

*Med Care*. Author manuscript; available in PMC 2015 October 24.

Published in final edited form as:

*Med Care*. 2011 October ; 49(10): 932–939. doi:10.1097/MLR.0b013e318215d5e2.

## Using the Johns Hopkins Aggregated Diagnosis Groups (ADGs) to predict mortality in a general adult population cohort in Ontario, Canada

Peter C. Austin, PhD<sup>(1),(2),(3)</sup>, Carl van Walraven, MD, MSc<sup>(1),(5),(6)</sup>, Walter P. Wodchis, PhD<sup>(1),(2),(4)</sup>, Alice Newman, MSc<sup>(1)</sup>, and Geoffrey M. Anderson, MD, PhD<sup>(1),(2)</sup>

<sup>(1)</sup>Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada

<sup>(2)</sup>Department of Health Management, Policy and Evaluation, University of Toronto

<sup>(3)</sup>Dalla Lana School of Public Health, University of Toronto

<sup>(4)</sup>Toronto Rehabilitation Institute, Toronto, Ontario

<sup>(5)</sup>Ottawa Hospital Research Institute, Ottawa, Ontario

<sup>(6)</sup>Department of Medicine, University of Ottawa, Ottawa, Ontario

### Abstract

**Background**—Administrative health care databases are increasingly used for health services and comparative effectiveness research. When comparing outcomes between different treatments, interventions or exposures, the ability to adjust for differences in the risk of the outcome occurring between treatment groups is important. Similarly, when conducting health care provider profiling, adequate risk-adjustment is necessary for conclusions about provider performance to be valid. There are limited validated methods for risk-adjustment in ambulatory populations using administrative health care databases.

**Objectives**—To examine the ability of the Johns Hopkins' Aggregated Diagnosis Groups (ADGs) to predict mortality in a general ambulatory population cohort.

**Research Design**—Retrospective cohort constructed using population-based administrative data.

**Subjects**—All 10,498,413 residents of Ontario, Canada between the ages of 20 and 100 years who were alive on their birthday in 2007. Subjects were randomly divided in derivation and validation samples.

**Measures**—Death within one year of the subject's birthday in 2007.

**Results**—A logistic regression model consisting of age, sex, and indicator variables for 28 of the 32 ADG categories had excellent discrimination: the c-statistic (equivalent to the area under the ROC curve) was 0.917 in both derivation and validation samples. Furthermore, the model demonstrated very good calibration. In comparison, the use of the Charlson comorbidity index or

the Elixhauser comorbidities resulted in a minor decrease in discrimination compared to the use of the ADGs.

**Conclusions**—Logistic regression models using age, sex, and the John Hopkins ADGs were able to accurately predict one-year mortality in a general ambulatory population of subjects.

### Keywords

comorbidity; administrative data; Aggregated Diagnosis Groups; Adjusted Clinical Groups; health services research; Charlson comorbidity index; Elixhauser comorbidities

## 1. Introduction

The ability to characterize the comorbidity burden of a population is of great importance in many areas of health services and comparative effectiveness research. When using observational or non-randomized studies to compare outcomes between subjects receiving different treatments, exposures, or interventions, the ability to adjust for systematic differences in outcome risk between treatment groups can reduce bias when comparing outcomes between treatment groups. Furthermore, adjusting for comorbidity burden allows for valid risk-adjusted estimates of the performance of different health care providers [1].

Several methods have been derived for summarizing the comorbidity burden of a patient population using administrative data. Charlson et al. derived and validated a weighted index of comorbidities for predicting mortality in hospitalized general medical patients, which was subsequently adapted by Deyo et al. for use with the International Classification of Diseases (ICD-9-CM) diagnosis and procedure codes that are frequently used in electronic health care administrative data [2,3]. The use of the Deyo-Charlson comorbidity index for risk-adjustment is ubiquitous in health services research. Similarly, Elixhauser et al. developed a method to classify comorbidities in hospitalized patients using diagnoses coded using the ICD-9-CM diagnosis codes in administrative data [4]. Both of these schemes have been updated for use with the ICD-10 diagnosis classification scheme [5,6]. Schneeweiss et al. proposed that the number of unique drugs prescribed be used for risk-adjustment purposes [7]. Another method used for risk-adjustment is the Chronic Disease Score, which uses outpatient pharmacy records [8]. A limitation of the former two approaches is their reliance on hospitalization records, which limits their utility in ambulatory populations of patients. A limitation of the latter two methods is their use of prescription records. In many jurisdictions, data on drug prescribing are not available for the entire population.

