost, hospitalization, readmission, or resource use, rather than by number of medical specialties involved in a patient's care. No existing comorbidity instrument was explicitly designed to estimate the complexity of a patient's interaction with the health care system, like Specialty Intensity Score that we develop in this article.

In the sections that follow, we describe the process for developing the Specialty Intensity Score, by using ICD-10-CM codes to estimate the number of doctors with unique specialties a patient sees. We then test whether the Specialty Intensity Score is empirically different from the commonly used Charlson Comorbidity Index by quantifying the predictive value of each in estimating a patient's probability of experiencing a patient safety incident in the hospital. Our results show that the two scores are sufficiently distinct, and that the Specialty Intensity Score has higher predictive power for safety incidents.

TABLE 1 Comorbidity scores in the literature

| Comorbidity Score or Index | Year of Publication | Pharmacy-Based Data? | Response Variable(s) of Interest | Target Patient Base |
|---|---------------------|----------------------|--|--|
| Cumulative Illness Rating Scale ²⁵ | 1968 | No | Degree of impairment | General |
| Kaplan/Feinstein index ²⁶ | 1974 | No | Mortality | Patients with diabetes mellitus |
| Charlson comorbidity index ²⁷ | 1987 | No | Mortality | General |
| Diagnostic cost groups ²⁸ | 1988 | No | Cost | General |
| Chronic disease score ²⁹ | 1992 | Yes | Physician ratings of physical disease severity, patient-rated health status, mortality and hospitalization | |
| Ambulatory care groups ³⁰ | 1992 | No | Ambulatory care resource use | General |
| Number of prescribed medications ²⁹ | 1992 | Yes | Mortality, hospitalization, physician-rated disease severity, patient-rated health status | High utilizers of ambulatory health care |
| Satariano index ³¹ | 1994 | No | Survival | Breast cancer patients |
| Index of co-existent diseases ³² | 1995 | No | Postoperative complications and 1-year health-related quality of life | Patients undergoing total hip replacement |
| Total illness burden index ³³ | 1995 | No | Functional status outcomes | General |
| Elixhauser comorbidity score ³⁴ | 1998 | No | Mortality | General |
| Silliman score ³⁵ | 1999 | No | Mortality, physical function | Breast cancer patients |
| Klabunde outpatient and inpatient indices ³⁶ | 2000 | No | Mortality | Breast cancer and prostate cancer patients |
| Geriatric index of comorbidity ³⁷ | 2002 | No | Cognitive status, depressive symptoms, functional status, somatic health | Elderly patients |
| Frailty index ³⁸ | 2002 | No | Frailty, ability to react to stress | General |
| Washington university head and neck comorbidity index ³⁹ | 2002 | No | Prognosis | Patients with head and neck squamous cell carcinoma |
| RxRisk score ⁴⁰ | 2003 | Yes | Cost | Veterans health administration population |
| Medication regimen complexity index ⁴¹ | 2004 | Yes | Outcomes measures | Patients with moderate to severe chronic obstructive pulmonary disease |
| Multipurpose Australian comorbidity scoring system ⁴² | 2005 | No | Mortality, hospital readmission | General |
| Simplified comorbidity index ⁴³ | 2005 | No | Prognosis | General |
| Rheumatic disease comorbidity index ⁴⁴ | 2007 | No | Mortality, hospitalization | Patients with rheumatoid arthritis |
| Pharmacy-based comorbidity index ⁴⁵ | 2013 | Yes | Hospitalization | General |
| CirCom score ⁴⁶ | 2014 | No | Mortality | Cirrhosis patients' |
| Multimorbidity index ⁴⁷ | 2015 | No | Health-related quality of life | Patients with rheumatoid arthritis |
| Drug derived complexity index ⁴⁸ | 2016 | Yes | Death, hospitalization, readmission | General |
| Medicines comorbidity index ⁴⁹ | 2017 | Yes | Death, hospitalization | General |
| Specialty intensity score | | No | Number of doctors with unique specialties | General |

