View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Vanderbilt Electronic Thesis and Dissertation Archive

EVALUATION OF AN ELECTRONIC HEALTH RECORD ADAPTATION OF A 4-YEAR MORTALITY RISK INDEX

By

Eugenia R McPeek Hinz

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Biomedical Informatics

August, 2012

Nashville, Tennessee

Approved:

Professor Joshua C. Denny MD MS

Professor Michael Matheny MD MS MPH

Professor Nancy M Lorenzi PhD MLS MA

Professor Josh F Peterson MD MPH

Professor William M Gregg MD MS

Copyright © 2012 by Eugenia McPeek Hinz MD
All Rights Reserved

ACKNOWLEDGEMENTS

First I would like to thank my committee for their guidance and support in these projects. Everyone added their own unique perspectives that proved invaluable to me and this work. My advisor, Josh Denny in particular has provided me continual encouragement, support and freedom to pursue this research from the first research question to the end results contained herein. The other members of my committee, Michael Matheny, Josh Peterson, Bill Gregg and Nancy Lorenzi have provided outstanding support through our ongoing conversations about this work from study design inception to data collection and analysis. I also extend my thanks to Raquel Zink and Lisa Bastarache whose technical support and discussions on data integrity taught me much.

I also need to thank Anderson Spickard, III who encouraged and allowed me to participate in the project with the National Board of Medical Education and through which I was introduced to the many physicians and staff of the Vanderbilt Medical Group. The physicians in particular were understanding and accommodating with all aspects of these studies sharing their experiences and allowing me to interview their patients. Thanks are due as well to Connie Davis for her help in allowing us to perform the survey portion of the study as well as Cymethia Brown, Deborah Mitchell, Amira Zukanovic, Pat Jackson and the other medical assistants so agreeable to participating in the second portion of the study.

I would also like to thank the Departments of Biomedical Informatics, the faculty and staff has been greatly supportive in this tour through graduate education. If my husband would let me, I imagine, I would stay in school forever so much have I enjoyed learning new things. My student colleagues have also been sources of help and

encouragement; I could not have made it through without you all. Finally I would thank Rischelle Jenkins, Alayne Ford and Teresa Gillespie for their assistance these past few years. Rischelle in particular has always been a smiling face and very helpful for a multitude of items from my first day of interviewing to my scheduling of my defense. My funding through the National Library of Medicine training grant 5T15LM007450-11 has made all of this possible.

And lastly, I want to thank my wonderful husband and three lovely daughters who have always been supportive of me following my dreams and willing to take this journey through Nashville.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iii
LIST OF TABLES vi
LIST OF FIGURES vii
LIST OF ABBREVIATIONS viii
I. INTRODUCTION 1
II. USING AN ELECTRONICALLY ADAPTED MORTALITY RISK INDEX TO STRATIFY PRIMARY CARE PATIENTS 6
Authors 6
Introduction and Background 6
Methods 8
Results 14
Discussion 19
Conclusion 22
Acknowledgements 23
III. AMBULATORY EVALUATION OF A 4-YEAR MORTALITY RISK INDEX BY SURVEY AND ELECTRONIC HEALTH RECORD ADAPATION 24
Introduction 24
Methods 25
Results 30
Discussion 42
Conclusion 44
Acknowledgement 44
IV. SUMMARY 45
Summary of Findings 45
Limitations 46
Future Directions 48
V. APPENDIX A ICD9 for Lee Disease Determination 51
VI. APPENDIX B Charlson-Deyo ICD9 and Point Value Determinations 61
VII. APPENDIX C Survey 766
VIII.References 788

LIST OF TABLES

Table Page

1. Demographics and Mortality Risk Point Attribution 11
2. Physician Practice Profile by Gender and Risk Stratum 14
3. Demographics by Mortality Risk 16
4. Physician Demographics and Mortality Estimates of their patient panels 18
5. Demographics of Survey Population and Electronic Equivalent 31
6. Electronic Proxies Performance to Predict Survey Data 32
7. Discrepant cases for Cancer by Survey versus Electronic Adaptation 33
8. Discrepant cases for Chronic Lung Disease by Survey versus Electronic Adaptation 34
9. Discrepant cases for Diabetes Mellitus by Survey versus Electronic Adaptation 35
10. Discrepant cases for CHF by Survey versus Electronic Adaptation 36
11. Discrepant cases for Smoking by Survey versus Electronic Adaptation 37
12. Discrepant cases for ADLs by Survey versus Electronic Adaptation 38
13. Survey Positive Functional Status adjudicated as Negative for Impairment 39
14. Survey Positive Functional Status adjudicated as Positive for Impairment 39
15. Adjudicated Control Standard versus Electronic Version 40
16. Adjudicated Control Standard versus Survey Version 40
17. Spearman Correlation Coefficients 41

LIST OF FIGURES

Figure

1. Kaplan-Meier Failure By Risk Stratums 15

2. AUC for Lee versus Charlson-Deyo at Two Years 17

3. eLee Risk Points by Physician 19

4. Functionality Questions 28

LIST OF ABBREVIATIONS

Abbreviation	Definition
AUC	Area Under the Receiver Operating Characteristic Curve
BMI	Body Mass Index
CHF	Congestive Heart Failure
COPD	Chronic Obstructive Pulmonary Disease
DM	Diabetes Mellitus
EDW	Enterprise Data Warehouse
EHR	Electronic Health Record
eLee	Electronic EHR-Adaptation of the Lee 4-year Mortality Risk Index
ICD9	International Classification of Diseases Ninth Version
IDI	Integrated Discrimination Improvement
JCAHO	Joint Commission on Accreditation of Healthcare Organizations
Lee	original Lee 4-year Mortality Risk Index
MRN	Medical Record Number
PPV	Positive Predictive Value
PCP	Primary Care Physician
ROC	Receiver Operating Characteristic

CHAPTER I

INTRODUCTION

Quality care metrics provide a means to compare physicians across practices. These measures are based on scientific evidence and attempt to reflect national guidelines for standard of care or typical practice parameters for disease management. For example ambulatory care physicians are scored on the percentage of diabetic patients with good control as evidenced by a HgbA1c value less than 8 mg/dl. Unfortunately these goals lack methods to adjust the contribution of patient health status to physician quality metrics outcomes. Consequently, physicians generally give little credence to such data scoring results as inadequate measures of the quality of the care they are providing, often voicing the complaint that "my patients are sicker" as a cause for a lower quality metric than their colleagues.(1,2) That sentiment is supported by evidence of patient co-morbidities exerting a stronger influence on quality metric outcomes than the effect of the physician.(3,4)

Mortality risk and severity of illness scores have been used for decades as surrogates of patient disease burden to help risk adjust populations to analyze outcomes.(5–14) Measures developed for outpatient populations have also typically relied on subjective descriptions of health status with variable success.(8,10,15–18) The burden of data collection by measure of person time effort limits scalability of mortality indices for application across larger populations.(19) Electronic Health Records (EHR) are tremendous sources of patient data, and can secondarily be used to provide information on

physicians through the patients they see, notes they write, diagnoses they make, and the orders they generate. We hypothesized that EHR data could bridge the gap in defining components of an ambulatory mortality risk index in a systematic and replicable manner. There is a strong need for an automated implementation of an outpatient mortality risk prediction tool to support individual patient clinical decision support and provide risk adjustment for population management.

Defining disease burden for an ambulatory population is complex because of greater variability in health status and the typical overall better health of the average outpatient population compared to the inpatient one.(20) Limited duration risk indices such as the modified forms of the Charlson or Elixhauser, rely primarily on easily obtained administrative International Classification of Diseases Ninth Edition (ICD9) data and benefit from shorter term predictions in higher risk populations.(7,21) The longer duration mortality risk indices, such as the (Lee 4 year Mortality Risk Index) Lee and Schonberg 5 year mortality risk index, share common components of patient age, gender, BMI, risk behaviors, co-morbid diseases and functional status.(10,12,22) Functional status is unique to these measures and previously shown to predict mortality by itself.(10,11)

Lee et al. described a 4 year Mortality Risk Index for outpatient populations over 50 years of age that has been validated in two separate populations.(22) To calculate mortality risk, the Lee model uses a 12-item patient survey of age, gender, presence of 4 specific chronic diseases (Congestive Heart Failure (CHF), Chronic Lung Disease, Diabetes Mellitus (DM), Cancer), Body Mass Index (BMI), current smoking status and functional status to derive the risk score. Functional status represents up to one third of the total score, thus potentially contributing a significant portion to a patient's final mortality risk

estimate. The Lee index has been used to better delineate frailty and vulnerability in older populations.(23) Its application to a variable-aged outpatient population using EHR data has not been fully investigated.

The Lee index is unique from the other longer duration tools in that all of its measures are available either directly from or with closely related proxies in the EHR. Measurement of functional status has been a part of Vanderbilt's EHR via a clinical intake questionnaire administered to most outpatients as a part of Joint Commission on Accreditation of Healthcare Organizations (JCAHO) requirements in the form of a binary yes/no response. While limited in comparison to the four functional status questions in the Lee it provides a reasonable proxy to this data. The Schonberg 5-year mortality risk score is the next most similar tool to the Lee but includes a subjective patient health status question and asks the patient to provide the number of hospitalizations for the prior year. Hospitalization data is challenging to collect from a particular EHR system given discontinuity in care as patients often seek care at different hospitals given changing health insurance, geographic relocations, and personal preferences.

Overview of work presented in this thesis

This thesis describes two studies evaluating the utility, discrimination and validity of an outpatient mortality risk score using an electronic adaptation of the Lee 4-year mortality risk model.

In our initial study (**CHAPTER II**), we evaluated the accuracy of the Lee EHR-adaptation (eLee) to predict 2-year mortality. We implemented an EHR-adaptation of the

Lee in an outpatient primary care population using all patients seen by the fourteen study providers. We calculated patient mortality risk from the date of enrollment and followed patients for up to two years for the outcome of death. Using the total points assigned by the Lee and subsequent data of death at two years, discrimination of this model was determined by calculating the area under the receiver operator characteristic curve (AUC). We also applied a Cox proportional hazards model to three condensed stratums of mortality risk denoted as Low, Intermediate and High to represent clinically relevant and potentially actionable levels of risk. To evaluate differences in physician patient panels, we compared the aggregate percentage of high and intermediate risk patients in each physician's panel.

In the second study (**CHAPTER III**), we compared the eLee representing our EHR-adaptation to the original Lee survey for biases of information gathering and retrieval, and overall accuracy of survey and electronic data. To better match the questions of functional status ascertained by the Lee survey, we implemented an expanded form of the functional status questions into the EHR and trained intake personnel on use of the new tool. The new functionality consisted of an initial screening question of functional status that expanded to the four specific questions from the original Lee survey.

To evaluate the similarities and differences of the Lee and Schonberg indices, we surveyed adult patients in ambulatory primary care clinics with a combined survey instrument. Subsequently we applied the same eLee to the patients surveyed to directly compare the electronic equivalent score to the survey results. We completed chart reviews of discrepancies between the survey and electronic results seen due to unrelated biases with both methodologies to establish an adjudicated standard. Both the survey and the

electronic EHR-adaptation showed excellent correlation with the adjudicated standard with the electronic version.

CHAPTER IV discusses the overall conclusions from the two studies along with their limitations. We explore the potential future directions for applying this research in the clinical realms of individual level scores for personalized decision support and on the population level for risk adjustment for physician panels.

CHAPTER II

USING AN ELECTRONICALLY ADAPTED MORTALITY RISK INDEX

TO STRATIFY PRIMARY CARE PATIENTS

Authors
Eugenia R. McPeek Hinz MD, Michael E Matheny MD MS MPH,
Josh F Peterson MD MPH, Nancy M Lorenzi PhD MLS MA, Anderson Spickard III MD MS,
Hui Nian MS, Qingxia Chen PhD, William M Gregg MD MS, Joshua C Denny MD MS

Introduction and Background

Mortality risk and severity of illness scores have been used for decades as surrogates of patient disease burden to help risk adjust a population for utilization outcomes.(5–14) Measures developed for outpatient populations have typically relied on subjective descriptions of health status, often assessed through patient survey, along with administrative data with variable success.(8,10,15–18) The burden of data collection of these subjective components limits their scalability for larger populations.(19) An automated implementation of an outpatient mortality risk prediction tool could enable individualized clinical decision support and risk adjustment for population management. Toward this end, we developed and evaluated an EHR-adaptation of the Lee et al. 4-year mortality index (eLee) in an outpatient primary care population and evaluated its ability to predict 2-year mortality.(22) In addition, we evaluated mortality risk variation among providers' patient populations.

Some of the first attempts to characterize disease burden were developed for inpatient populations as a means to normalize patient effects on clinical outcomes. The Severity of Illness Index was an early attempt to stratify patient co-morbidity as a means to understand resource utilization habits of doctors and cost for hospitalizations.(5,24–27) The well-known Charlson 1-year mortality risk measure, designed from hospitalized patient population has been widely used in modified forms to define risk of death based upon the cumulative measure of 13 ICD9 diseases categories.(7,9,28) Extending the Charlson risk score, Elixhauser et al. defined more co-morbidity conditions using a large administrative data set to predict clinical outcomes for inpatient admissions.(21) A common characteristic of these and similar measures is the use of diagnoses and demographics to calculate a risk of death. These metrics have been well validated and benefit from the accuracy connected with short-term predictions in high risk populations.