The Johns Hopkins Adjusted Clinical Groups (ACGs)<sup>®</sup> are a person-focused, diagnosis-based method of categorizing subjects' illnesses. The ACG system assigns each International Classification of Disease (ICD) codes (-9 version, -9-CM version, or -10 version) to one of 32 diagnosis clusters known as Aggregated Diagnosis Groups (ADG). Individual diseases or conditions are placed into a single ADG based on five clinical dimensions: duration of the condition; severity of the condition; diagnostic certainty; etiology of the condition; and specialty care involvement [9–12]. ICD codes within the same ADG are similar in both clinical criteria and expected need for healthcare resource. Each individual may have diagnoses belonging to between zero and 32 ADGs. The 32 ADGs can

be collapsed into 12 Collapsed ADGs (CADG). As with the ADGs, a given diagnosis belongs in only one CADG; however, subjects can have multiple diagnoses, each of which can be within different CADGs. Finally, subjects are assigned to exactly one of 106 ACGs. Subjects within the same ACG are expected to have similar healthcare resource utilization. The reader is referred elsewhere for a greater discussion of the ADG and ACG methodology [9–12]. Importantly, the ADG/ACG definitions do not rely solely on the use inpatient health administrative data, but also use data contained in ambulatory health care data. Thus, ICD diagnosis codes obtained from physician billing claims can be used in addition to ICD diagnosis codes contained in electronic hospital discharge abstracts.

The Johns Hopkins ADGs and ACGs were developed for predicting health care resource utilization. Several studies have examined the ability of these classifications to predict health care use. However, there is a paucity of research into the ability of these comorbidity classification schemes to predict mortality. A few studies have examined the ability of the Johns Hopkins ACGs to predict mortality [13–17]. Depending on the patient population and the duration of follow-up for determining mortality, c-statistics for models that included the ACGs in addition to age and/or sex ranged from 0.701 to 0.768. To the best of our knowledge, no studies have examined the ability of the ADG comorbidity classification scheme to predict mortality. There are fewer ADGs than there are categories in the ACG classification. There are two potential advantages to the use of ADGs to predict mortality compared to using ACGs. First, there is the potential for more parsimonious regression models. Second, since patient risk may be related to multiple conditions and there are  $2^{32}$  possible combinations of ADGs, use of the ADGs may permit more accurate mortality prediction compared to the use of the ACGs.

There is increasing interest in using administrative health care data to compare the effects of treatments, interventions, or exposures in non-hospitalized or ambulatory populations. Mortality is an outcome that is frequently used in health services and comparative effectiveness research. The objective of the current study was to determine whether ADGs can be used to accurately predict mortality in a general adult population cohort.

## 2. Methods

### 2.1 Data sources

In the Canadian province of Ontario, all medically necessary services are provided within a single-payor public health care system, with no parallel private system. The Ontario Health Insurance Plan (OHIP) is a government-funded universal health insurance program that funds physician services, while hospital, long term care, and home care services are funded by the Ministry of Health and Long-Term Care. These services are provided to all residents of Ontario, without deductibles or co-payments. Furthermore, prescription drug coverage is provided to all residents over the age of 65 years as well as to those on social assistance.

We used four different population-based administrative health care databases that were linked by encrypted health number. The Registered Persons Database (RPDB) contains basic demographic information on all Ontarians who were ever eligible for Ontario's universal health care insurance program. The RPDB contains information on each resident's date of

birth, sex, and date of death (if applicable). The Canadian Institute for Health Information (CIHI) Discharge Abstract Database (DAD) contains information on all inpatient hospitalizations in the province of Ontario. For each hospitalization record, there are 25 fields for recording diagnoses made on the patient during the course of the hospitalization. Since 2002, diagnoses have been coded using the International Classification of Disease, 10<sup>th</sup> Revision (ICD-10) coding scheme. The Ontario Health Insurance Plan (OHIP) physician billing database contains billing claims submitted by Ontario physicians to the provincial universal health insurance program. Each claim contains a fee code describing the type of service provided, and a diagnosis code denoting a reason for the service. The diagnosis field is coded using a truncated version of the ICD-9 coding scheme [18]. The Ontario Mental Health Reporting System (OMHRS) collects data on patients in adult-designated inpatient mental health beds. This includes beds in general, provincial psychiatric, and specialty psychiatric facilities. The OMHRS contains data on reasons for admission and for discharge and on psychiatric and non-psychiatric diagnoses.