2 | METHODS

2.1 | Data

The study utilized data from the Center for Medicare and Medicaid Services (CMS) Limited Data Set of Standard Analytical Files. This

included Medicare Part A and Part B medical claims for a 5% random sample of Medicare beneficiaries. The specific data utilized in the project come from claims from the Inpatient and Carrier files for the year of 2015. The inpatient file contained ICD-9-CM and ICD-10-CM diagnosis codes, which indicate the health conditions associated with patients' inpatient visits. The inpatient file also

contained National Provider Identifier (NPI) codes for the attending, operating, and other physicians who provided services to patients during their inpatient visits. The carrier file contained NPI codes for referring and performing physicians. Specialties of the referring, performing, attending, operating, and other physicians were determined by mapping their NPI codes to their provider taxonomy codes, using data from the National Plan and Provider Enumeration System.

The overarching goal of this article is to use patient diagnoses to estimate the number of specialties that will be required during a hospital visit and then to use this specialty estimate as a score of the patient's need for medical coordination (the SI score). We aim to validate the score by testing its success at predicting the occurrence of health complications. This analysis requires three types of data: (a) a list of diagnosis categories, (b) a measure of physician medical specialty usage, and (c) a measure of inpatient medical complications. The following three sections explain how we used 2015 CMS inpatient data to create these three types of variables. All analyses were conducted using R statistical software. Before undertaking work with this deidentified administrative data, all members of the research team underwent training in research ethics through the Collaborative Institutional Training Initiative and reviewed institutional policies for work with data, as required by the Institutional Review Board at our institution. We ensured that our research design and processes met the protocol for working with patient administrative data, which required protection of data but did not require direct consent from patients in the data set.

TABLE 2 Patient safety indicators

| |
|--|
| PSI 02: Death in low-mortality diagnosis-related groups |
| PSI 03: Pressure ulcer |
| PSI 06: Iatrogenic pneumothorax |
| PSI 07: Central venous catheter-related blood stream infection |
| PSI 08: In hospital fall with hip fracture |
| PSI 09: Perioperative hemorrhage or hematoma |
| PSI 10: Postoperative acute kidney injury requiring dialysis |
| PSI 11: Postoperative respiratory failure |
| PSI 12: Perioperative pulmonary embolism or deep vein thrombosis |
| PSI 13: Postoperative sepsis |
| PSI 14: Postoperative wound dehiscence |

TABLE 3 Descriptive statistics

| | Lower Bound Specialty Count | Upper Bound Specialty Count | Patient Safety Event Indicator | Charlson Comorbidity Index |
|----------|-----------------------------|-----------------------------|--------------------------------|----------------------------|
| Mean | 4.04 | 4.37 | 0.002398 | 2.55 |
| Median | 4 | 4 | 0 | 2 |
| Variance | 3.10 | 3.78 | 0.002393 | 4.60 |
| Minimum | 0 | 0 | 0 | 0 |
| Maximum | 15 | 16 | 1 | 19 |

Note: N = 277 264.

2.1.1 | Clustering diagnosis categories

With 14 500 ICD-9-CM diagnosis codes and 70 000 ICD-10-CM codes,⁵⁰ it is neither practical nor insightful to use individual diagnosis codes as the explanatory variables in a model of specialty usage. Many codes are nearly identical to each other, so treating each code as a distinct covariate would cost a large amount of computational power for very little explanatory gain. Therefore, it is useful to group the individual diagnosis codes into clinically meaningful categories and to use these categories as the explanatory variables. For this article, we experimented with three different diagnosis categorizations. We ultimately chose to use the World Health Organization subchapter categorization, since it was constructed based on medical classification rather than cost considerations. See the appendix for the other classifications we considered.