Defining disease burden for an outpatient population can be more challenging because of greater variability in health status and diversity of reasons for seeking care. In addition, longer-term mortality risk may be better predicted with functional status which indicates how well patients are physically coping with their cumulative disease burden.(11) The Lee index was designed and validated for an outpatient populations over 50 years of age and uses a 12 item patient survey to measure age, gender, presence of 4 specific chronic diseases, body mass index (BMI), current smoking status and functional status to derive the risk score. Functional status represents up to one third of the total score potentially contributing a significant portion to a patient's final mortality risk estimate. The Lee has been used to better delineate frailty and vulnerability in older populations.(23) Its application to a variable-aged outpatient population using EHR data has not been fully

investigated.

Methods

Study Design

We conducted this study among fourteen primary care physicians from a large academic medical center from September 1st 2009 to February 7th 2012. The physicians were recruited by personal invitation by one of the authors (AS). Four clinicians declined to participate. For the other fourteen, informed consent was obtained during an introductory meeting. Of the fourteen physicians, eleven had outpatient Internal Medicine practices, one was a Geriatrician and two were Internal Medicine/Pediatric doctors. All patients seen by these providers were included the study. Informed consent of participant patients was waived. The Vanderbilt Institutional Review Board approved this project.

Patient Cohort

We required all subjects to have an outpatient visit during the initial cohort-defining period to eliminate patients seen during inpatient admissions. The date of this visit established the beginning of follow up and the time point at which the mortality risk score was calculated. To assure attribution fidelity of physician patient pairs, we ordered the sum of established outpatient visits (99211, 99212, 99213, 99214, 99215) for each patient by doctor and then correlated an outpatient visit during the enrollment period with their maximum corresponding study provider. This process removes cross-covering physician care within suites and ensures that the physician most responsible for that patient was correctly attributed. Manual review of twelve patients randomly selected from the original cohort found six patients were excluded appropriately. Of the other six included in the

final cohort, five had the primary care physician (PCP) of record as their study provider and one had a different PCP but clearly was being cared for by the algorithm identified provider. This was verified by the presence of multiple outpatient visits with the algorithm provider as well as orders for labs and physical therapy by the same doctor.

Data Collection

Each physician was given a survey asking relevant provider characteristics including years in practice and gender. We used the Learning Portfolio system, a web-based application linked to the Vanderbilt EHR, to identify and collect data on all patients for whom the study physicians authored clinical notes (inpatient or outpatient).(29) Demographics, billing codes, and other structured data associated with each patient record were extracted from the EHR and the Enterprise Data Warehouse. Patients were followed for the end point of mortality through February 7th 2012.

We manually created lists of ICD9 codes for diabetes, cancer (excluding nonmelanoma skin cancer), lung disease and heart disease to match diagnoses in the Lee survey (Appendix A). As a comparison metric, we also calculated the Deyo modified Charlson mortality risk index. For the Charlson-Deyo version we relied on the originally reported ICD9 codes and updated them to reflect changes in coding since the time it was originally published (Appendix B).(9) Reliability of current smoking status limited complete identification of likely smokers. Consequently we used ICD9 for tobacco use and flags from elsewhere in the EHR to identify smokers.

Measurement of functional status was obtained from a structured (XML) clinic intake questionnaire. This value represented the response to the following question, "Are you having any problems with walking, feeding yourself, bathing, dressing or other daily

activities, that you would like to discuss today?" Spot verification for twenty charts confirmed that from the information available in the patient charts that 19 of these patients did have needs for assistance in functional activities documented in their record via clinic notes or clinical communications.

Outcome Definition/Ascertainment

All-cause mortality data was collected from the Social Security Death Index and internal EHR documentation for defining deaths.(30) We performed secondary verification of all deaths and date of death by cross validation using National Death Index and outside sources such as Hospice records, funeral notices/obituaries or newspaper reports (for those more recently deceased).

Electronic Mortality Risk Measurement

We calculated the Lee for all patients using EHR data directly or as proxies for each of the elements captured by the Lee twelve-question survey. The scores were determined directly for measures of age and gender. Individuals with a single ICD9 code matching at least one of the disease proxy ICD9s for either index were considered to have the disease for purposes of risk calculation. Table 1 shows the assigned risk factors, definition and points assigned along with demographics of our cohort.

Table 1: Demographics and Mortality Risk Point Attribution

Risk Factor	Risk Factor Definitions	Points	Patient count (%)
Age	18-59 + 60-64 65-69 70-74 75-79 80-84 >=85	0 1 2 3 4 5 7	1661 (49.9) 373 (11.2) 364 (10.9) 316 (9.5) 257 (7.8) 198 (6.0) 157 (4.7) Total (3326)
Male Gender	Men > 50 years Men < 50 years†	2 1	867 (26.1) 295 (8.9)
Diabetes	250.*	1	809 (24.3)
Cancer	151.*, 152.*, 153.*, 154.* 162.*, 164.* 189.*, 197.*, 199.* 200.*, 201.*, 202.*, 203.* 204.*, 205.* 206.*, 207.*, 208.*, 209.* 230.*, 231.*	2	128 (3.9)
Lung Disease	232, 233, 234 416.9, 490.*, 491.*, 493.* 398.41	2	118 (3.6)
Heart Disease	425.2 428.*	2	294 (8.8)
BMI < 25 (only for patients ≥50 years†)	Calculated from EHR Height and Weight Or Weight <=135 lbs†	1	622 (18.7)
Tobacco Use	ICD9 305.1 Internal EHR "smoker" flag	2	157 (4.7)
ADLs†	Structured EHR collection	2	244 (7.3%)
Height Missing for 6.9% **† denotes modification of the Lee model for cohort**			

In the Lee, impaired functional status has a range of values from 1 to 7. In our EHR, functional status is only represented as a binary marker so we evaluated the data to determine the optimal equivalent value for this question for patients with impaired functional status. The integrated discrimination improvement method was applied to the data set to determine a value of 2 ($p = 0.040$) as most effective proxy for the model.(31)

Since the original validation study of the Lee risk score was completed on outpatients age ≥50 years, we made several empirical modifications to allow better application to patients under the age of 50. Males under the age of 50 were given 1 risk point (instead of 2) and BMI < 25 did not contribute risk points considered a risk factor for those under the age of 50. We calculated BMI using heights and weights from the EHR. For patients with missing height information (6.9%), we assumed that an individual with a height of 5' 2" and greater would have a BMI less than 25 when the measured weight was 135 lbs or less.

Statistical Analysis

Baseline characteristics were compared using t-test for continuous variables, and the χ^2 test for categorical variables. Time in days from the patient's enrollment visit between the dates of September 1st 2009 and February 7th 2010 to their death or end of follow up February 7th 2012 was thereafter determined. Kaplan-Meier graphs for cumulative mortality were considered separately for risk strata.

We subdivided the thirteen strata of mortality risk into three composite risk levels of low, intermediate and high risk. Incidence rates by risk points were compared together for discovery of natural break points. The incidence rate from the lowest risk strata doubled between points 5 and 6 and tripled as compared to baseline between 8 and 9

points. Low risk was assigned percentage values 1-8% (0-5 points), intermediate risk between 9-20% (6-8 points) and high risk 28-64% (9+ points).

To compare the risk of death between low, intermediate and high-risk subgroups, we calculated hazard ratios (HRs) using Cox proportional hazards models. Since the Lee was derived using age, gender, BMI, smoking status and activities of daily living, these covariates were not adjusted for separately in the analysis. We evaluated the impact of individual provider and physician gender as covariates in the Cox model.

Model discrimination and calibration were assessed by calculating area under the receiver operator characteristic curve (AUC). The Charlson-Deyo model is intended for 1 year risk definition and the Lee measure for patient age 50 and greater. Consequently we compared the AUC for 1 and 2 years for the Charlson-Deyo analysis and by limiting to patients age of 50+ to contrast the original model intentions to our implementation. Finally we evaluated the contribution of age to the model and compared that to the Lee model in total. Integrated discrimination improvement (IDI) was also presented to compare the performances of different models in predicting two-year mortality. Specifically we first derived the predicted probabilities of two-year mortality for each patient from a univariable logistic regression on the risk score, and then calculated IDI based on the method proposed by Pencina, et al.(31)

Aggregate practice panel mortality risk was evaluated for the study physicians. Five physicians were eliminated for further study due to insufficient patient numbers (n< 75) due to part-time status. We evaluated differences between physician's practices for their overall patient risk profile compared to their gender by t-test. We used a Cox proportional hazards model with outcome of death to study physician effect by including physician

identification as a fix effect. Data analysis was completed within STATA version 11.1 (StataCorp College Station TX) and R 2.14.1 (www.r-project.org/).

Results

The initial cohort identified in February of 2010 included 6859 patients who had had encounters with the 14 study physicians over the previous five months. Only 3326 patients remained after removing inpatient encounters and applying physician patient pairing refinement. Mean age of the participants was 58.7 years (standard deviation [SD] 16.5, range 18-102) with 65.1% females and 34.9% males. Male physicians saw older patients on average than female physicians (61.5 vs. 54.4 years, $p< 0.001$). Likewise, male physicians saw more high risk patients (5.9% vs. 2.3%, $p< 0.001$).

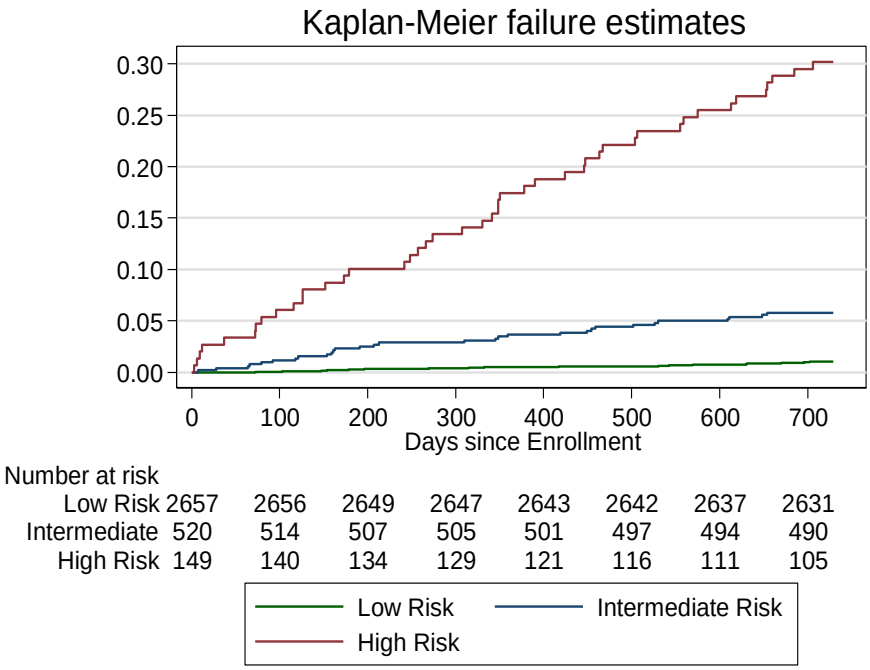
Table 2: Physician Practice Profile by Gender and Risk Stratum

	Female Physicians	**Male Physicians**
Patient Gender		
female	84.4%	52.7%
male < 50	6.1%	10.6%
male ≥ 50	9.5%	36.7%
Mortality Risk		
Low	89.5%	73.8%
Intermediate	8.2%	20.3%
High	2.3%	5.9%
Mean Age (years±SD)	54.4 ±16.3	61.5 ±16.0

A total of 103 patients died during the 2 year study period. Figure 1 shows the cumulative mortality of each risk group. The AUC for the Lee model continuous risk score

was 0.855. For comparison, the AUC for the Charlson index at 2 years was 0.813 (Figure 2). The percentage of patients who died during the two years after enrollment as defined by their low to high risk stratum was 1.1%, 5.8% and 30.2%.(See Table 3) The Hazard ratio (HR) for intermediate risk was 6.0 (95% CI 3.7–9.7) with a p value of < 0.001 and 37.3 (22.7– 61.2) for high risk with a p < 0.001. Multivariable analysis adjusting for physician and gender identified one physician with significantly higher hazard ratio compared to peers. Review of this physician' four patients who died found these individuals were almost entirely classified as low risk due to younger age and had other diseases not captured by the Lee with significant co-morbidity, such as end stage liver disease. There was significant evidence of difference in providers panel for Hazard Ratios which was verified by a likelihood ratio test with a (p = 0.017).(32)

Figure 1: Kaplan-Meier Failure By Risk Stratums



To better understand the contribution of age and functional status to the mortality risk score, we compared the AUC values for the Lee model by age points and by exclusion of the functional status component. With an AUC of 0.778 for age points alone, age yielded a significant proportion of the final model results but the total model improved the discriminatory power to 0.855 with a $p < 0.001$ when compared to age alone. This was also confirmed by a significant IDI of 0.079 between the two models (95% CI, 0.056 to 0.103, $p<0.001$). The addition of a binary functional status marker mildly improved model discrimination, with an AUC without functional status 0.842 versus with of 0.855, and the corresponding IDI was estimated to be 0.018 (95% CI was 0.007 to 0.030, p=0.002)