## 2.2 Study subjects

The study sample consisted of all subjects in the RPDB who were alive and eligible on their birthday in 2007. Each subject's birthday in 2007 served as the subject-specific index date. We excluded subjects who were aged less than 20 years or older than 100 years on the index date.

For each subject we determined whether they died within the 365 days following their index date. For each subject, we identified all diagnoses associated with all hospital admissions from the CIHI DAD and all physician billing claims in the OHIP database for physician services provided in the two years prior to the index date. For each subject, we used the Johns Hopkins ACG® software program to collapse these diagnoses to the 32 ADGs. Thus, for each subject, we determined whether an ICD diagnosis code within each of the 32 ADGs had occurred in the two years prior to the index date.

## 2.3 Statistical Methods

For each of the 32 ADGs, we compared the probability of death within 365 days of the index date for those with and without diagnoses in the given ADG using the Chi-squared test.

In order to evaluate the performance of ADG-based methods for predicting mortality in a sample that was independent of the sample in which regression models were derived, we used a random number generator to divide the overall sample into approximately equally sized derivation and validation samples. Using the subjects in the derivation sample, we used logistic regression to regress the occurrence of death within 365 days of the index date on age (assuming a linear relationship between age and the log-odds of death), sex, and 32 indicator variables representing the presence or absence of the 32 ADGs. This model will be called the 'full logistic regression model'. Backwards variable elimination with a significance level of 0.05 for variable retention was used to develop a parsimonious logistic regression model for predicting mortality. The resultant model will be called the "final logistic regression model".

Model discrimination in the derivation sample was assessed using the c-statistic [19]. We used the final logistic regression model to obtain the predicted probability of mortality for each subject in the validation sample. The predictive accuracy of the model developed in the derivation sample was assessed in the validation sample using the c-statistic.

Model calibration was assessed in several manners. First, using the final logistic regression model, predicted probabilities of mortality were obtained for each subject in the validation sample. The validation sample was divided into 100 approximately equally sized groups using the centiles of predicted probability of mortality (each centile in the validation sample consisted of approximately 52,500 subjects). Within each of the 100 groups in the validation sample, we determined both the mean predicted probability of mortality based on the final logistic regression model and the observed probability of mortality amongst subjects in that group. The relationship between observed and predicted mortality was then examined graphically. Second, calibration-in-the-large was determined [20]. Calibration-in-the-large compares the mean predicted probability of death in the validation sample with the observed probability of mortality in the validation sample. Third, we determined the calibration slope (deviation of the calibration slope from unity denotes miscalibration) [20]. To do so, we used logistic regression to regress the occurrence of death within one year of the index date in the validation sample on the linear predictor of mortality obtained using the regression coefficients from the final logistic regression model (estimated in the derivation sample) applied to the subjects in the validation sample. In order to determine whether the performance of the full logistic regression was solely due to age and sex, we repeated the above process with a regression model that included only age and sex. We also repeated the above process with a regression model that only contained indicator variables for the 32 ADGs and that excluded age and sex.

Analyses in health services research and pharmacoepidemiology are often restricted to those over the age of 65 years, since these subjects are eligible for Medicare in the United States and for provincial drug insurance coverage in Ontario. In order to determine whether the final logistic regression model performed differently in different age strata, we stratified each of the derivation and validation samples into two strata. The first consisted of subjects aged less than 65 years, while the second consisted of subjects aged 65 years and over. Within each of the two strata in the derivation sample we re-estimated the coefficients for the final logistic regression model. We then obtained predictions of the probability of death within 365 days for each subject in each of the two strata in the validation sample. The predictive accuracy of the final regression model was assessed in each stratum in the validation sample using the c-statistic.