2.1.2 | Specialty usage

The Centers for Medicare and Medicaid Services inpatient claims data include information on patients' utilization of referring, performing, attending, operating, and other physicians. A patient's usage of a physician is recorded using the physician's unique National Provider Identifier (NPI) code. Because each physician is allowed to list more than one specialty, and we do not have a way of knowing which specialty a patient saw that physician for, we devised several different methods for estimating the number of physician specialties that each patient utilized. We used an upper-bound and lower-bound approach and ran robustness checks across the different methods, which are described in more detail in an appendix. Table 3 gives descriptive statistics about

TABLE 4 LASSO-penalized Poisson model summaries

| | Lower Bound Specialty Count | Upper Bound Specialty Count |
|--|-----------------------------|-----------------------------|
| Number of predictors | 127 | 95 |
| Lambda | 0.00992 | 0.01982 |
| Root mean squared error | 1.548 | 1.696 |
| Root mean squared error/standard deviation | 0.879 | 0.872 |

TABLE 5 Logistic regression model summaries

| | Patient Safety Event ~ SI Score 1 | Patient Safety Event ~ SI Score 1 + Charlson | Patient Safety Event ~ SI Score 2 | Patient Safety Event ~ SI Score 2 + Charlson |
|-------------------|-----------------------------------|--|-----------------------------------|--|
| Intercept | -9.85*** | -9.76*** | -9.75*** | -9.66*** |
| SI score | 0.87*** | 0.92*** | 0.79*** | 0.83*** |
| Charlson | - | -0.11*** | - | -0.11*** |
| Residual deviance | 4567.9 | 4548.3 | 4582.7 | 4562.9 |
| AIC | 4571.9 | 4554.3 | 4586.7 | 4568.9 |
| VIF | - | 1.09625 | - | 1.10069 |

Note: * $P < .05$, ** $P < .01$, *** $P < .001$.

the upper and lower bound of doctor specialties that a patient utilized. Table 4 compares results under both the lower-bound and upper-bound estimates of physician specialties utilized.

2.1.3 | Inpatient complications

We utilized the Agency for Healthcare Research and Quality's Patient Safety Indicators⁵¹ to estimate the number of medical complications each patient experienced while in the hospital. Table 2 gives the list of Patient Safety Indicators that we included in the study. Table 3 provides descriptive statistics on the prevalence of patient safety events, as observed in our data.

2.2 | Methodological approach

To predict the average number of specialist doctors utilized with the diagnosis subchapter indicator variables, the first stage of our analysis employed a least absolute shrinkage and selection operator (LASSO) penalized Poisson model (Table 4). This approach is appropriate and valuable in our context for several reasons. First, the Poisson model assumes that the response variable is count data, as opposed to binary or continuous. Our response variable, the number of specialties seen in an inpatient visit, is a count variable. The Poisson model also assumes that the mean and variance of the response are roughly equal. Therefore, the Poisson generalized linear model is appropriate for our specialty count variables. Second, using a shrinkage method is appropriate because our data are very wide: there are 279 predictors. Generalized linear models tend to have problems with multicollinearity when there are such a large number of predictors. LASSO combats this problem by both shrinking the size of predictor coefficients and performing subset selection.

In the second stage of our analysis, we evaluate the validity of the SI score by testing its power in predicting the occurrence of patient safety events in a logistic regression. We use residual deviance and the Akaike Information Criterion as metrics of the explanatory power of our model. After estimating the simple logistic regression, we introduce the Charlson index as another covariate in

the logistic regression, for comparison against the SI Score. We use likelihood ratio tests to determine if the Charlson explains any additional variation in patient safety events and variance inflation factor (VIF) tests to determine if the Charlson and SI score are significantly collinear. We carry out this entire process twice: once for the SI score created using the lower bound specialty count variable and once for the SI score created using the upper bound.

3 | RESULTS

Our analysis yielded two results. The first is the generation of the Specialty Intensity Score, which is a measure that we designed for health services researchers to utilize in their own studies, particularly studies that involve multiple chronic condition patients. The second result of our study estimates the power of the Specialty Intensity Score for predicting the occurrence of a negative patient safety event during a hospital stay, and the comparison of that score to the existing Charlson Comorbidity Score.