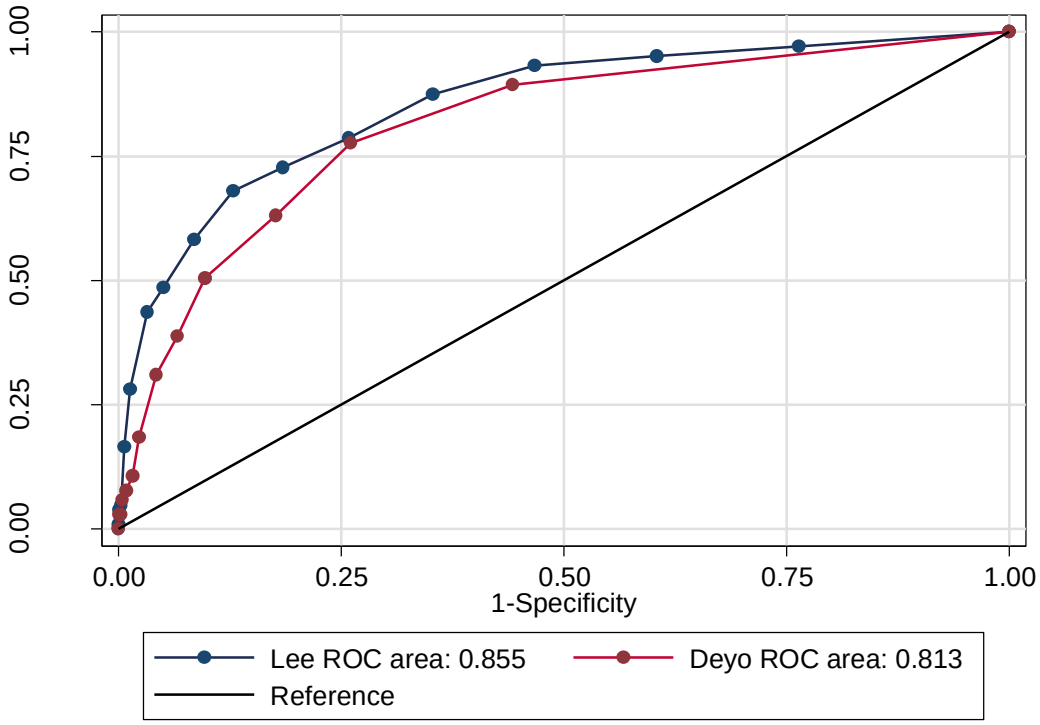
Table 3: Demographics by Mortality Risk

Mortality Risk	**Number of Patients**	**Percentage of the Population**	**Number of Deaths**	**2 Year Mortality Incidence**
Low (1-8%)	2657	79.9%	28	1.1%
Intermediate (9-20%)	520	15.6%	30	5.8%
High (28-64%)	149	4.5%	45	30.2%

The Charlson-Deyo AUC when applied to the same cohort study population for its intended model was 0.820 (95% CI 0.768-0.872) at one year. The effectiveness of the Charlson-Deyo model decreases mildly at two years to 0.813 (95% CI 0.771–0.8549) but not significantly from the Lee model (p=0.089). However, the corresponding IDI was estimated to be 0.069 (95% CI, 0.033 to 0.104) and suggested that Lee model was a

significant improvement in predicting two-year mortality compared to Charlson-Deyo model (p<0.001). The Lee model for persons 50 years and older at two years gives an AUC of 0.843 (95% CI 0.801-0.885). Overall the Lee outperformed the Charlson-Deyo at every cut point in the ROC curve at two years (Figure 2).

Figure 2: AUC for Lee versus Charlson-Deyo at Two Years



Heterogeneity in physician practice mortality profiles:

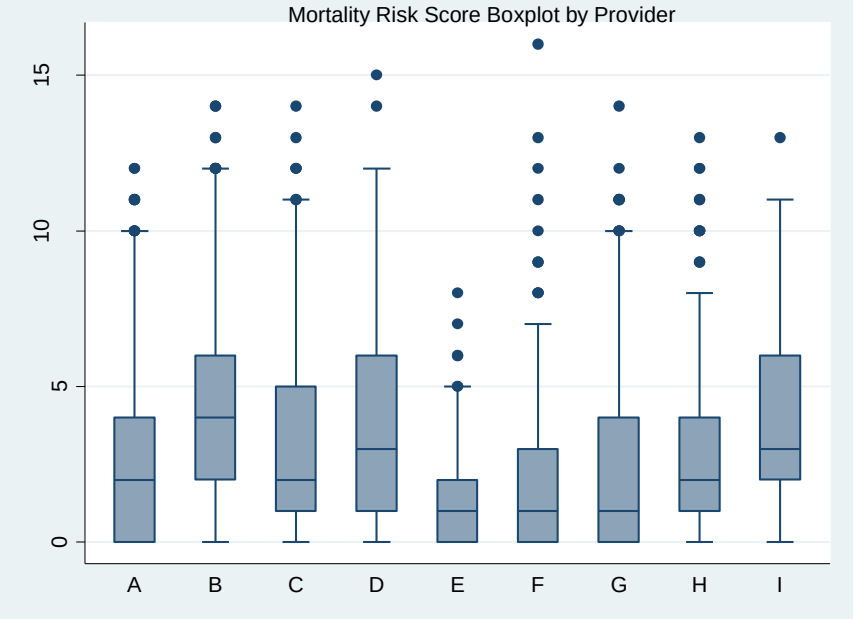
Using the risk stratification definitions of low, intermediate and high risk, a little over twenty percent of the study cohort population fell into the higher risk categories. Physician practices showed significant differences in total percentage of intermediate and high mortality risks patients combined, with one physician with only 4.4% versus 30.4% risk for another of higher risk patients (p < 0.001). (See Table 4) Factors such as years in

practice were not associated with the mean age of their patient population (p= 0.25). Physician gender was a significant predictor of patient's gender with respect to female physicians with a majority of female patient (83.8%) and males greater than age 50 with male primary care physicians (84.7%) (Table 2, $p < 0.001$).

Table 4: Physician Demographics and Mortality Estimates of their patient panels

Clinician	Gender	Years in Practice	No. patients captured	Mean Age of Practice(SD)	Percentage High/Intermediate Risk
A	Female	16	347	57.2 ± 14.9	12.7 (9.6-16.6)
B	Male	20	1023	64.9 ± 13.9	30.4 (27.7 -33.3)
C	Male	16	220	56.8 ± 17.2	18.6 (14.1-24.3)
D*	Male	14	316	63.4 ± 15.0	28.2 (23.5-33.4)
E	Female	8	90	50.9 ± 14.5	4.4 (1.7-10.9)
F	Female	8	278	52.4 ± 16.3	8.3 (5.6-12.1)
G	Female	6	518	54.6 ± 17.0	11.0 (8.6 - 14.0)
H**	Male	5	352	51.9 ± 17.7	15.3 (12.0-19.5)
I	Male	3	89	62.8 ± 13.9	28.1 (18.8- 37.4)
***Geriatrician **Internal Medicine/Pediatric**					

Figure 3: eLee Risk Points by Physician



Discussion

This study demonstrates that using EHR and billing data the Lee Four Year Mortality Risk model successfully predicted 2-year mortality in an outpatient population. With an AUC of 0.855, our model had excellent discrimination consistent with original Lee results even when adapted to EHR proxies and applied to a general outpatient population.(33) This model offers a scalable electronically equivalent method to derive severity of illness for a general outpatient population comparable to the original survey version. Given the growth of adoption of EHRs driven by the HITECH Act (Health Information Technology for Economic and Clinical Health Act), automatic implementation of electronic risk scores may allow more personalized care recommendations and provide greater insight into care quality.(34)

To provide potentially clinically significant stratums of risk, we chose the use cut points of low, intermediate and high risk as done similarly before by Cagne et. al. The

utility of these types of cut offs allows for an individual risk score which can then inform care on a personalized level. For example, current population based recommendations for colonoscopy screening recommend stopping at the age of 75. In our study, 21% of patients greater than 75 years of age fell into a low risk stratum and could potentially still benefit from screening. Conversely patients in a higher risk category may benefit from different quality metric goals.

We found differences in physician practices for proportion of higher risk patients in their panel not related to years in practice or mean age of the patients in the clinician panel. Such population differences are important when considering differences in patient outcomes by physician. Low and intermediate risk patients may be a better population to consider physician quality measures by selecting the patients populations for which providers have more influence on outcomes.(35) Robust risk adjustment methods are critical for effective provider profiles and engagement of clinicians.(36)

Overall the Lee model performed slightly better than the Charlson-Deyo across a general outpatient population. (, though this difference may not be clinically significant (AUC XX vs YY). Josh the IDI does indicate clinical significance) The greatest differences in the Lee model's effectiveness were seen in an older population, possibly due to the Lee inclusion of functional status and age in the model. We noted that the Lee did not include certain life-threatening disease such as end stage liver disease, which adversely affected its performance for some younger patients. Future work to extend this model would benefit from expanding the potential diseases considered in the model. Additional incremental benefits may be seen with more detailed and structured data assessment of functional status exists.

This implementation of an electronic equivalent Lee index allowed determination of clinically relevant risk stratums. The Lee does not include patient or physician assessment of health status as a component of the score and was thus more easily translated into an electronic version. We are only aware of one other electronic implementation of the Lee index that evaluated effectiveness of colorectal cancer screening with mortality risk.(37) Ultimately, scalable risk defining models such as these have the potential to improve patient care through personalized clinical decision support and allow for risk adjustment to populations.

Limitations caution interpretation of these results. First, several changes were made to the original Lee algorithm to make it applicable to EHR data and to include patients under the age of 50. These modifications to the model for BMI and male gender were determined clinically and have not been validated on a test cohort. We also approximated functional status using a binary value instead of the more granular 7-point scale in the original Lee survey-based method. By assigning a value of only 2 additional points for functional status, we likely underestimated the risk stratum for some members of the cohort. If the true varied functional status values were available the significance of effect on the model would have likely been more evident.

Finally we looked at only two years of mortality data instead of four years. Even with the shorter time period of observation and changes in calculations, differences in magnitude of effect for the different risk stratums were evident and could be used clinically. Future efforts should evaluate EHR based methods over longer time periods to determine if the effect seen thus far continues.

Our study was small with only fourteen clinicians enrolled and 3326 of their

patients from one academic medical center evaluated. We employed a definition of physician patient pairs based upon maximum number of established care visits for a study provider, increasing the probability of fidelity of who the primary care provider was for that patient. However, this has a secondary effect of establishing a minimum level of data available for score determination that may serve to minimize missing data and represents patients who tend to see a provider more regularly. Alternately patients without a primary care provider or who serially use acute care services may be underrepresented.

Our application required only one of the study ICD9 diagnoses to have a point risk assigned. As a result, disease risk points may have been given for a diagnosis that the patient may not actually have, since ICD9 codes can have poor positive predictive values.(38) Data for current smoking status was the most difficult to determine as there were multiple methods to define this data and underestimation of the population is likely for this marker.

Conclusion

We found that an EHR-derived 4-year mortality risk measure can successfully risk stratify an outpatient population and predict two year mortality outcomes. Physician practices demonstrated differences in the numbers of intermediate and high-risk patients, but mortality outcomes were not physician-dependent. Physician and patient same gender co-segregation contributed to aggregate co-morbidity of physician practices. This measure was calculated using readily available EHR data and has potential for real-time implementation to assist in personalized clinical decision support and to adjust for the risk of physician practices to better understand their outcomes.

Acknowledgements

This work was supported by the NLM Training grant: 3T15LM007450-08S1 and a grant from the National Board of Medical Examiners.

CHAPTER III

AMBULATORY EVALUATION OF A 4-YEAR MORTALITY RISK INDEX

BY SURVEY AND ELECTRONIC HEALTH RECORD ADAPATION

Introduction

In the modern age of medicine, better care for common co-morbidities such as heart disease, cancer and infection allows more adults to live active productive lives well into their eighth and ninth decades of life.(39) As consequence, the boundaries of healthfulness are less well defined leading to potentially inappropriate care. For example, previous studies have found application of age only screening recommendation led to screening of cancer patients with limited potential benefit or arbitrary limitation of cancer screening based upon age for patients in good health.(40,41) To respond to these changes, some clinical care guidelines are beginning to incorporate life expectancy into decisions of care recommendations.(42) Longer duration validated prognostic tools exist but the reproducibility of their applications in a clinical setting is not well defined.(12,22)

In Chapter II, we found the accuracy of our EHR-adaptation of the Lee index extended to a general adult population with an AUC of 0.855 at two years. This value is very comparable to the original validated measure at four years of 0.82. The two main limitations of this study were an unknown accuracy of the electronic proxy components used for determining co-morbid diseases and binary functional status indicator. The application of ICD9s to capture the four diseases in that study may cause false positive

results due to poor positive predictive poor associated with ICD9s.(43,44)

To evaluate the similarities and differences of the Lee survey versus the electronic EHR-adaptation, in this study, we surveyed adult patients in ambulatory primary care clinics with the original Lee survey instrument and compared the results to electronic scores. To better match the questions of functional status with the survey, we implemented an expanded form of the functional status questions into the EHR. We completed chart reviews of discrepancies between the survey and electronic results to define the biases associated with both methodologies to establish an adjudicated standard for comparison.

Methods

Overview

We preformed a cross sectional analysis of primary care patient's responses to a survey on measures of health status and compared these results to their corresponding electronic score for a four year mortality index. We surveyed 415 primary care patients at the time of their clinic visit at a large academic medical center from November 2011 to February 2012 for the Lee 4-year and Schonberg 5-year mortality risk indices. Patients consented for enrollment were verbally administered a combined 14 question survey. We derived the electronic equivalent mortality risk score for each patient from EHR and billing data. The Vanderbilt Institutional Review Board approved this project.

Survey Creation and Administration

We used the 12 question Lee survey with minor changes to include former tobacco use as well as current tobacco use. We then added two new questions on health status and the number of prior hospitalizations to allow for calculation of the Schonberg score after the Lee questions. Thus, the final survey was a combination of the Lee and Schonberg (Appendix C). The survey was administered verbally to patients by three study personnel (the author, a board-certified internist and two trained research assistants, who were observed intermittently by the author for quality control).