Health services use is heavy during the final year of life [21–22]. We conducted a sensitivity analysis to address the concern that increased health services use during the last year of life would give rise to the opportunity for greater documentation of comorbidity. We excluded subjects who died within 365 days of their index date. In the sample of all subjects who survived for at least one year from their index date, we examined the ability of our final logistic regression model to predict the occurrence of death between 366 and 730 days following the index date. Methods similar to those described above were used for this sensitivity analysis.

Finally, for comparative purposes, we examined the ability of two alternative comorbidity coding schemes to predict 1-year mortality in our overall population cohort. First, we calculated the Charlson comorbidity index [3,5] and the Elixhauser comorbidities [4,5] using data from hospitalization occurring in the two years prior to the index date (alternatively, researchers may include out-patient records in addition to in-patient records when determining whether a given comorbidity was present [23]). Diagnoses for coding the Charlson comorbidity index were obtained from the CIHI DAD. For the Elixhauser comorbidities, diagnoses were obtained from the CIHI DAD. For mental and addiction Elixhauser comorbidities, the OMHRS database was also used to identify occurrences of the given diagnoses. Subjects who had not been hospitalized in the previous two years had their Charlson score set to zero. Similarly, these subjects had their values of each of the 30 Elixhauser comorbidities set to absent. For the Charlson comorbidity score, a logistic regression model to predict the probability of one-year mortality using the Charlson comorbidity index and age and sex was developed in the derivation sample. For the Elixhauser comorbidities, the coefficients for a logistic regression model with age, sex, and indicator variables for the 30 Elixhauser comorbidities were estimated in the derivation sample. The accuracy of each of these two models was assessed in the validation sample using the c-statistic.

### 3. Results

The study sample consisted of 10,498,413 subjects between the ages of 20 and 100 years. The median age was 46 (25<sup>th</sup> and 75<sup>th</sup> percentiles: 34 and 59, respectively). Women comprised 51% of the study sample. The prevalence of each of the 32 ADGs in the study sample is described in Table 1. The prevalence of the individual ADGs ranged from a low of 0.5% (ADG: See and Reassure) to a high of 43.8% (ADG: Signs/Symptoms: Uncertain). Note that since the ADGs are not mutually exclusive, subjects with a diagnosis in a given ADG can also have diagnoses within other ADGs. While 43.8% of subjects had at least one diagnosis within the latter ADG (ADG: Signs/Symptoms: Uncertain), only 0.8% of subjects had diagnoses that lay solely within this ADG. The number of distinct ADGs in which subjects had diagnoses ranged from 0 (15.4% of subjects) to 25 (< 6 subjects). The median number of distinct ADG categories was 4 (the 25<sup>th</sup> and 75<sup>th</sup> percentiles were 2 and 7, respectively), while 95% of subjects had diagnoses in 11 or fewer ADGs.

Overall, 85,007 (0.81%) subjects died within 365 days of their index date. The 1-year mortality rate for those with and without each ADG is described in Table 1. For all 32 of ADGs, there was a statistically significant difference in the probability of mortality between those with and without the ADG ( $P < 0.001$ ). The statistical significance of many of these associations may be driven primarily by the very large size of our population sample.

The predictive accuracy of the different regression models are summarized in Table 2. The c-statistic of the full logistic regression model (age, sex, and indicator variables for the 32 ADGs) was 0.917 in both the derivation and validation samples. The c-statistic of regression model with only age and sex as predictor variables was 0.883 in both the derivation and validation samples. The c-statistics of the logistic regression that contained only indicator variables for the 32 ADGs was 0.866 and 0.864 in the derivation and validation samples,

respectively. The final logistic regression model contained 30 covariates: age, sex, and indicator variables for 28 ADGs. The c-statistic of the final logistic regression model was 0.917 in both the derivation and validation samples. For comparative purposes, the c-statistic of the logistic regression that used the Charlson comorbidity index and age and sex was 0.906 in both the derivation and validation samples. The c-statistic of the logistic regression model that used age, sex, and the 30 Elixhauser comorbidities was 0.909 in both the derivation and validation samples.

To assess the sensitivity of this finding to the particular split of the sample, we repeatedly split the original sample into derivation and validation samples 100 times. In each derivation sample we estimated the coefficients of the final regression model and applied these coefficients to the corresponding validation sample. The c-statistics ranged from 0.915 to 0.919, with a median of 0.917.