3.1 | Predicting specialty usage with diagnosis subchapters

Researchers can create SI Scores for patients in their own data using Table C1 from our appendix. To calculate the SI Score for a given patient, we exponentiate the sum of the intercept and the coefficients of the patient's diagnosis subchapters. For example, if there is a patient with diagnoses in subchapters 59, 111, and 116, then we exponentiate the sum of 1.0826 (the intercept), 0.0422 (the Sub59 coefficient), 0.0043 (the Sub111 coefficient), and 0.0390 (the Sub116 coefficient). As a result of this calculation, we estimate that this type of patient will use an average of 3.216 specialties. This is the patient's SI score using the lower bound specialty count variable.

$$SIS = e^{1.0826 + 0.0422 \text{ Sub59} + 0.0043 \text{ Sub111} + 0.0390 \text{ Sub116}} = 3.2160. \quad (1)$$

The lower bound specialty count model achieves a root mean squared error of 1.548 on the holdout data. This means that if we only

count physicians' primary specialty in our calculation of patient specialty usage, our model's predictions of specialty usage are an average of 1.548 away from the patient's actual usage. For reference, the SD of SpecCount1 variable in the holdout data is 1.761. Because our model's root mean squared error is smaller than the SD by a factor of 0.879, we can conclude that our model is superior to the naive estimator.

3.2 | Predicting patient safety event occurrence with predicted specialty usage

Although both the SI score and Charlson index are significant predictors of patient safety events, the models show that the SI score is the more important one, both in terms of coefficient size and deviance reduction (Table 5). This was expected; the Patient Safety Indicators measure iatrogenic hospital complications, which are problems with the hospital's care, not problems with the patient's body. It is intuitive that these complications would be more strongly associated with a score of care coordination need, such as the SI score, than a score of biological frailty, such as the Charlson.

Whether we used the lower or upper bound of specialty usage, the logistic regressions demonstrated that the SI score and Charlson index are significant, unique covariates of patient safety event occurrence. This shows that our analysis is robust to the inclusion of physicians' secondary specialty in the specialty count calculation. That being said, we recommend that researchers use SI Score 1 over SI Score 2. This choice makes intuitive and empirical sense for several reasons. First, patients are more likely to use just a physician's primary specialty than they are to use both their primary and secondary specialties during a visit. Thus, the likelihood that we are overcounting by including both physician specialties in a calculation of total patient specialty usage is greater than the likelihood that we are undercounting by disregarding secondary specialties. Second, because we consider the ability of the SI score to explain patient safety events to be the SI score's validity as a measure of care coordination need, we want to choose the SI score that has the strongest relationship with patient safety event occurrence. The logistic models with SI Score 1 achieve lower deviances and Akaike Information Criteria than the models with SI Score 2, both with and without the inclusion of the Charlson index. Thus, SI Score 1 is empirically a better proxy for patient need for care coordination. For these reasons, we recommend using SI Score 1, and we include the coefficient table for the SI Score 1 model in the appendix.

To determine if the Charlson explains a significant amount of new variation in the probability of patient safety events when introduced into these logistic regressions, we use variance inflation factors (VIFs) and likelihood ratio tests (LRTs). The VIF helps us understand if there is an excessive amount of multicollinearity between the SI score and Charlson variables. In other words, it indicates if the SI score and Charlson index are distinct scores. If the VIF is very high, then there is no need for researchers to account for both scores because they are very similar. However, if the VIF is low, then the two indexes

are unique. In the logistic regressions with the Charlson, the VIF ranges from 1.096 with SI Score 1 to 1.101 with SI Score 2, both of which are well below the rule of thumb threshold of 5. We conclude that the SI score and Charlson index are distinct scores, each offering unique information about inpatient visits. The variance inflation factors (VIFs) can show that multicollinearity is an issue, but they cannot show the usefulness of the SI score and Charlson index in explaining patient safety event occurrence. For this, we use LRTs. These tests compare the residual deviance of the models without the Charlson to the models including it. If adding the Charlson to the logistic regression significantly reduces the residual deviance, then we can conclude that adding the Charlson explains a significant amount of additional variation in the probability of a patient safety event. For the SI Score 1 model, the difference in deviance is 19.59 ($P < 0.001$). For the SI Score 2 model, the deviance difference is 19.83 ($P < 0.001$). Thus, we can conclude that the Charlson does contribute significant additional explanatory power to models predicting patient safety events.