Study Cohort

We approached patients 18 years of age or older for survey participation generally immediately before or after their visits with the physician. We collected age and gender for patients not surveyed but also seen in the clinic on the days of collection. Physicians specialties for the two clinics where patients were interviewed included two geriatricians and the rest were primarily internal medicine(IM) or IM/Pediatric practices. We consented and surveyed patients in the privacy of the examining room with the patient and any accompanying persons. For patients unable to answer on their own, answers from the family or chaperone if familiar with the patient were used instead. Dementia patients occasionally had family members who responded differently than the patient and their answers were recorded instead.

Data Collection

We transcribed the original survey results into a REDCap database

(http://www.project-redcap.org/) along with patient identifying data of name, date of encounter and medical record number (MRN).(45) We exported REDCap survey data into the STATA statistical package and calculated the Lee and Schonberg prediction results from the survey (Table 5). Equivalent electronic demographics, billing codes, and other structured data associated with each patient were extracted from the EHR and the Enterprise Data Warehouse in a similar fashion as described in Chapter 2.

Common to the Lee and the Schonberg indices are components of functional status to their risk prediction scores. In the prior study (Chapter II), we used an electronically-completed clinic intake questionnaire to approximate functional status. This questionnaire is filled out by clinic staff when a patient is brought into a clinic room and includes questions on reason for visit and included the question "Are you having any problems with walking, feeding yourself, bathing, dressing, or other daily activities that you would like to talk about today" to address the patient's functional status. Upon completion, this data is represented in a codified form in XML in the EHR. To better approximate the Lee measures electronically, we modified the XML-based Clinical Intake Questionnaire. The prior question was changed into a two-part process which starts with a screening question: "Does the patient have any problems with walking 2-3 city blocks, bathing, dressing, or remembering important things?" This question captures the key elements of the four questions included in the Lee. If the patient responds in the affirmative, the form expands (using javascript elements in the web interface of our EHR) to ask four additional questions related specific measures of independence with activities of daily living (See Figure 3). All clinical staff responsible for placing patients in examination rooms and completing the intake questionnaire were trained individually on the new changes.

Figure 4: Functionality Questions

1. Has a health or memory problem caused the patient to have any difficulty with bathing or showering?
2. Has a health or memory problem caused the patient to have any difficulty with managing money such as paying bills or keeping track of expenses?
3. Has a health problem caused the patient to have any difficulty with walking 2-3 city blocks without resting or assistance with a mobility device?
4. Has a health or memory problem caused the patient to have any difficulty with pulling or pushing large objects, such as a large living room chair or small sofa?

Adjudicated Standard

Due to potential biases in the responses to the survey, we evaluated survey responses that differed from electronic ones using manual chart review. Age, gender and electronic BMI measurements demonstrated close correlation and thus were not evaluated further. For the four disease categories and tobacco use components of the Lee score, the primary author reviewed patient charts for discrepancies between survey responses and electronic data. All cases of correlation by survey and electronic data for presence or absence of a measure were considered valid. Chart review for evidence of disease included review of the patient problem list, physician notes, ICD9s or identification of medications or laboratory values consistent with disease processes.

Chart review of discrepancies between responses for limitations of functionality between the survey and electronic forms were more challenging due to variable presence of sufficient

chart documentation of the patient's functional status/limitations. Thus, a global assessment of presence or absence of "functional status impairment" was deemed sufficient. One reviewer reviewed all discrepant cases. A second reviewer reviewed a random set of 50 of the 121 discrepant cases for calculation of interrater agreement using Cohen's Kappa and percent agreement.

We created new disease and functional status control variables based upon the chart reviews of the discrepant cases to evaluate the strengths and limitations of the survey and electronic Lee adaptation. Points for the adjudicated functional status control were given based upon the patient's survey results for those cases where chart reviews were positive.

Statistical Analysis

Baseline characteristics of the patient cohort were compared using t-test for continuous variables, and the χ^2 test for categorical variables. All analyses assumed a two-tailed distribution. We compared the accuracy of the gold standard combined electronic and chart review Lee points to the EMR-derived elements to the actual results of the Lee survey points score using recall, precision, F-measure (2 * ((precision * recall)/(precision +recall)). Pair wise population comparisons for binary presence of disease and smoking components were evaluated by McNemar's exact score. The Schonberg score, Lee, and eLee were compared to each other and to the adjudicated standard using Spearman correlation coefficient. Data analysis was completed within STATA version 11.1 (StataCorp College Station TX) and R 2.14.1 (www.r-project.org/).

Results

We interviewed a total of 410 patients in two primary care clinics from November 2011 to February 2012. One patient withdrew from the study after survey completion, leaving 409 study subjects. Mean age of the participants was 58.3 years (standard deviation [SD] ± 17.3, range 19-99) with 58.2% females and 41.8% males (Table 5). Patient-reported BMI was moderately lower than that found electronically, 27.1 ± 5.1 versus 29.4 ± 11.9.

The Schonberg questions of health status and hospitalization yielded similar result as the original paper for health status and moderately more individuals, hospitalized twice in the past year, 11.5% in our survey versus 5.4% in the original cohort. Spot review of charts for health status confirmed the subjective nature of this measure with three patients above the age of 80 reporting their health as excellent even though two had been hospitalized twice in the past year, one of whom was survey positive for cancer and the other patient not hospitalized had admitted trouble with memory and walking. On the other end of the spectrum, three patients in their 50-60's (59,68,57) described their health as poor with limited findings in their charts for significant diseases, besides well controlled diabetes and morbid obesity, or much by way of functional limitations.

Table 5: Demographics of Survey Population and Electronic Equivalent

	Survey n= 409	Electronic n = 409
Age (years) [L, S]	58.3 ± 17.6	58.3 ± 17.6
Gender- Female [L,S]	58.2%	58.2%
DM [L,S]	83 (20.3%)	90 (22.0%)
Cancer [L,S]	50 (12.2%)	28 (6.9%)
Lung Dz [L,S]	36 (8.8%)	26 (6.4%)
CHF [L]	25 (6.1%)	31 (7.6%)
Smoking [L,S]	Never [L,S] 53.4% Former [S] 32.8% Current [L,S] 13.7%	Yes/Current/flag 11.0%
BMI (kg/m^2) [L,S]	28.8 ± 7.7 27.1 ± 5.1 for those with an EHR-defined height	29.4 ± 11.9 Missing height 13.7% (56)
ADL/Bathing deficits[S,L]	50 (12.2%)	18 (4.9%)
Walking deficits[S,L]	145 (35.5%)	39 (10.7%)
Managing Money deficits[L]	42 (10.3%)	21 (5.7%)
Trouble pushing large Objects[L]	115 (28.2%)	35 (9.7%)
Health Status[S]	Excellent(0) 11.3% Very Good(1) 31.0% Good(2) 34.2% Fair(3) 18.6% Poor(4) 4.9%	
Hospitalized[S]	None 73.4% Once 15.1% Twice 11.5%	
Race		A 2.9% B 21.8% I 0.2% U 2.7% W 72.4%
Values represent Mean±SD. L = Lee 4-yr Model component; S = Schonberg 5-yr Model component		

Table 6 shows the rates of Lee component measures in the survey compared to the electronic results. Survey results reported presence of cancer in 12.2% and chronic lung disease 8.8% of patients. These values exceeded the rates found electronically: cancer was found in 6.9% and chronic lung disease in 6.4% of patients, respectively. Survey data was

similar to electronic data for the rates of diabetes mellitus (20.3% vs. 22.0%) and congestive heart failure (6.1% vs. 7.6%). None of these subpopulations showed much population similarity by McNemar testing indicating significant differences in results.

Table 6: Electronic Proxies Performance to Predict Survey Data

	Recall	Precision	Accuracy	F-measure	McNemar's Exact
DM	82.2%	89.2%	93.9%	0.856	0.230
Lung Dz	60.0%	41.7%	92.4%	0.492	0.071
Cancer	75.0%	42.0%	91.2%	0.538	<0.001
CHF	48.4%	64.0%	94.1%	0.552	0.308
ADL's	92.1%	33.3%	84.6%	0.490	<0.001
Walk	89.7%	24.1%	72.1%	0.728	<0.001
Push Large Object	74.3%	22.6%	76.0%	0.346	<0.001
Money	57.1%	28.6%	90.5%	0.406	0.001
Bathing	72.2%	26.0%	89.7%	0.382	<0.001
Smoke Current	77.8%	62.5%	92.4%	0.694	0.071

* Yellow highlighted McNemar's score for equality demonstrates the lack of correlation of the survey and EHR-adaptation of scores.

Cancer Chart Review

Review of the discrepant cancer cases found that of the 29 survey positive and electronic negative patients, 11 patients had remote history of cancer, greater than 10 or more years ago, 5 had previously treated cancers but not recently classified as positive. Of the other 13 cases, 7 had nonmelanoma skin cancers, 2 noncancerous resections and 6 had no specific chart evidence of cancer. There were 7 survey negative for whom electronic evidence of cancer was found. One of these had a cancer under treatment elsewhere, 5 had ICD9s for cancer testing, and 1 had a precancerous state with one older apparently

erroneous ICD9 denoting transformation. After chart review 16 historical cancers by survey were reclassified as true positive and 6 electronic cases redefined as false positives.

Table 7: Discrepant cases for Cancer by Survey versus Electronic Adaptation

Cancer	Electronic Positive	Electronic Negative
Survey Positive	21	29 - 11 remote cancer hx >10 years - 5 with cancer previously treated <10 years - 7 other skin cancers - 2 noncancerous resections - 6 no specific history found
Negative for Cancer	7 - 1 positive cancer treated elsewhere - 5 noncancerous ICD9 for testing - 1 precancerous state with one older transformation ICD9 from 5 years ago not repeated	352

Chronic Lung Disease Chart Review

For the chronic lung disease cases, we reviewed 21 patients denoted survey positive and electronic negative (Table 8). A majority (14 out of 21), were found by chart review to be ICD9 positive, mainly with ICD9 496 representing a general code Chronic Obstructive Pulmonary Disease (COPD) and two others with significant lung disease such as bronchiectasis. All of these were categorized as true positives on chart review. These patients were false negatives in the electronic algorithm due to failure of capture of their ICD9 for these clinics in the electronic data warehouse. For the other 5 individuals, 2 had mild lung disease and 3 had no chart evidence of lung disease. The survey negative,

electronic positive subset consisted of 11 patients. Seven patients were true positives for the electronic algorithm with 5 COPD patients and 2 other significant lung diseases, such as pulmonary fibrosis. The remaining 4 patients had no chart evidence of disease.

Table 8: Discrepant cases for Chronic Lung Disease by Survey versus Electronic Adaptation

Chronic Lung Disease	Electronic Positive	Electronic Negative
Survey Positive	15	21 - 14 with COPD by ICD9 in chart - 2 with other significant Lung disease - 2 with mild lung disease - 3 with no chart evidence of disease
Negative for Lung Disease	11 - 5 with COPD disease, - 2 other significant lung disease with COPD ICD9 not relevant - 4 with no chart evidence of significant Lung Disease	362

Diabetes Mellitus Chart Review

Twenty-five diabetes mellitus (DM) cases were dissimilar. Nine patients were survey positive but electronic negative cases, 3 were deemed positive by history and ICD9s, the remaining six were negative by chart review. For the 16 survey negative, electronic positive cases, 12 were deemed positive with ICD9s and evidence of impaired fasting glucose or borderline Diabetes by chart review. The other 4 cases were negative by chart review.

Table 9: Discrepant cases for Diabetes Mellitus by Survey versus Electronic Adaptation

Diabetes Mellitus	Electronic Positive	Electronic Negative
Survey Positive for DM	74	9 - 3 positive (1 with history of gestational DM) - 6 negative by chart review
Negative for DM	16 - 12 considered impaired fasting glucose or borderline DM by chart review - 4 negative by chart review	310

Congestive Heart Failure Chart Review

The final disease component under consideration was congestive heart failure (CHF). With a McNemar score of 0.308 between the survey and the electronic score directly, this measure showed the best population similarity of any of the Lee measures. For the 9 survey positive, electronic negative cases, 3 were positive for isolated CHF ICD9s mostly remotely and the other 6 lacked evidence of CHF, often having a history of non-CHF heart disease such as endocarditis, which is not considered in the Lee "CHF" category. Considering the 15 survey negative, electronic positive cases, 10 were positive by chart review for heart disease. Chart evidence for CHF included diagnosis in patient notes including diastolic heart failure, a less than normal ejection fraction and current medication of diuretics and ace inhibitors. The other 5 electronic positive, survey negative cases did not have CHF.

Table 10: Discrepant cases for CHF by Survey versus Electronic Adaptation

CHF	Electronic Positive	Electronic Negative
Survey Positive	16	9 - 3 positive for CHF mostly remotely 5+ years ago - 6 negative or with other significant heart disease
Negative for CHF	15 - 10 positive for CHF by chart review - 5 with other significant heart disease not CHF	369

Smoking Chart Review

Using the Schonberg version of the smoking question, we found that 53.4% of patients reported never using tobacco, 32.8% were former users and only 13.7% admitted current tobacco abuse. The electronic measure for current tobacco identified 11.0% of the population as current smokers. Simplifying the survey data to "current" or "not currently", using tobacco, the McNemar result is 0.071 suggesting the two populations are dissimilar (Table 11). Review of patient charts for discrepant cases found of the 21 electronic negative but survey positive patients, 18 still smoking, 2 recently quit and 1 reportedly only used in college where the possibility of social tobacco use appeared to still be likely. Of the 10 survey-negative cases but electronically positive cases, 5 appeared to still be smoking by chart review and 5 appeared to have quit a few years back.