The adjusted odds ratios for the association between age, sex, and the 28 ADGs and 1-year mortality in the derivation sample are reported in Table 3. The adjusted odds ratios for the 28 ADGs ranged from a low of 0.656 (ADG: Chronic specialty: Unstable-Ear, Nose, Throat) to a high of 2.880 (ADG: Psychosocial: Recurrent or Persistent, Unstable).

The final logistic regression model predicted probabilities of 1-year mortality for each subject in the derivation sample. These predicted probabilities ranged from a low of 0.000023 to a high of 0.8975. The relationship between the observed probability of death and the mean predicted probability of death across the 100 strata determined by the centiles of predicted probability of death in the validation sample is described in the left panel of Figure 1. A dotted diagonal line has been superimposed on the figure. Points on this diagonal line denote perfect concordance between observed and predicted mortality. In the 94 subgroups with the lowest mean predicted probability of death, there was almost perfect concordance between the mean predicted probability of death and the proportion of subjects that died within one year of the index date. In all but the top three strata, the absolute difference between the observed probability of death and the mean predicted probability of death was less than 0.01. In the highest three strata, the difference between the observed probability of death and the mean predicted probability of death were 0.0141, 0.0175, and  $-0.0344$ , respectively. Across the 100 strata, the median difference between the observed probability of death the mean predicted probability of death was  $-0.00025$  (25<sup>th</sup> and 75<sup>th</sup> percentiles:  $-0.00044$  and  $0.00004$ , respectively). We repeated the process of dividing the sample into derivation and validation components four additional times and determined the calibration of the model in the validation sample after estimating the regression coefficients in the derivation sample. The four resultant calibration plots were indistinguishable from that presented in the left panel of Figure 1. The center and right panels of Figure 1 depict the concordance between predicted and observed mortality for the Charlson and Elixhauser methods, respectively. Both of these methods had calibration that was comparable to that of the ADG model.

The final logistic regression model demonstrated excellent calibration-in-the-large, with an intercept of 0.0066 (95% CI:  $-0.0035$  to  $0.0166$ ). The difference in log-odds between predictions and observed outcomes was not statistically significantly different from zero (P

= 0.2014). The calibration slope was equal to 0.9961 (95% CI: 0.9903 to 1.0019). Thus, the final logistic regression model displayed no lack of calibration in the validation sample. The Charlson and Elixhauser models displayed acceptable calibration-in-the-large and calibration slopes that were not different from 1.

We examined the performance of the final logistic regression model in subjects under the age of 65 years. The c-statistic of the final logistic regression model was 0.824 and 0.819 in the derivation and validation samples, respectively. In subjects over the age of 65 years, the c-statistic of the final logistic regression model was 0.816 and 0.814 in the derivation and validation samples, respectively. We speculate that the decreased discrimination when the sample was stratified by age was due to fewer individual differences in demographic and diagnostic profiles in the more homogeneous subsamples.

After excluding subjects who died within 365 days of their index date, the c-statistic of the final logistic regression model when used for predicting the probability of mortality between 366 and 730 days of the index date was 0.905 in both the derivation and validation samples.

#### 4. Discussion

We examined the ability of logistic regression models using age, sex, and the Johns Hopkins Aggregated Diagnosis Groups (ADGs) to predict the probability of death within one year in a general population cohort. We used a large, population-based sample consisting of all Ontarians between the ages of 20 and 100 years who were alive on their birthday in 2007. We found that logistic regression models based on age, sex, and the ADGs accurately predicted mortality in this population sample. A logistic regression model consisting of age, sex, and 28 ADGs had excellent discrimination and calibration.

In a review of comorbidity scores to control for confounding in administrative database research, Schneeweiss and Maclure found that the c-statistics for four versions of the Charlson score and two versions of the Chronic Disease Score ranged, depending on the population and exposure, from 0.64 to 0.77 for in-hospital or 30-day mortality [24]. While our study population consisted of ambulatory patients, the performance of ADGs for predicting mortality performed very favorably compared to that of previous comorbidity scores for predicting mortality.