4 | LIMITATIONS

Our use of the Centers for Medicare and Medicaid Services claims dataset limits the study for several reasons. All patients in our dataset are Medicare patients, who are older and sicker than the general population. It is possible that optimal care coordination for Medicare patients may be systematically different than optimal care coordination for younger patients. It is possible that older adults are less likely to see a specialist for additional conditions compared with younger people, particularly if some of their diagnoses are long-standing problems that they have dealt with separately in the past. On the other hand, it may also be possible that this population is more likely to see a specialist compared to younger patients based on the fact that they are more likely to be insured. Therefore, our estimates could be biased in either direction relative to the U.S. population as a whole. The fact that our data came from 2015 may also make it less valid as time goes on. As medical guidelines change over time, doctors may develop different practices for recommending when a patient should consult with a new specialist. Also, changes in telemedicine and coordination practices resulting from the COVID-19 pandemic may dramatically alter the average number of specialists that are considered standard for patients to see, rendering our scoring system outdated.

There are several assumptions that we had to make when constructing our measure. First, we assume that patients saw a doctor from every medical specialty that they needed expertise from. This assumption could be violated by good care coordination practices that some organizations engage in. For example, if a patient's primary care doctor made phone calls to specialists to get advice for that patient, then that patient would not appear to have consulted these doctors in our dataset, and we would underestimate the number of areas of medical expertise they needed. Conversely, we assumed that patients needed special medical expertise from every doctor that they saw. It is possible that some of these doctors were checking in on the patients for routine medical check-ups, unrelated to their specialties.

Finally, we were unable to observe which specialty the patient was utilizing. Doctors could list up to two specialties and we had no way of knowing which of the two specialties was relevant to the patient. To account for this, we devised a maximum estimate of doctors seen and a minimum estimate and compared these approaches, as described in previous sections.

5 | DISCUSSION

In this article, we have developed the Specialty Intensity Score, which uses ICD-10-CM codes to score a patient's need for medical expertise drawn from different, unique medical specialties. We have done this by using clusters of ICD-10-CM codes to predict the number of specialist doctors from unique specialties that a patient sees during a hospital visit. To test the validity of our Specialty Intensity Score, we checked whether our score had power in predicting a patient's probability of experiencing a patient safety event, and whether that predictive power was independent of the Charlson comorbidity index. Our measure has significantly more predictive power in estimating a patient's probability of a patient safety event compared to the existing Charlson comorbidity index. This finding validates the need for our score as an independent source of meaning for researchers exploring issues relating to complex hospital patients with many comorbidities.

The reason our Specialty Intensity Score adds something unique over existing measures of comorbidity listed in Table 1. Scores developed using mortality, prognosis, functionality, or patient health theoretically capture a patient's biological vulnerability. Scores using utilization, such as resource use, cost, and hospitalization, capture the intensity of investment, but not the complexity of care. Our score, on the other hand, captures an important feature of the nature of the patient's interaction with the health care system: how many doctors are likely to be needed for their care. Therefore, our measure is uniquely positioned for studying care coordination, when combined with coordination-sensitive outcomes.

Patients needing to draw medical knowledge from a broader range of medical specialties will either need to see many different specialist doctors or else find some alternative way of getting personalized medical information from those specialties. For example, one alternative approach would be for the patient to see a primary care doctor or hospitalist and for that hospitalist to consult separately with specialist doctors as needed. Another model would be for a team of specialists to each meet with the patient and also meet separately as a team with one another, as has sometimes been practiced at the Mayo Clinic.⁵² It might also be possible for electronic medical records to facilitate coordination across a team of practitioners⁵³ or a non-MD patient advocate to assist in the coordination process.⁵⁴ Some researchers have proposed the use of systems engineering to facilitate and design optimal transfer of medical information across multiple disciplines of expertise.^{55,56} A systems engineer, for example, might put together checklists for transfer of information between patients or might determine which specialist doctor should be seen first or which specialist's advice should override another's when there is a conflicting recommendation.

The Specialty Intensity Score is designed to be used in other studies of both efficiency and clinical effectiveness of different care coordination practices. By identifying which patients are the most likely to need knowledge from a broader spectrum of medical specialties, our score can allow researchers to identify the patients most likely to be impacted by good or bad coordination practices and to look at the effects of particular practices on those patients separately from patients with less need for coordination.