Table 11: Discrepant cases for Smoking by Survey versus Electronic Adaptation

Smoking	Electronic Positive	Electronic Negative
Survey Positive	35	21 - 18 currently smoking - 2 recently quit - 1 only used in college
Negative for Smoking	10 - 5 smoking currently, generally just a few cigarette daily - 5 quit a few years back	343

Chart Review of Functional Status

We found 174 (42.5%) individuals had positive responses to any of the four measures of functional status by the survey compared to only 63 (15.4%) by the intake questionnaire ($p< 0.001$) (Table 12). Survey respondents reported the most trouble with walking several blocks for 145 individuals (35.5%) or inability with pushing a large object like a living room chair for 115 patients (28.2%).

The clinic intake personnel completed the new intake questionnaire and answered the functional status question on 89.5% of patients in our survey. While those without intake results appeared similar by age, mean 56.3 versus 56.0 years to patients with a response of "no" to the baseline functional status question, they tended to report better measures of health status by Lee survey, 1.36 versus 1.68 ($p = 0.053$).

Electronic capture for these same measures found 39 (10.7%) with walking deficits and 35 (9.7%) with trouble pushing large objects. For the survey measures of ADL's and managing money the self reported deficits of 50 (12.2%) patients and 42 (10.3%) with electronic correlates of 18 (4.9%) and 21 (5.7%). We reviewed the 5 patients that were survey negative but electronic positive. Chart review suggested three had functional status limitations. The two electronic algorithm false positives included one individual with a

stress fracture and the other with no clear evidence of any limitations.

Table 12: Discrepant cases for ADLs by Survey versus Electronic Adaptation

ADL	Electronic Positive	Electronic Negative
Survey Positive	58	116 - 44 considered positive - 72 without significant chart evidence
Negative for ADL	5 - 3 considered positive - 2 negative with 1 with recent stress fracture injury	230

Review of survey positive but electronic negative cases found 44(37.9%) patients with some evidence of functional status limitations but a majority of patients 72 (62.1%) had no significant evidence by chart review to back up functional limitations. For example, many cases ruled as negative by chart review had evidence of non-limiting chronic pain, such as knee osteoarthritis. For these patients, no appreciable limitations could be determined as evidenced by review of the chart notes – many of the patients walked for exercising regularly or performed yard work, some even noting regular running programs or playing tennis. Additionally, exam notes supported lack of functional limitations with descriptions like normal gait and strength by the primary care, orthopedic or neurology physicians. Many of the adjudicated positive cases had dementia or Parkinson's disease with cognitive limitations inhibiting the ability to share functional status completely with the intake personnel. See Tables 13 and 14 for descriptions of example patients.

Review by a second practicing physician of a random set of 50 of the discordant cases for functional status, identified 18 cases positive for functional status limitations by chart review for a Kappa score of 0.48 consistent with moderate agreement between

reviewers. For the 58 concordant cases by survey and electronic measure the patient self-reported functional status was used for the adjudicated standard.

Table 13: Survey Positive Functional Status adjudicated as Negative for Impairment

70s female	Chart description by Geriatrician as Independent with ADLs and no memory impairment
50s female	with shoulder injury still working with e survey results positive for all four measures
60s female	with OA of knees walking only 2 hours a day instead of 3 hours
60s male	with ACL tear but still playing tennis and working in yard: complaint of trouble walking and pushing large objects
50s female	with knee pain but lives alone and works in real estate
70s female	Walking ¾ mile 4 times a week for exercise
60s male	semi retired plumber, jogs, plays golf and lifts weights for exercise

Table 14: Survey Positive Functional Status adjudicated as Positive for Impairment

30s male	with dwarfism complaints of trouble walking only
80s male	described as frail in clinical notes, also with mild memory impairment
70s female	with severe dementia, accompanied by daughter
50s female	with MS and recent fall with rib injury
60s male	Male with difficulty standing in chart notes

Correlation of Survey and Electronic Indices of Predicted Mortality

The new adjudicated results for the four diseases, smoking and functional status were added together with the standard covariates of age, gender and BMI to provide an adjudicated control. Subsequently we compared the control to the survey and electronic results for individual disease and smoking components (Tables 15 and 16). The electronic algorithm for diabetes, cancer and CHF showed improved population correlation. An undetermined error with chronic lung disease exportation from the EDW limited its

correlation to the survey results being more effective as well as smoking.

Table 15: Adjudicated Control Standard versus Electronic Version

	Recall	Precision	F-measure	McNemar's Exact
DM	96.6%	95.6%	0.961	0.706
Lung Dz	57.9%	84.6%	0.687	0.007
Cancer	81.5%	78.6%	0.800	0.763
CHF	89.7%	83.9%	0.867	0.480
Smoke Current	88.4%	67.6%	0.766	0.011

Table 16: Adjudicated Control Standard versus Survey Version

	Recall	Precision	F-measure	McNemar's Exact
DM	86.5%	92.8%	0.895	0.157
Lung Dz	81.6%	86.1	0.838	0.564
Cancer	96.4%	52.9%	0.683	<0.001
CHF	65.5%	76.0%	0.703	0.317
Smoke Current	91.4%	94.6%	0.930	0.767

A total of five separate Lee scores were calculated for comparison. They included:

1) the original Lee 4-year survey score;

2) the Schonberg 5-year survey score;

3) the expanded functional status (0-7 points) electronic score;

4) the 2-point original electronic score;

5). and an adjudicated "gold" standard

comprised of concordant survey and electronic positives and chart reviewed results for the four diseases categories, current smoking status, and functional impairment (as a the self

report survey score).

We evaluated that difference in the total scores of the five separate scores using Spearman's correlation coefficient (Table 17). Using the adjudicated standard for comparison, the 2- point functional status and extended functional status scores had exactly the same score of 0.936 (95% CI 0.915-0.953) compared with the Lee survey with a score of 0.924 (95% CI 0.896-0.947). All three versions of showed excellent correlation and overlapping confidence intervals indicating statistical equivalence of the three scores to each other. The correlation of the 2-point and expanded electronic scores of 0.996 (95% CI 0.001-0.998) highlights the effective similarities of the functional status point attribution. The Schonberg score against the original Lee score had a very good coefficient of 0.883 highlighting similarity of these scoring methods despite variations in the risk point assignment.

Table 17: Spearman Correlation Coefficients

Spearman (95% CI)	**Lee Survey**	**Schonberg**	**2 point eLee**	**0-7 points eLee**
"Gold" Standard Score	0.924 (0.896-0.947)	0.795 (0.751-0.834)	0.936 (0.915-0.953)	0.936 (0.915-0.953)
Lee Survey Score		0.883 (0.858- 0.904)	0.858 (0.821- 0.889)	0.856 (0.820- 0.888)
Schonberg Score			0.726 (0.672-0.774)	0.727 (0.674-0.774)
2 points eLee				0.995 (0.991- 0.998)

Discussion

This study demonstrates the correlation of an EHR-adaptation of the Lee 4-year index to the original survey validating the similarity of electronic version, with a Spearman correlation coefficient above 0.90 for the survey and the electronic versions against an adjudicated standard, both methods had excellent calibration. The Lee EHR-adaptation index offers a scalable method for deriving disease burden in an outpatient population.

According to chart review, both the survey and the electronic scores had errors in identification of co-morbidities and functional status. On the survey side, we found that patients often did not identify their diseases as accurately as their medical record, as has been found in other studies in relation to health literacy.(46) Survey respondents under-identified a personal history of diabetes and congestive heart failure and over-identified a history of cancer. The latter was typically due to recollections of non-melanoma skin cancers or testing for possible cancers that turned out to be negative. For functional status, patients identified short-term or non-limiting functional limitations related to pain, possibly due to inadequate understandings of the questions being asked. Capture of non-visual limitations such as dementia was less effective by the measure of intake questionnaires and was identified with survey method frequently and is a known problem of clinical recognition.(47) Expansion of the Lee EHR-adaptation to include dementia and medications provides an alternate means of more reliably obtaining that data when compared to the chart validation.

Errors in the electronic version of the Lee derived from missing data or false positives associated with use of a single ICD9 to identify a given disease. We identified gaps in ICD9 data from the EDW in particular to the ICD9s for chronic lung disease that did

not correlate with information available for review in the EHR. Review of the electronic capture by ICD9 for cancers in particular identified many patients with history of cancer more than 10 year before. The benefit of assigning full point risk for remote cancer history warrants further study. Simply limiting ICD9 data to less than five years and requiring multiple ICD9s to instantiate a disease category may decrease false positives.

We implemented capture of functional status within our EHR by changing the clinic intake questionnaire to better mirror the Lee survey. Correlation of the electronic extended functional status score versus a binary application of two points for functional status score were equivalent. This finding supports the use of a single composite functional status question to capture this measure in a more efficient manner. The lack of capture of the memory related functional status limitations via the intake form suggests an opportunity to use alternate means to screen for less visible functional limitations such as dementia and warrants consideration and further study. The lack of response to these questions may be biased by both social stigma of neurodegenerative diseases and the very presence of cognitive disease limiting patient's ability to respond.

Limitations caution interpretation of these results. First the size of the study is small with only 409 individuals and applied to a general Internal Medicine population. Our mean age was 58.3 years similar to our previous study with a mean age of 58.7 years. The original Lee survey included only patients > 50 years, and thus studied an older population (mean age 67 years). While race is not a component of the original Lee index, the possible effect of our cohort composition is unknown and included almost twice the percentage of African Americans in this study versus the original Lee study.

By using only subjective self reported functional status determinations, we had

limited means for verifying actual and amount of functional or mental disability. This limited our determination of absolute truth and led to our use of an adjudicated standard group to represent the closest intentions to a gold standard. Biases in respondent's survey to diseases and functional status responses lead to variable over and under representation of the survey covariates. The processes of clinic intake personnel appeared to limit capture of all patients with limitations especially for those with memory or less physically evident functionality problems.

Conclusion

We found that an EHR-derived 4-year mortality risk measures effectively correlates with the original Lee survey version for risk determination. By identifying sources of possible errors in the electronic scoring algorithm, changes to improve the fidelity of capture risk covariates allows for improved more uniform scoring. An EHR-adaptation approach to risk scoring has the benefit of automation allowing for possible inclusion in a clinical setting to inform patient care on an individual level or in aggregate on a physician panel or a clinic level for risk adjustment.

Acknowledgement

This work was supported by NLM training grant 5T15LM007450-11 and the National Board of Medical Examiners.

CHAPTER IV

SUMMARY

Summary of Findings

These two studies explore the application of 4-year mortality risk score calculated from EHR data to define disease burden in a general ambulatory population. The first study demonstrated that an electronic adaptation of the Lee index using EHR-derived data as proxies for the original survey data was able to predict patient death at two years. Physician practices differed considerably in proportions of higher predicted mortality risk individuals, due in part to gender co-segregation by patients to similar gender physicians. The second study compared direct patient survey-defined risk scores (using the original model) to the electronically-defined model. We also studied the effectiveness of an expanded EHR form implementation to capture functional status, an element important in prior studies but not often captured well in EHRs. The outcome of the latter study in defining ascertainment gaps informs on potential improvements for the EHR-adaptation. Compared to an adjudicated "Gold" standard formed through chart review, both the survey and electronic versions performed well. Surprisingly, the expanded functional status form did not improve over the former binary implementation consisting of a simple screening question regarding the patients' functional status. These results favor the use of a single composite functionality question, although cognitive deficits may be inadequately captured through this method. Future studies should investigate the application of a functional

status question with EHR-mined diagnoses such as dementia, which are less obvious to a casual observer and significantly impact functional status.

These results demonstrate the effectiveness of an electronic algorithm to reproduce a previously validated mortality risk measure. Such a tool has the potential for larger scale more automated implementation which could be applied in clinical settings for individual level decision support or in aggregate population level risk adjustment. Methods that apply to a general adult population allows for risk adjustment in a complete fashion that implementation with a limited age would not.

Limitations

Limitations of these studies caution their interpretations. The primary limitation of these studies is the relatively small numbers of individuals and physicians used for comparison, and the length of time (two vs. four years) for which mortality was assessed. We validated the ability of the eLee to discriminate risk of death for two years instead of four years as the original study. The AUC value at two years for our model 0.86 was slightly higher than four year validation set of 0.82. As our data saw a small decrease in the AUC from year one to year two, the outcome at four years would likely be decreased over the two year results.

A second general limitation is derived from our electronic application of the Lee index applies outside of the intended and validated population of 50 years and greater. As a consequence of broadening the index to a general adult population, we applied clinically modifications to the risk definitions for BMI and male gender under the age of 50 that are not validated for this portion of the population.

In regards to our electronic implementation, we used EHR data as proxy equivalents for survey components of diseases, BMI, smoking status and functional status. For the electronic definition of diseases we substituted ICD9 codes and only required one value to be included. The effect of this definition was to include patients with possibly errant diagnoses codes. For the BMI proxy, a minority of patients did not have height information and even fewer of those were treated to risk points due to a weight of less than 135 pounds so the overall contribution to these changes is limited. For smoking status indicator we saw an improvement in the electronic capture of this measure from the initial study results in 2009 to the values captured in 2012 due to EHR improvements to capture this information over time. Finally for functional status, we used a binary value as a proxy of functional status with limited validation of its correspondence to measures of functional independence. We evaluated the utility of an expanded electronic functional status measurement as well as the validity of these proxies in general in the final study.