One advantage of using the Johns Hopkins ADGs is its application to non-hospitalized cohorts. Adaptations of the Charlson comorbidity index for use with ICD-9-CM or ICD-10 diagnostic codes are frequently used for comorbidity adjustment when estimating effects of exposures and treatments using administrative health care data, or for comparing outcomes across different health care providers. However, the original Charlson comorbidity score was derived in hospitalized general medical patients, and was initially validated in female oncology patients [3]. Furthermore, coding of the Deyo-Charlson index is designed for settings in which all subjects have been hospitalized, limiting its utility in ambulatory subjects. In contrast, ADGs can be determined for all subjects who have accessed the health care system, whether in an ambulatory setting or in a hospital setting. This permits comorbidity adjustment to be conducted in ambulatory populations, as well as in



hospitalized populations. For comparative purposes, we examined the predictive ability of logistic regression models that incorporated the Charlson comorbidity index or the Elixhauser comorbidities using data obtained from hospitalizations in the two years prior to the index date. We found that the use of these two models resulted in a minor decrease in discrimination compared to the model that incorporated the ADGs. However, calibration was approximately comparable across the three methods. Furthermore, it should be noted that use of a logistic regression model consisting of only age and sex resulted in only a modest decrease in discrimination compared to the other three regression models.

When choosing between risk-adjustment based on ADGs and risk-adjustment based on either the Charlson or Elixhauser comorbidities, one must consider several competing issues. Arguments in favor of an approach based on the ADGs include the minor increase in discrimination compared to the latter two approaches. Furthermore, the use of ADGs may have greater face validity since the ACG/ADG system was not designed primarily for use in hospitalized patients. Arguments in favor of the of the latter two approaches include the fact that the use of the ADGs requires a user license which typically requires a fee, whereas the Charlson and Elixhauser coding algorithms are non-proprietary and can be used without payment. It should be noted that the fee for using the ACG software may be nominal when used for research or academic purposes. A further relative disadvantage to the use of ADGs is that the assignment of ICD-9/10 diagnosis codes to ADGs is via a proprietary algorithm. Thus, the ADGs may be less transparent than the Charlson and Elixhauser comorbidity adjustment methods, for which the assignment of ICD-9/10 codes to different categories is explicitly described. As a consequence, it may difficult for researchers using the ADGs to fully explore their data so as to understand their results. Despite this lack of transparency, ACGs have been successfully used to predict mortality in several patient populations [13–16].

We have shown that the ADGs can be used to accurately predict 1-year mortality in a general population cohort. However, their utility for predicting mortality in specific disease populations or for predicting other outcomes needs to be examined in future studies. When regression models using age, sex, and the ADGs are used for predicting mortality in other general population cohorts, we recommend that researchers recalibrate the model by estimating the regression coefficients in their specific populations, rather than using regression coefficients estimated in our sample.

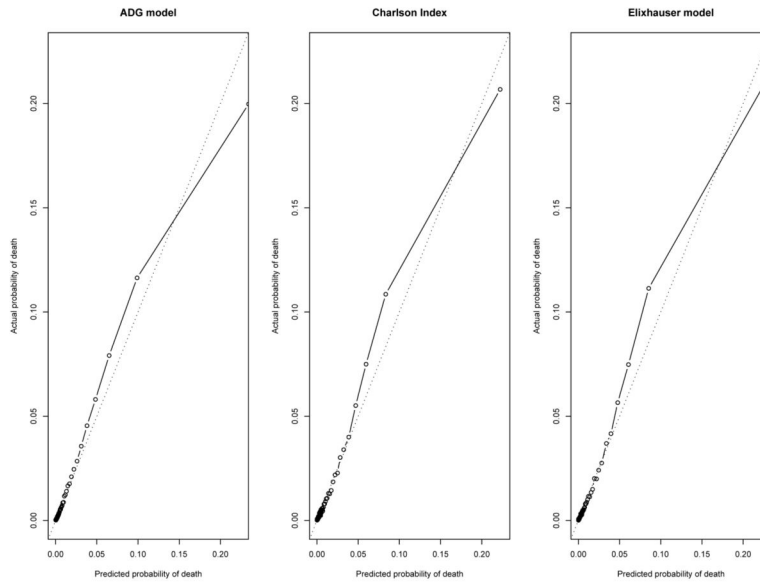
In conclusion, a logistic regression model that used age, sex, and the Johns Hopkins Aggregated Diagnosis Groups accurately predicted mortality in a general population cohort. This method may be useful for risk-adjustment or comorbidity adjustment in health services research when comparing mortality between health care providers or when using observational studies to estimate the effects of exposures, treatments, and interventions on mortality.