FUNDING

This work was supported by St. Olaf College, by the Collaborative Undergraduate Research and Inquiry program at St. Olaf College and by the Frank Gary Endowment of St. Olaf College. The funding sources had no involvement in the study design; collection, analysis, and interpretation of data; writing of the report; or the decision to submit the report for publication.

AUTHOR CONTRIBUTIONS

Conceptualization and Supervision: Ashley Hodgson and Thomas Bernardin

Funding acquisition and rResources: Thomas Bernardin

Formal Analysis and Investigation: Ashley Hodgson, Thomas Bernardin, Benjamin Westermeyer, Ahmed Noman, Ella Hagopian, Tyler Radtke

Thomas Bernardin, Benjamin Westermeyer, Ahmed Noman, Ella Hagopian, Tyler Radtke

Methodology and Visualization: Thomas Bernardin, Benjamin Westermeyer, Ahmed Noman, Ella Hagopian, Tyler Radtke

Data curation and Software: Benjamin Westermeyer, Ahmed Noman, Ella Hagopian, Tyler Radtke

Project administration and Supervision: Ashley Hodgson and Thomas Bernardin

Writing—Original Draft Preparation: Ashley Hodgson and Benjamin Westermeyer

Writing—Review and Editing: Ashley Hodgson, Benjamin Westermeyer and Tyler Radtke

All authors have read and approved the final version of the manuscript.

Ashley Hodgson had full access to all of the data in this study and takes complete responsibility for the integrity of the data and the accuracy of the data analysis.

TRANSPARENCY STATEMENT

Ashley Hodgson affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the CMS [https://www.cms.gov/Research-Statistics-Data-and-Systems/

Files-for-Order/LimitedDataSets]. Restrictions apply to the availability of these data, which were used under license for this study.

CONFLICT OF INTEREST STATEMENT

The authors of this article claim no conflicts of interest.