In the second study we, we applied a research survey tool in a clinical setting where time constraints limited the interview results to a simple answers. The original study benefited from evaluating over 20 different measures of functional and disease status allowing more fine differentiation of true limitations and actual disease. In our study patient health literacy was not included as a part of the study. The subsequent chart review of discordant cases revealed significant disparity in patients understanding of the diseases they reported.

To capture all possible Lee index responses on functional status we expanded the electronic capture of dependence measures within the clinic intake questionnaire. Patient responses to these measures are captured by the medical assistants for either the legacy

method of one functional status question or with the newer version. In the first study less than 8% of the population was electronically defined as positive. Chart review of these patients did indeed support their dependence issues. In the second study after training the clinic intake personnel on the new method, capture of electronic functional status limitations doubled. Even though the mean age of the two studies was similar, the patient base of the second study included two geriatricians, so the possibility of increased functionality limitations for the second study population could be true. Still the potential for a Hawthorne effect is also possible and should be considered in the difference in electronic capture. Finally there also appeared to be some component of intake personnel ascertainment bias of patient robustness in capturing functional status needs electronically. This was especially evidenced by the differences in better self defined health status values for those whose primary functional status question was not even answered in the first place. Along with the intake personnel's apparent judgment bias of physical functioning, there also some evidence by chart review of the survey positive electronically negative functional status patients, that limitations such as dementia were not noted electronically as well as the physical functional status limitations. The second study in particular, identified the biases in data collection from both methods of scoring. The electronic version benefits from being able to be iteratively improved and eventually automated.

Future Directions

These studies lay the groundwork in establishing the equivalency of the electronic Lee version to the original survey in form of measurement and to a lesser extent outcome

determinations. The use of EHR data proxies appears to be as effective in categorizing for the components of the Lee with caveats identified in the second study to be noted. The next step entails validating the eLee by applying the new model in a large general adult population. Using the lessons of the second study modifications of the electronic algorithm for disease and functional status determination are warranted. Specifically, disease definitions should include the following limitations of restricting ICD9 diseases inclusions to only those associated to the patient within the past five years and with at least two values in a disease group. These changes should eliminate erroneous ICD9 values but need investigation to determine potential effect on sensitivity and positive predictive value for disease identification. Further studies for broaden disease definitions may also be fruitful but need evaluation too before addition. For instance the inclusion of lab results of hyperglycemia may aide in DM identification or the use of steroid inhaler medications to identify chronic lung disease. Finally broadening of disease categories similar in composition to the Charlson-Deyo would increase the capture of potentially life threatening diseases that may be more common in younger populations and are not caught by the current four chronic disease subset. The benefit of this approach would be the potential use of points already established in the Charlson formulation. This type of modifications would require full statistical validation processes in both training and evaluation sets before use.

Future larger studies using retrospective approaches to determine risk by EHR data and then to determine outcome of death in four years can appropriately validate the discrimination and calibration of this tool in a general internal medicine population.

Potentially future application of the eLee may inform patient level clinical decision support and in aggregate population level risk adjustment.

APPENDIX A ICD9 for Lee Disease Determination

icd9	dz_process
494.1	LUNGDZ
494	LUNGDZ
494.0	LUNGDZ
492.1	LUNGDZ
492.2	LUNGDZ
492.3	LUNGDZ
492.4	LUNGDZ
492.5	LUNGDZ
202.80	CANCER
428.1	CHF
428.2	CHF
428.3	CHF
428.4	CHF
428.9	CHF
429.9	CHF
250.0	DM
250.00	DM
250	DM
250.01	DM
250.02	DM
250.03	DM
250.10	DM
250.1	DM
250.11	DM
250.12	DM
250.13	DM
250.20	DM
250.2	DM
250.21	DM
250.22	DM
250.23	DM
250.30	DM
250.3	DM
250.31	DM
250.32	DM
250.33	DM
250.40	DM
250.4	DM

250.41	DM
250.42	DM
250.43	DM
250.50	DM
250.5	DM
250.51	DM
250.52	DM
250.53	DM
250.60	DM
250.6	DM
250.61	DM
250.62	DM
250.63	DM
250.70	DM
250.7	DM
250.71	DM
250.72	DM
250.73	DM
250.80	DM
250.8	DM
250.81	DM
250.82	DM
250.83	DM
250.90	DM
250.9	DM
250.91	DM
250.92	DM
250.93	DM
151	CANCER
151.0	CANCER
151.1	CANCER
151.2	CANCER
151.3	CANCER
151.4	CANCER
151.5	CANCER
151.6	CANCER
151.8	CANCER
151.9	CANCER
152	CANCER
152.0	CANCER
152.1	CANCER

152.2	CANCER
152.3	CANCER
152.8	CANCER
152.9	CANCER
153	CANCER
153.1	CANCER
154	CANCER
154.1	CANCER
162	CANCER
162.0	CANCER
162.2	CANCER
162.3	CANCER
162.4	CANCER
162.5	CANCER
162.8	CANCER
162.9	CANCER
164	CANCER
164.0	CANCER
164.1	CANCER
164.2	CANCER
164.3	CANCER
164.8	CANCER
164.9	CANCER
189	CANCER
189.0	CANCER
189.1	CANCER
189.2	CANCER
189.3	CANCER
189.8	CANCER
189.9	CANCER
197	CANCER
197.0	CANCER
197.1	CANCER
197.2	CANCER
197.3	CANCER
197.4	CANCER
197.5	CANCER
197.6	CANCER
197.7	CANCER
197.8	CANCER
199	CANCER

199.0	CANCER
199.1	CANCER
200	CANCER
200.0	CANCER
200.1	CANCER
200.2	CANCER
200.3	CANCER
200.4	CANCER
200.5	CANCER
200.6	CANCER
200.7	CANCER
200.8	CANCER
201	CANCER
202	CANCER
202.0	CANCER
202.1	CANCER
202.2	CANCER
202.3	CANCER
202.4	CANCER
202.5	CANCER
202.6	CANCER
202.7	CANCER
202.8	CANCER
202.9	CANCER
203	CANCER
203.0	CANCER
204	CANCER
204.0	CANCER
204.1	CANCER
205	CANCER
205.0	CANCER
205.1	CANCER
206	CANCER
207	CANCER
207.0	CANCER
207.1	CANCER
207.2	CANCER
208	CANCER
209	CANCER
209.0	CANCER
209.1	CANCER

209.2	CANCER
209.3	CANCER
209.4	CANCER
209.5	CANCER
209.6	CANCER
230	CANCER
231	CANCER
232	CANCER
233	CANCER
234	CANCER
515	LUNGDZ
491	LUNGDZ
491.2	LUNGDZ
491.21	LUNGDZ
493	LUNGDZ
493.2	LUNGDZ
492	LUNGDZ
492.0	LUNGDZ
492.8	LUNGDZ
493.82	LUNGDZ
493.9	LUNGDZ
398.91	CHF
428.0	CHF
428.1	CHF
428.9	CHF
305.1	TOBACCOUSE
V15.82	TOBACCOUSE
174	CANCER
174.0	CANCER
174.1	CANCER
174.2	CANCER
174.3	CANCER
174.4	CANCER
174.5	CANCER
174.6	CANCER
174.8	CANCER
174.9	CANCER
185	CANCER
493.20	LUNGDZ
493.2	LUNGDZ
493.12	LUNGDZ

493.11	LUNGDZ
493.10	LUNGDZ
493.1	LUNGDZ
493.00	LUNGDZ
493.0	LUNGDZ
496	LUNGDZ
196	CANCER
196.0	CANCER
196.1	CANCER
196.2	CANCER
196.3	CANCER
196.5	CANCER
196.6	CANCER
196.8	CANCER
196.9	CANCER
153.0	CANCER
153.2	CANCER
153.3	CANCER
153.4	CANCER
153.5	CANCER
153.6	CANCER
153.7	CANCER
153.8	CANCER
153.9	CANCER
186	CANCER
186.0	CANCER
186.9	CANCER
193	CANCER
142	CANCER
142.0	CANCER
142.1	CANCER
142.2	CANCER
142.8	CANCER
142.9	CANCER
145	CANCER
145.0	CANCER
145.1	CANCER
145.2	CANCER
145.3	CANCER
145.4	CANCER
145.5	CANCER

145.6	CANCER
145.8	CANCER
145.9	CANCER
195	CANCER
195.0	CANCER
195.1	CANCER
195.2	CANCER
195.3	CANCER
195.4	CANCER
195.5	CANCER
195.8	CANCER
198	CANCER
198.0	CANCER
198.1	CANCER
198.2	CANCER
198.3	CANCER
198.4	CANCER
198.5	CANCER
198.6	CANCER
198.7	CANCER
198.8	CANCER
198.81	CANCER
198.82	CANCER
198.89	CANCER
147	CANCER
147.0	CANCER
147.1	CANCER
147.2	CANCER
147.3	CANCER
147.8	CANCER
147.9	CANCER
180	CANCER
180.0	CANCER
180.1	CANCER
180.8	CANCER
180.9	CANCER
172	CANCER
172.1	CANCER
172.2	CANCER
172.3	CANCER
172.4	CANCER

174.5	CANCER
172.6	CANCER
172.7	CANCER
172.8	CANCER
172.9	CANCER
141	CANCER
141.0	CANCER
141.1	CANCER
141.2	CANCER
141.3	CANCER
141.4	CANCER
141.5	CANCER
141.6	CANCER
141.8	CANCER
141.9	CANCER
146	CANCER
146.0	CANCER
146.1	CANCER
146.2	CANCER
146.3	CANCER
146.4	CANCER
146.5	CANCER
146.6	CANCER
146.7	CANCER
146.8	CANCER
146.9	CANCER
161	CANCER
161.0	CANCER
161.1	CANCER
161.2	CANCER
161.3	CANCER
161.8	CANCER
161.9	CANCER
201	CANCER
201.0	CANCER
201.00	CANCER
201.01	CANCER
201.02	CANCER
201.03	CANCER
201.04	CANCER
201.05	CANCER

201.06	CANCER
201.07	CANCER
201.08	CANCER
201.1	CANCER
201.10	CANCER
201.11	CANCER
201.12	CANCER
201.13	CANCER
201.14	CANCER
201.15	CANCER
201.16	CANCER
201.17	CANCER
201.18	CANCER
201.2	CANCER
201.20	CANCER
201.21	CANCER
201.22	CANCER
201.23	CANCER
201.24	CANCER
201.25	CANCER
201.26	CANCER
201.27	CANCER
201.28	CANCER
201.4	CANCER
201.40	CANCER
201.41	CANCER
201.42	CANCER
201.43	CANCER
201.44	CANCER
201.45	CANCER
201.47	CANCER
201.48	CANCER
201.5	CANCER
201.50	CANCER
201.51	CANCER
201.52	CANCER
201.53	CANCER
201.54	CANCER
201.55	CANCER
201.56	CANCER
201.57	CANCER

201.58	CANCER
201.6	CANCER
201.60	CANCER
201.61	CANCER
201.62	CANCER
201.63	CANCER
201.64	CANCER
201.65	CANCER
201.66	CANCER
201.67	CANCER
201.68	CANCER
201.7	CANCER
201.70	CANCER
201.71	CANCER
201.72	CANCER
201.73	CANCER
201.74	CANCER
201.75	CANCER
201.76	CANCER
201.77	CANCER
201.78	CANCER
201.9	CANCER
201.90	CANCER
201.91	CANCER
201.92	CANCER
201.93	CANCER
201.94	CANCER
201.95	CANCER
201.96	CANCER
201.97	CANCER
201.98	CANCER
493.21	LUNGDZ
493.22	LUNGDZ

APPENDIX B Charlson-Deyo ICD9 and Point Value Determinations

icd9	dz	point_value
410	MI	1
410.1	MI	1
410.2	MI	1
410.3	MI	1
410.4	MI	1
410.5	MI	1
410.6	MI	1
410.7	MI	1
410.8	MI	1
410.9	MI	1
411	MI	1
411.1	MI	1
411.89	MI	1
412	MI	1
429.7	MI	1
398.91	CHF	1
425.2	CHF	1
428	CHF	1
428	CHF	1
428.2	CHF	1
428.21	CHF	1
428.22	CHF	1
428.3	CHF	1
428.32	CHF	1
428.4	CHF	1
428.9	CHF	1
440.2	PVD	1
440.21	PVD	1
440.23	PVD	1
443.1	PVD	1
443.9	PVD	1
997.7	PVD	1
430	Cerebrovascular_dz	1
431	Cerebrovascular_dz	1
432	Cerebrovascular_dz	1
432.9	Cerebrovascular_dz	1
433	Cerebrovascular_dz	1
433	Cerebrovascular_dz	1

433.1	Cerebrovascular_dz	1
433.2	Cerebrovascular_dz	1
436	Cerebrovascular_dz	1
437	Cerebrovascular_dz	1
437	Cerebrovascular_dz	1
437.1	Cerebrovascular_dz	1
437.3	Cerebrovascular_dz	1
437.4	Cerebrovascular_dz	1
437.5	Cerebrovascular_dz	1
437.9	Cerebrovascular_dz	1
438	Cerebrovascular_dz	1
438.8	Cerebrovascular_dz	1
438.89	Cerebrovascular_dz	1
438.9	Cerebrovascular_dz	1
747.81	Cerebrovascular_dz	1
997.02	Cerebrovascular_dz	1
290	Dementia	1
290	Dementia	1
290.1	Dementia	1
290.2	Dementia	1
290.3	Dementia	1
290.41	Dementia	1
290.42	Dementia	1
294.1	Dementia	1
331.19	Dementia	1
331.82	Dementia	1
416.9	Chronic_Pulmonary_Dz	1
490	Chronic_Pulmonary_Dz	1
491	Chronic_Pulmonary_Dz	1
491.2	Chronic_Pulmonary_Dz	1
491.21	Chronic_Pulmonary_Dz	1
493	Chronic_Pulmonary_Dz	1
493	Chronic_Pulmonary_Dz	1
439.1	Chronic_Pulmonary_Dz	1
493.2	Chronic_Pulmonary_Dz	1
492	Chronic_Pulmonary_Dz	1
492	Chronic_Pulmonary_Dz	1
492.8	Chronic_Pulmonary_Dz	1
710.1	Connective_tissue_Dz	1
710.9	Connective_tissue_Dz	1
373.34	Connective_tissue_Dz	1