## References

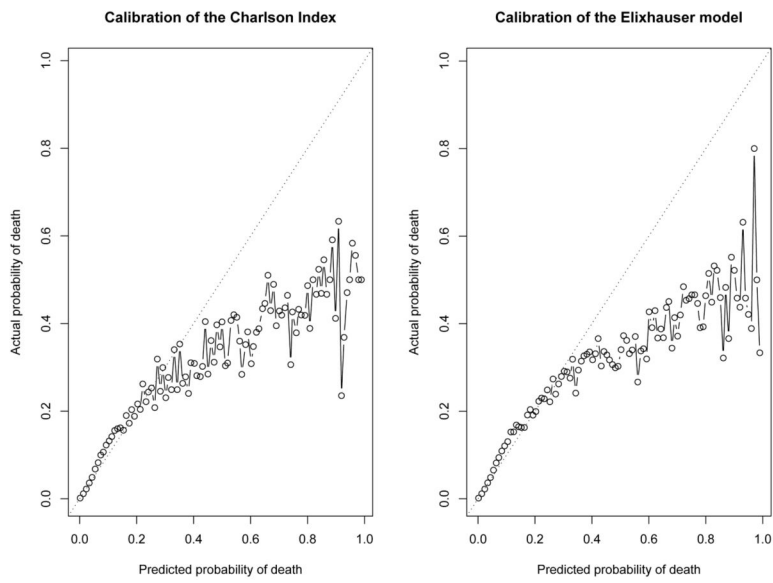
1. Iezzoni, LI., editor. Risk Adjustment for Measuring Health Outcomes. 2. Chicago, IL: Health Administration Press; 1997.

2. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and Validation. *Journal of Chronic Disease*. 1987; 40:373–383.
3. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology*. 1992; 45:613–619. [PubMed: 1607900]
4. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Medical Care*. 1998; 36:8–27. [PubMed: 9431328]
5. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical Care*. 2005; 43:1130–9. [PubMed: 16224307]
6. Sundararajan V, Henderson T, Perry C, Muggivan A, Quan H, Ghali WA. New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality. *Journal of Clinical Epidemiology*. 2004; 57:1288–1294. [PubMed: 15617955]
7. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *American Journal of Epidemiology*. 2001; 154:854–864. [PubMed: 11682368]
8. Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *Journal of Clinical Epidemiology*. 1992; 45:197–203. [PubMed: 1573438]
9. Weiner, Jonathan P., editor. The Johns Hopkins University Bloomberg School of Public Health, Health Services Research & Development Center. . The Johns Hopkins ACG® Case-Mix System Version 6.0 Release Notes. The Johns Hopkins University; Apr. 2003
10. Johns Hopkins University. [Site accessed July 29, 2010] Johns Hopkins ACG Case-Mix Adjustment System. Available at: <http://www.acg.jhsph.edu>
11. Weiner J, Starfield B, Steinwachs D, Abramson J. Development and application of a population-oriented measure of ambulatory care case-mix. *Medical Care*. 1991; 29:452–472. [PubMed: 1902278]
12. Starfield B, Weiner J, Murla P. Ambulatory care groups: A categorization of diagnoses for research and management. *Health Services Research*. 1991; 26:53–74. [PubMed: 1901841]
13. Pietz K, Petersen LA. Comparing self-reported health status and diagnosis-based risk adjustment to predict 1- and 2–5 year mortality. *Health Services Research*. 2007; 42:629–643. [PubMed: 17362210]
14. Berlowitz DR, Hoenig G, Cowper DC, Duncan PW, Vogel WB. Impact of comorbidities on stroke rehabilitation outcomes: Does the method matter? *Arch Phys Med Rehabil*. 2008; 89:1903–1906. [PubMed: 18929019]
15. Petersen LA, Pietz K, Woodard LD, Byrne M. Comparison of the predictive validity of diagnosis-based risk adjusters for clinical outcomes. *Medical Care*. 2005; 43:61–67. [PubMed: 15626935]
16. Fan VS, Maciejewski ML, Liu CF, McDonnell MB, Fihn SD. Comparison of risk adjustment measures based on self-report, administrative data, and pharmacy records to predict clinical outcomes. *Health Services and Outcomes Research Methodology*. 2006; 6:21–36.
17. Reid RJ, Roos NP, MacWilliam L, Frolich N, Black C. Assessing population health care need using a claims-based ACG morbidity measure: A validation analysis in the province of Manitoba. *Health Services Research*. 2002; 37:1345–1364. [PubMed: 12479500]
18. [Site accessed December 7, 2010] [http://www.health.gov.on.ca/english/providers/pub/ohip/physmanual/physmanual\\_mn.html](http://www.health.gov.on.ca/english/providers/pub/ohip/physmanual/physmanual_mn.html) (See Section 4.15)
19. Harrell, FE, Jr. *Regression Modeling Strategies: With applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer-Verlag; 2001.
20. Steyerberg, EW. *Clinical Prediction Models*. New York, NY: Springer-Verlag; 2009.
21. Lubitz JD, Riley GF. Trends in Medicare payments in the last year of life. *New England Journal of Medicine*. 1993; 328:1092–1096. [PubMed: 8455667]
22. Hogan CH, Lunney J, Gabel J, Lynn J. Medicare beneficiaries' costs of care in the last year of life. *Health Affairs*. Jul-Aug;2001 20(4):188–195.
23. Schneeweiss S, Wang PS, Avorn J, Glynn RJ. Improved comorbidity adjustment for predicting mortality in Medicare populations. *Health Services Research*. 2003; 38:1103–1120. [PubMed: 12968819]