ORCID

Ashley Hodgson  <https://orcid.org/0000-0001-7434-6774>

REFERENCES

- Care Coordination Measures Atlas, Updated June 2014. 385.
- Busse R, Blümel M. Payment systems to improve quality, efficiency and care coordination for chronically ill patients—a framework and country examples. “Triple Aim” Future Health Care Madr Span Sav Banks Found FUNCAS. 2015;
- Parekh AK, Barton MB. The challenge of multiple comorbidity for the US health care system. *JAMA*. 2010;303(13):1303-1304.
- Parekh AK, Goodman RA, Gordon C, Koh HK, Conditions HIW o MC. Managing multiple chronic conditions: a strategic framework for improving health outcomes and quality of life. *Public Health Rep*. 2011;126(4):460-471.
- Beckmann U, Gillies DM, Berenholtz SM, Wu AW, Pronovost P. Incidents relating to the intra-hospital transfer of critically ill patients. *Intensive Care Med*. 2004;30(8):1579-1585.
- McDonald KM, Sundaram V, Bravata DM, Lewis R, Lin N, Kraft SA, et al. Closing the quality gap: a critical analysis of quality improvement strategies (Vol. 7: Care Coordination). 2007
- Hearld LR, Carroll N, Hall A. The adoption and spread of hospital care coordination activities under value-based programs. *Am J Manag Care*. 2019;25(8):397-404.
- de Bakker DH, Struijs JN, Baan CA, et al. Early results from adoption of bundled payment for diabetes care in The Netherlands show improvement in care coordination. *Health Aff (Millwood)*. 2012;31(2):426-433.
- Baicker K, Levy H. Coordination versus competition in health care reform. *N Engl J Med*. 2013;369(9):789-791.
- Fagan PJ, Schuster AB, Boyd C, et al. Chronic care improvement in primary care: evaluation of an integrated pay-for-performance and practice-based care coordination program among elderly patients with diabetes. *Health Serv Res*. 2010;45(6p1):1763-1782.
- Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res*. 2010;10(1):247.
- Andrews CM, Darnell JS, McBride TD, Gehlert S. Social work and implementation of the affordable care act. *Health Soc Work*. 2013;38(2):67-71.
- Tsiachristas A, Dikkers C, Boland MRS, Rutten-van Mülken MPMH. Exploring payment schemes used to promote integrated chronic care in Europe. *Health Policy*. 2013;113(3):296-304.
- Kringos D, Boerma W, Bourgueil Y, et al. The strength of primary care in Europe: an international comparative study. *Br J Gen Pract*. 2013; 63(616):e742-e750.
- Hannigan B, Simpson A, Coffey M, Barlow S, Jones A. Care coordination as imagined, care coordination as done: findings from a cross-national mental health systems study. *Int J Integr Care* [Internet]. 2021;18(3):1-14. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6137622/>.
- Penm J, MacKinnon NJ, Strakowski SM, Ying J, Doty MM. Minding the gap: factors associated with primary care coordination of adults in 11 countries. *Ann Fam Med*. 2017;15(2):113-119.
- McDonald K, Schultz E, Albin L, Pineda N, Lonhart J, Sundaram V, et al. Coordination Atlas Version 4. AHRQ Publ. 2014; (14-0037-EF).
- Update on the adoption of health information technology and related efforts to facilitate the electronic use and exchange of health information: A report to congress. Off Natl Coord Health Inf Technol. 2013; 82.
- Mattocks KM, Cunningham K, Elwy AR, et al. Recommendations for the evaluation of cross-system care coordination from the VA state-of-the-art working group on VA/non-VA care. *J Gen Intern Med*. 2019;34(1):18-23.
- Risco E, Sauch G, Alberio A, et al. Spanish validation of the “user reported measure of care coordination” questionnaire for older people with complex, chronic conditions. *Int J Environ Res Public Health*. 2020;17(18):6608.
- Zlateva I, Anderson D, Coman E, Khatri K, Tian T, Fifield J. Development and validation of the medical home care coordination survey for assessing care coordination in the primary care setting from the patient and provider perspectives. *BMC Health Serv Res*. 2015;15(1):1-11.
- Quinn M, Robinson C, Forman J, Krein SL, Rosland A-M. Survey instruments to assess patient experiences with access and coordination across healthcare settings: available and needed measures. *Med Care*. 2017;55(Suppl 7 1):S84.
- Lemke M, Kappel R, McCarter R, D'Angelo L, Tuchman LK. Perceptions of health care transition care coordination in patients with chronic illness. *Pediatrics*. 2018;141(5):1-8.
- Care Coordination Tier Assignment Tool, Version 1.0 Health Care Home Initiative [Internet]; 2021. Available from: <https://www.health.state.mn.us/facilities/hchomes/payment/training/complexteirtool.pdf>
- Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc*. 1968;16(5):622-626.
- Kaplan MH, Feinstein AR. The importance of classifying initial comorbidity in evaluating the outcome of diabetes mellitus. *J Chronic Dis*. 1974;27(7):387-404.
- Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383.
- Ellis RP, Ash A. Refining the Diagnostic Cost Group Model: A Proposed Modification to the AAPCC for HMO Reimbursement. Brandeis Health Policy Research Consortium, Boston University; 1988.
- Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol*. 1992;45(2):197-203.
- Weiner JP, Starfield BH, Lieberman RN. Johns Hopkins ambulatory care groups (ACGs). A case-mix system for UR, QA and capitation adjustment. *HMO Pract*. 1992;6(1):13-19.
- Satariano WA, Ragland DR. The effect of comorbidity on 3-year survival of women with primary breast cancer. *Ann Intern Med*. 1994;120(2):104-110.
- Greenfield S, Apolone G, McNeil BJ, Cleary PD. The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. *Med Care*. 1993;31(2):141-154.
- Greenfield S, Sullivan L, Dukes KA, Silliman R, D'Agostino R, Kaplan SH. Development and testing of a new measure of case mix for use in office practice. *Med Care*. 1995;33(4):AS47-AS55.
- Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27.
- Silliman RA, Lash TL. Comparison of interview-based and medical-record based indices of comorbidity among breast cancer patients. *Med Care*. 1999;37(4):339-349.
- Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol*. 2000; 53(12):1258-1267.
- Rozzini R, Frisoni GB, Ferrucci L, et al. Geriatric index of comorbidity: validation and comparison with other measures of comorbidity. *Age Ageing*. 2002;31(4):277-285.