695.4	Connective_tissue_Dz	1
530.2	Ulcer_Dz	1
531	Ulcer_Dz	1
531	Ulcer_Dz	1
531.01	Ulcer_Dz	1
531.1	Ulcer_Dz	1
531.1	Ulcer_Dz	1
531.11	Ulcer_Dz	1
531.2	Ulcer_Dz	1
531.2	Ulcer_Dz	1
531.21	Ulcer_Dz	1
531.3	Ulcer_Dz	1
531.3	Ulcer_Dz	1
531.4	Ulcer_Dz	1
531.4	Ulcer_Dz	1
531.41	Ulcer_Dz	1
531.5	Ulcer_Dz	1
531.5	Ulcer_Dz	1
531.51	Ulcer_Dz	1
531.6	Ulcer_Dz	1
531.6	Ulcer_Dz	1
531.61	Ulcer_Dz	1
531.7	Ulcer_Dz	1
531.7	Ulcer_Dz	1
531.71	Ulcer_Dz	1
531.9	Ulcer_Dz	1
531.9	Ulcer_Dz	1
532	Ulcer_Dz	1
532	Ulcer_Dz	1
532.01	Ulcer_Dz	1
532.1	Ulcer_Dz	1
532.1	Ulcer_Dz	1
532.11	Ulcer_Dz	1
532.2	Ulcer_Dz	1
532.2	Ulcer_Dz	1
532.21	Ulcer_Dz	1
532.2	Ulcer_Dz	1
532.3	Ulcer_Dz	1
532.31	Ulcer_Dz	1
532.4	Ulcer_Dz	1
532.4	Ulcer_Dz	1

532.41	Ulcer_Dz	1
532.5	Ulcer_Dz	1
532.5	Ulcer_Dz	1
532.51	Ulcer_Dz	1
532.6	Ulcer_Dz	1
532.6	Ulcer_Dz	1
532.61	Ulcer_Dz	1
532.7	Ulcer_Dz	1
532.7	Ulcer_Dz	1
532.71	Ulcer_Dz	1
532.9	Ulcer_Dz	1
532.9	Ulcer_Dz	1
532.91	Ulcer_Dz	1
533	Ulcer_Dz	1
533	Ulcer_Dz	1
533.01	Ulcer_Dz	1
533.1	Ulcer_Dz	1
533.1	Ulcer_Dz	1
533.11	Ulcer_Dz	1
533.2	Ulcer_Dz	1
533.2	Ulcer_Dz	1
533.21	Ulcer_Dz	1
533.2	Ulcer_Dz	1
533.3	Ulcer_Dz	1
533.31	Ulcer_Dz	1
533.4	Ulcer_Dz	1
533.4	Ulcer_Dz	1
533.41	Ulcer_Dz	1
533.5	Ulcer_Dz	1
533.5	Ulcer_Dz	1
533.51	Ulcer_Dz	1
533.6	Ulcer_Dz	1
533.6	Ulcer_Dz	1
533.61	Ulcer_Dz	1
533.7	Ulcer_Dz	1
533.7	Ulcer_Dz	1
533.71	Ulcer_Dz	1
533.9	Ulcer_Dz	1
533.9	Ulcer_Dz	1
533.91	Ulcer_Dz	1
534	Ulcer_Dz	1

534	Ulcer_Dz	1
534.01	Ulcer_Dz	1
534.1	Ulcer_Dz	1
534.1	Ulcer_Dz	1
534.11	Ulcer_Dz	1
534.2	Ulcer_Dz	1
534.2	Ulcer_Dz	1
534.21	Ulcer_Dz	1
534.2	Ulcer_Dz	1
534.3	Ulcer_Dz	1
534.31	Ulcer_Dz	1
534.4	Ulcer_Dz	1
534.4	Ulcer_Dz	1
534.41	Ulcer_Dz	1
534.5	Ulcer_Dz	1
534.5	Ulcer_Dz	1
534.51	Ulcer_Dz	1
534.6	Ulcer_Dz	1
534.6	Ulcer_Dz	1
534.61	Ulcer_Dz	1
534.7	Ulcer_Dz	1
534.7	Ulcer_Dz	1
534.71	Ulcer_Dz	1
534.9	Ulcer_Dz	1
534.9	Ulcer_Dz	1
534.91	Ulcer_Dz	1
347.1	Cerebrovascular_dz	1
362.34	Cerebrovascular_dz	1
430	Cerebrovascular_dz	1
431	Cerebrovascular_dz	1
432	Cerebrovascular_dz	1
432.1	Cerebrovascular_dz	1
432.9	Cerebrovascular_dz	1
434.01	Cerebrovascular_dz	1
434.11	Cerebrovascular_dz	1
434.91	Cerebrovascular_dz	1
435	Cerebrovascular_dz	1
435	Cerebrovascular_dz	1
435.1	Cerebrovascular_dz	1
435.2	Cerebrovascular_dz	1
435.3	Cerebrovascular_dz	1

435.9	Cerebrovascular_dz	1
438	Cerebrovascular_dz	1
250	DM	1
250	DM	1
250	DM	1
250.01	DM	1
250.02	DM	1
250.03	DM	1
250.1	DM	1
250.1	DM	1
250.11	DM	1
250.12	DM	1
250.13	DM	1
250.2	DM	1
250.2	DM	1
250.21	DM	1
250.22	DM	1
250.23	DM	1
250.3	DM	1
250.3	DM	1
250.31	DM	1
250.32	DM	1
250.33	DM	1
196.2	Metastatic_Solid_Tumor	6
196	Metastatic_Solid_Tumor	6
456.3	Mod_Severe_Liver_Dz	3
456.2	Mod_Severe_Liver_Dz	3
456.1	Mod_Severe_Liver_Dz	3
572.8	Mod_Severe_Liver_Dz	3
572.3	Mod_Severe_Liver_Dz	3
250.8	DM	1
250.8	DM	1
250.81	DM	1
250.82	DM	1
250.83	DM	1
250.9	DM_end_Organ_Damage	2
250.9	DM_end_Organ_Damage	2
250.91	DM_end_Organ_Damage	2
250.92	DM_end_Organ_Damage	2
250.93	DM_end_Organ_Damage	2
342	Hemiplegia	2

342	Hemiplegia	2
342.1	Hemiplegia	2
342.8	Hemiplegia	2
342.9	Hemiplegia	2
343.1	Hemiplegia	2
343.4	Hemiplegia	2
344	Hemiplegia	2
438.2	Hemiplegia	2
438.2	Hemiplegia	2
438.5	Hemiplegia	2
585.3	Mod_Severe_Renal_Dz	2
585.4	Mod_Severe_Renal_Dz	2
585.5	Mod_Severe_Renal_Dz	2
585.6	Mod_Severe_Renal_Dz	2
585.9	Mod_Severe_Renal_Dz	2
586	Mod_Severe_Renal_Dz	2
587	Mod_Severe_Renal_Dz	2
588	Mod_Severe_Renal_Dz	2
582	Mod_Severe_Renal_Dz	2
581	Mod_Severe_Renal_Dz	2
581.1	Mod_Severe_Renal_Dz	2
581.2	Mod_Severe_Renal_Dz	2
581.8	Mod_Severe_Renal_Dz	2
581.81	Mod_Severe_Renal_Dz	2
581.89	Mod_Severe_Renal_Dz	2
581.9	Mod_Severe_Renal_Dz	2
250.4	DM_end_Organ_Damage	2
250.4	DM_end_Organ_Damage	2
250.41	DM_end_Organ_Damage	2
250.42	DM_end_Organ_Damage	2
250.43	DM_end_Organ_Damage	2
250.5	DM_end_Organ_Damage	2
250.5	DM_end_Organ_Damage	2
250.51	DM_end_Organ_Damage	2
250.52	DM_end_Organ_Damage	2
250.53	DM_end_Organ_Damage	2
250.6	DM_end_Organ_Damage	2
250.6	DM_end_Organ_Damage	2
250.61	DM_end_Organ_Damage	2
250.62	DM_end_Organ_Damage	2
250.63	DM_end_Organ_Damage	2

250.7	DM_end_Organ_Damage	2
250.7	DM_end_Organ_Damage	2
250.71	DM_end_Organ_Damage	2
250.72	DM_end_Organ_Damage	2
250.73	DM_end_Organ_Damage	2
151	tumor	2
151	tumor	2
151.1	tumor	2
151.2	tumor	2
151.3	tumor	2
151.4	tumor	2
151.5	tumor	2
151.6	tumor	2
151.8	tumor	2
151.9	tumor	2
152	tumor	2
152	tumor	2
152.1	tumor	2
152.2	tumor	2
152.3	tumor	2
152.8	tumor	2
152.9	tumor	2
153	tumor	2
153.1	tumor	2
154	tumor	2
154.1	tumor	2
162	tumor	2
162	tumor	2
162.2	tumor	2
162.3	tumor	2
162.4	tumor	2
162.5	tumor	2
162.8	tumor	2
162.9	tumor	2
164	tumor	2
164	tumor	2
164.1	tumor	2
164.2	tumor	2
164.3	tumor	2
164.8	tumor	2
164.9	tumor	2

189	tumor	2
189	tumor	2
189.1	tumor	2
189.2	tumor	2
189.3	tumor	2
189.8	tumor	2
189.9	tumor	2
209	tumor	2
209	tumor	2
209.1	tumor	2
209.2	tumor	2
209.3	tumor	2
209.4	tumor	2
209.5	tumor	2
209.6	tumor	2
230	tumor	2
230.1	tumor	2
230.2	tumor	2
230.3	tumor	2
230.4	tumor	2
230.5	tumor	2
230.6	tumor	2
230.7	tumor	2
230.8	tumor	2
230.9	tumor	2
231	tumor	2
231.1	tumor	2
231.2	tumor	2
231.8	tumor	2
231.9	tumor	2
232	tumor	2
232	tumor	2
232.1	tumor	2
232.2	tumor	2
232.3	tumor	2
232.4	tumor	2
232.5	tumor	2
232.6	tumor	2
232.7	tumor	2
232.8	tumor	2
232.9	tumor	2

233	tumor	2
233	tumor	2
233.1	tumor	2
233.2	tumor	2
233.3	tumor	2
233.3	tumor	2
233.31	tumor	2
233.39	tumor	2
233.4	tumor	2
233.5	tumor	2
233.6	tumor	2
233.7	tumor	2
233.9	tumor	2
234	tumor	2
234	tumor	2
234.8	tumor	2
203	tumor	2
203	tumor	2
204	tumor	2
204	tumor	2
204.1	tumor	2
205	tumor	2
205	tumor	2
205.1	tumor	2
206	tumor	2
207	tumor	2
207	tumor	2
207.1	tumor	2
207.2	tumor	2
208	tumor	2
200	tumor	2
200	tumor	2
200.1	tumor	2
200.2	tumor	2
200.3	tumor	2
200.4	tumor	2
200.5	tumor	2
200.6	tumor	2
200.7	tumor	2
200.8	tumor	2
201	tumor	2

201	tumor	2
201	tumor	2
201.01	tumor	2
201.02	tumor	2
201.03	tumor	2
201.04	tumor	2
201.05	tumor	2
201.06	tumor	2
201.07	tumor	2
201.08	tumor	2
201.1	tumor	2
201.1	tumor	2
201.11	tumor	2
201.12	tumor	2
201.13	tumor	2
201.14	tumor	2
201.15	tumor	2
201.16	tumor	2
201.17	tumor	2
201.18	tumor	2
201.2	tumor	2
201.2	tumor	2
201.21	tumor	2
201.22	tumor	2
201.23	tumor	2
201.24	tumor	2
201.25	tumor	2
201.26	tumor	2
201.27	tumor	2
201.28	tumor	2
201.4	tumor	2
201.4	tumor	2
201.41	tumor	2
201.42	tumor	2
201.43	tumor	2
201.44	tumor	2
201.45	tumor	2
201.46	tumor	2
201.47	tumor	2
201.48	tumor	2
201.5	tumor	2

201.5	tumor	2
201.51	tumor	2
201.52	tumor	2
201.53	tumor	2
201.54	tumor	2
201.55	tumor	2
201.56	tumor	2
201.57	tumor	2
201.58	tumor	2
201.6	tumor	2
201.6	tumor	2
201.61	tumor	2
201.62	tumor	2
201.63	tumor	2
201.64	tumor	2
201.65	tumor	2
201.66	tumor	2
201.67	tumor	2
201.68	tumor	2
201.7	tumor	2
201.7	tumor	2
201.71	tumor	2
201.72	tumor	2
201.73	tumor	2
201.74	tumor	2
201.75	tumor	2
201.76	tumor	2
201.77	tumor	2
201.78	tumor	2
201.9	tumor	2
201.9	tumor	2
201.91	tumor	2
201.92	tumor	2
201.93	tumor	2
201.94	tumor	2
201.95	tumor	2
201.96	tumor	2
201.97	tumor	2
201.98	tumor	2
202	tumor	2
202	tumor	2