38. Rockwood K, Mitnitski AB, MacKnight C. Some mathematical models of frailty and their clinical implications. *Rev Clin Gerontol*. 2002;12(2): 109-117.
39. Piccirillo JF, Lacy PD, Basu A, Spitznagel EL. Development of a new head and neck cancer-specific comorbidity index. *Arch Otolaryngol Neck Surg*. 2002;128(10):1172-1179.
40. Sloan KL, Sales AE, Liu C-F, et al. Construction and characteristics of the RxRisk-V: a VA-adapted pharmacy-based case-mix instrument. *Med Care*. 2003;41(6):761-774.
41. George J, Phun Y-T, Bailey MJ, Kong DC, Stewart K. Development and validation of the medication regimen complexity index. *Ann Pharmacother*. 2004;38(9):1369-1376.
42. Holman CDJ, Preen DB, Baynham NJ, Finn JC, Semmens JB. A multi-purpose comorbidity scoring system performed better than the Charlson index. *J Clin Epidemiol*. 2005;58(10):1006-1014.
43. Colinet B, Jacot W, Bertrand D, et al. A new simplified comorbidity score as a prognostic factor in non-small-cell lung cancer patients: description and comparison with the Charlson's index. *Br J Cancer*. 2005;93(10):1098-1105.
44. Michaud K, Wolfe F. Comorbidities in rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2007;21(5):885-906.
45. Dong Y-H, Chang C-H, Shau W-Y, Kuo RN, Lai M-S, Chan KA. Development and validation of a pharmacy-based comorbidity measure in a population-based automated health care database. *Pharmacother J Hum Pharmacol Drug Ther*. 2013;33(2):126-136.
46. Jepsen P. Comorbidity in cirrhosis. *World J Gastroenterol WJG*. 2014; 20(23):7223-7230.
47. Radner H, Yoshida K, Mjaavatten MD, et al. Development of a multimorbidity index: impact on quality of life using a rheumatoid arthritis cohort. *Semin Arthritis Rheum*. 2015;45(2):167-173.
48. Robusto F, Lepore V, D'Ettorre A, et al. The drug derived complexity index (DDCI) predicts mortality, unplanned hospitalization and hospital readmissions at the population level. *Plos One*. 2016;11(2):e0149203.
49. Narayan SW, Nishtala PS. Development and validation of a medicines comorbidity index for older people. *Eur J Clin Pharmacol*. 2017;73(12): 1665-1672.
50. 2015 ICD-10-CM and GEMs | CMS [Internet]. General Equivalence Mappings (GEMs) – Diagnosis Codes and Guide (ZIP); 2021. Available from: <https://www.cms.gov/Medicare/Coding/ICD10/2015-ICD-10-CM-and-GEMs>
51. Agency for Healthcare and Research Quality - Patient Quality Indicators. PSI Overview; n.d. https://www.qualityindicators.ahrq.gov/modules/psi_overview.aspx. Accessed August 27, 2019.
52. Christensen CM, Grossman JH, Hwang J. The innovator's prescription. *Disruptive Solut For*. 2010;
53. Amir O, Grosz BJ, Gajos KZ, Swenson SM, Sanders LM. From care plans to care coordination: opportunities for computer support of teamwork in complex healthcare. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*; 2015. p. 1419-1428.
54. Schwartz L. Is there an advocate in the house? The role of health care professionals in patient advocacy. *J Med Ethics*. 2002;28(1):37-40.
55. Bisantz AM, Carayon P, Miller A, Khunlerkit A, Arbaje A, Xiao Y. Using human factors and systems engineering to improve care coordination. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA; 2012. p. 855-859.
56. Das P, Benneyan J, Powers L, Carmody M, Kerwin J, Singer S. Engineering safer care coordination from hospital to home: lessons from the USA. *Future Healthc J*. 2018;5(3):164.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Hodgson A, Bernardin T, Westermeyer B, Hagopian E, Radtke T, Noman A. Development of a specialty intensity score to estimate a patient's need for care coordination across physician specialties. *Health Sci Rep*. 2021;4:e303. <https://doi.org/10.1002/hsr2.303>