202.1	tumor	2
202.2	tumor	2
202.3	tumor	2
202.4	tumor	2
202.5	tumor	2
202.6	tumor	2
202.7	tumor	2
202.8	tumor	2
202.9	tumor	2
456.2	Mod_Severe_Liver_Dz	3
522.8	Mod_Severe_Liver_Dz	3
571	Mod_Severe_Liver_Dz	3
571.3	Mod_Severe_Liver_Dz	3
571.5	Mod_Severe_Liver_Dz	3
571.8	Mod_Severe_Liver_Dz	3
571.9	Mod_Severe_Liver_Dz	3
572	Mod_Severe_Liver_Dz	3
573	Mod_Severe_Liver_Dz	3
573.1	Mod_Severe_Liver_Dz	3
573.9	Mod_Severe_Liver_Dz	3
197	Metastatic_Solid_Tumor	6
197	Metastatic_Solid_Tumor	6
197.1	Metastatic_Solid_Tumor	6
197.2	Metastatic_Solid_Tumor	6
197.3	Metastatic_Solid_Tumor	6
197.4	Metastatic_Solid_Tumor	6
197.5	Metastatic_Solid_Tumor	6
197.6	Metastatic_Solid_Tumor	6
197.7	Metastatic_Solid_Tumor	6
197.8	Metastatic_Solid_Tumor	6
198	Metastatic_Solid_Tumor	6
198	Metastatic_Solid_Tumor	6
198.1	Metastatic_Solid_Tumor	6
198.2	Metastatic_Solid_Tumor	6
198.3	Metastatic_Solid_Tumor	6
198.4	Metastatic_Solid_Tumor	6
198.5	Metastatic_Solid_Tumor	6
198.6	Metastatic_Solid_Tumor	6
198.7	Metastatic_Solid_Tumor	6
198.8	Metastatic_Solid_Tumor	6
198.81	Metastatic_Solid_Tumor	6

198.82	Metastatic_Solid_Tumor	6
198.89	Metastatic_Solid_Tumor	6
199	Metastatic_Solid_Tumor	6
199	Metastatic_Solid_Tumor	6
199.1	Metastatic_Solid_Tumor	6
42	AIDS	6
V08	AIDS	6
496	Chronic_Pulmonary_Dz	1
500	Chronic_Pulmonary_Dz	1
501	Chronic_Pulmonary_Dz	1
502	Chronic_Pulmonary_Dz	1
503	Chronic_Pulmonary_Dz	1
504	Chronic_Pulmonary_Dz	1
505	Chronic_Pulmonary_Dz	1
506.4	Chronic_Pulmonary_Dz	1
710.1	Connective_tissue_Dz	1
710.1	Connective_tissue_Dz	1
710.2	Connective_tissue_Dz	1
710.3	Connective_tissue_Dz	1
710.4	Connective_tissue_Dz	1
710.5	Connective_tissue_Dz	1
710.8	Connective_tissue_Dz	1
714	Connective_tissue_Dz	1
714.1	Connective_tissue_Dz	1
714.2	Connective_tissue_Dz	1
714.81	Connective_tissue_Dz	1
725	Connective_tissue_Dz	1
196.1	Metastatic_Solid_Tumor	6
196	Metastatic_Solid_Tumor	6
456.21	Mod_Severe_Liver_Dz	3
456.2	Mod_Severe_Liver_Dz	3
456	Mod_Severe_Liver_Dz	3
572.4	Mod_Severe_Liver_Dz	3
572.2	Mod_Severe_Liver_Dz	3
588	Mod_Severe_Renal_Dz	2
588.1	Mod_Severe_Renal_Dz	2
588.8	Mod_Severe_Renal_Dz	2
588.81	Mod_Severe_Renal_Dz	2
588.89	Mod_Severe_Renal_Dz	2
588.9	Mod_Severe_Renal_Dz	2

196.3	Metastatic_Solid_Tumor	6
196.5	Metastatic_Solid_Tumor	6
196.6	Metastatic_Solid_Tumor	6
196.8	Metastatic_Solid_Tumor	6
196.9	Metastatic_Solid_Tumor	6

APPENDIX C Survey

Health Validation Study survey:

Participant Number: ___

Please circle or short answer this survey.

1. Age _____
2. Gender Male/Female
3. Weight _____

   Height _____
4. Has a doctor ever told you that you have diabetes or high blood sugar including borderline diabetes? (Yes / No)
5. Has a doctor ever told you that you that you have cancer or a malignant tumor, excluding minor skin cancers unless it was melanoma? (Yes / No)
6. Do you have a chronic lung disease (Emphysema/Chronic Bronchitis) that limits your usual activities or makes you need oxygen at home? (Yes / No)
7. Has a doctor ever told you that you have congestive heart failure? (Yes / No)
8. Have you smoked cigarettes in the past week? (Yes / No)

   If no:
   a. Never smoked (Less than 100 cigarettes in your entire life) (Yes)
   b. Former smoker? (Yes)
9. Because of a health or memory problem, do you have any difficulty with bathing or showering? (Yes / No)
10. Because of a health or memory problem, do you have any difficulty with managing your money- such as paying your bills and keeping track of expenses? (Yes / No)
11. Because of a health problem do you have any difficulty with walking several blocks? (Yes / No)

12. Because of a health or memory problem, do you have any difficulty with pulling or pushing large objects like a living room chair? (Yes / No)

13. In general, would you say your health is
    a. Excellent
    b. Very good
    c. Good
    d. Fair
    e. Poor

14. During the past 12 months, how many times were you hospitalized overnight?
    a. None
    b. Once
    c. Twice or more

REFERENCES

1. Werner RM, Asch DA. The unintended consequences of publicly reporting quality information. JAMA. 2005 Mar 9;293(10):1239–44.

2. Scholle SH, Roski J, Adams JL, Dunn DL, Kerr EA, Dugan DP, et al. Benchmarking physician performance: reliability of individual and composite measures. Am J Manag Care. 2008 Dec;14(12):833–8.

3. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. JAMA. 1999 Jun 9;281(22):2098–105.

4. Boyd CM. Clinical Practice Guidelines and Quality of Care for Older Patients With Multiple Comorbid Diseases: Implications for Pay for Performance. JAMA: The Journal of the American Medical Association. 2005 Aug 10;294(6):716–24.

5. Horn SD, Chachich B, Clopton C. Measuring Severity of Illness: A Reliability Study. Medical Care. 1983 Jul 1;21(7):705–14.

6. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit. Care Med. 1985 Oct;13(10):818–29.

7. 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):373–83.

8. Parkerson GR, Eugene Broadhead W, Chiu-Kit J, TSE. The Duke severity of illness checklist (DUSOI) for measurement of severity and comorbidity. Journal of Clinical Epidemiology. 1993 Apr;46(4):379–93.

9. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992 Jun;45(6):613–9.

10. Carey EC, Walter LC, Lindquist K, Covinsky KE. Development and validation of a functional morbidity index to predict mortality in community-dwelling elders. J Gen Intern Med. 2004 Oct;19(10):1027–33.

11. Covinsky KE, Justice AC, Rosenthal GE, Palmer RM, Landefeld CS. Measuring prognosis and case mix in hospitalized elders. The importance of functional status. J Gen Intern Med. 1997 Apr;12(4):203–8.

12. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. J Gen Intern Med. 2009 Oct;24(10):1115–22.

13. Hogan C, Lunney J, Gabel J, Lynn J. Medicare beneficiaries' costs of care in the last year of life. Health Aff (Millwood). 2001 Aug;20(4):188–95.

14. Starfield B, Lemke KW, Bernhardt T, Foldes SS, Forrest CB, Weiner JP. Comorbidity: Implications for the Importance of Primary Care in “Case” Management. Ann Fam Med. 2003 May;1(1):8–14.

15. Parkerson GR, Broadhead WE, Tse C-KJ. Health Status and Severity of Illness as Predictors of Outcomes in Primary Care. Medical Care. 1995 Jan 1;33(1):53–66.

16. Parkerson GR, Hammond WE, Yarnall KSH. Feasibility and Potential Clinical Usefulness of a Computerized Severity of Illness Measure. Arch Fam Med. 1994 Nov 1;3(11):968–73.

17. Reid RJ, MacWilliam L, Verhulst L, Roos N, Atkinson M. Performance of the ACG case-mix system in two Canadian provinces. Med Care. 2001 Jan;39(1):86–99.

18. Halling A, Fridh G, Ovhed I. Validating the Johns Hopkins ACG Case-Mix System of the elderly in Swedish primary health care. BMC Public Health. 2006;6:171.

19. Gibson RF, Haug PJ, Horn SD. Lessons from evaluating an automated patient severity index. J Am Med Inform Assoc. 1996 Oct;3(5):349–57.

20. Deyo RA, Carter WB. Strategies for Improving and Expanding the Application of Health Status Measures in Clinical Settings: A Researcher-Developer Viewpoint. Medical Care. 1992 May 1;30(5):MS176–MS186.

21. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care. 1998 Jan;36(1):8–27.

22. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. JAMA. 2006 Feb 15;295(7):801–8.

23. Ravaglia G, Forti P, Lucicesare A, Pisacane N, Rietti E, Patterson C. Development of an easy prognostic score for frailty outcomes in the aged. Age and Ageing. 2008 Mar 1;37(2):161 –166.

24. Horn SD, Horn RA. Reliability and validity of the Severity of Illness Index. Med Care. 1986 Feb;24(2):159–78.

25. Horn SD, Horn RA, Moses H. Profiles of physician practice and patient severity of illness. Am J Public Health. 1986 May;76(5):532–5.

26. Averill RF, McGuire TE, Manning BE, Fowler DA, Horn SD, Dickson PS, et al. A study of the relationship between severity of illness and hospital cost in New Jersey hospitals. Health Serv Res. 1992 Dec;27(5):587–12.

27. Horn SD, Sharkey PD, Buckle JM, Backofen JE, Averill RF, Horn RA. The Relationship between Severity of Illness and Hospital Length of Stay and Mortality. Medical Care. 1991 Apr 1;29(4):305–17.

28. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. J Clin Epidemiol. 1994 Nov;47(11):1245–51.

29. Spickard A 3rd, Gigante J, Stein G, Denny JC. Automatic capture of student notes to augment

mentor feedback and student performance on patient write-ups. J Gen Intern Med. 2008 Jul;23(7):979–84.

30. Lash TL, Silliman RA. A comparison of the National Death Index and Social Security Administration databases to ascertain vital status. Epidemiology. 2001 Mar;12(2):259–61.

31. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008 Jan 30;27(2):157–172; discussion 207–212.

32. Kleinbaum DG. Epidemiologic methods: the "art" in the state of the art. Journal of Clinical Epidemiology. 2002 Dec;55(12):1196–200.

33. Andrew P. B. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition. 1997 Jul;30(7):1145–59.

34. Blumenthal D. Launching HITECH. New England Journal of Medicine. 2010 Feb 4;362(5):382–5.

35. Kahn KL, Tisnado DM, Adams JL, Liu H, Chen W-P, Hu FA, et al. Does ambulatory process of care predict health-related quality of life outcomes for patients with chronic disease? Health Serv Res. 2007 Feb;42(1 Pt 1):63–83.

36. Spertus JA, Jones P, McDonell M, Fan V, Fihn SD. Health status predicts long-term outcome in outpatients with coronary disease. Circulation. 2002 Jul 2;106(1):43–9.

37. Synnestvedt MB, Weiner MG. Evaluation of the use of electronic health data to classify four-year mortality risk for older adults undergoing screening colonoscopies. AMIA Annu Symp Proc. 2011;2011:1361–5.

38. Benesch C, Witter DM, Wilder AL, Duncan PW, Samsa GP, Matchar DB. Inaccuracy of the International Classification of Diseases (ICD-9-CM) in Identifying the Diagnosis of Ischemic Cerebrovascular Disease. Neurology. 1997 Sep 1;49(3):660–4.

39. Butler RN. Population aging and health. BMJ. 1997 Oct 25;315(7115):1082–4.

40. Walter LC, Covinsky KE. Cancer Screening in Elderly Patients. JAMA: The Journal of the American Medical Association. 2001 Jun 6;285(21):2750 –2756.

41. Sima CS. Cancer Screening Among Patients With Advanced Cancer. JAMA: The Journal of the American Medical Association. 2010 Oct 13;304(14):1584.

42. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. JAMA. 2012 Jan 11;307(2):182–92.

43. Stein BD, Bautista A, Schumock GT, Lee TA, Charbeneau JT, Lauderdale DS, et al. The Validity of International Classification of Diseases, Ninth Revision, Clinical Modification Diagnosis Codes for Identifying Patients Hospitalized for COPD Exacerbations. Chest. 2012 Jan 1;141(1):87–93.

44. Allen LA, Yood MU, Wagner EH, Aiello Bowles EJ, Pardee R, Wellman R, et al. Performance of

Claims-based Algorithms for Identifying Heart Failure and Cardiomyopathy Among Patients Diagnosed With Breast Cancer. Medical Care. 2012 May;1.

45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics. 2009 Apr;42(2):377–81.

46. Gazmararian JA, Williams MV, Peel J, Baker DW. Health literacy and knowledge of chronic disease. Patient Education and Counseling. 2003 Nov;51(3):267–75.

47. Hinton L, Franz C, Reddy G, Flores Y, Kravitz R, Barker J. Practice Constraints, Behavioral Problems, and Dementia Care: Primary Care Physicians' Perspectives. Journal of General Internal Medicine. 2007;22(11):1487–92.