24. Schneeweiss S, Maclure M. Use of comorbidity scores for control of confounding in studies using administrative databases. *International Journal of Epidemiology*. 2000; 29:891–898. [PubMed: 11034974]



**Figure 1.** Calibration: Observed vs. predicted probability of death across the centiles of risk



**Figure 2.**  
Calibration of Charlson and Elixhauser models

**Table 1**

Prevalence of each ADG and 1-year mortality in each ADG

ADG	Total number of subjects in ADG	Percent of cohort in ADG (N=10,498,413)	1-year mortality among those with ADG	1-year mortality among those without ADG	P-value
Time Limited: Minor	2,319,066	22.1%	1.16%	0.71%	<0.001
Time Limited: Minor-Primary Infections	4,384,500	41.8%	1.08%	0.62%	<0.001
Time Limited: Major	477,677	4.5%	4.26%	0.65%	<0.001
Time Limited: Major-Primary Infections	782,064	7.4%	2.35%	0.69%	<0.001
Allergies	678,426	6.5%	0.49%	0.83%	<0.001
Asthma	542,076	5.2%	1.21%	0.79%	<0.001
Likely to Recur: Discrete	3,013,077	28.7%	1.29%	0.62%	<0.001
Likely to Recur: Discrete-Infections	1,722,137	16.4%	1.40%	0.69%	<0.001
Likely to Recur: Progressive	252,777	2.4%	7.11%	0.65%	<0.001
Chronic Medical: Stable	4,235,449	40.3%	1.49%	0.35%	<0.001
Chronic Medical: Unstable	1,792,443	17.1%	3.02%	0.35%	<0.001
Chronic Speciality: Stable-Orthopedic	197,517	1.9%	1.05%	0.81%	<0.001
Chronic Speciality: Stable-Ear, Nose, Throat	209,815	2.0%	1.85%	0.79%	<0.001
Chronic Speciality: Stable-Eye	542,570	5.2%	2.83%	0.70%	<0.001
Chronic Speciality: Unstable-Orthopedic	212,353	2.0%	1.37%	0.80%	<0.001
Chronic Speciality: Unstable-Ear, Nose, Throat	59,278	0.6%	1.01%	0.81%	<0.001
Chronic Speciality: Unstable-Eye	479,308	4.6%	2.13%	0.75%	<0.001
Dermatologic	1,310,379	12.5%	0.88%	0.80%	<0.001
Injuries/Adverse Effects: Minor	1,895,740	18.1%	1.11%	0.74%	<0.001
Injuries/Adverse Effects: Major	1,410,248	13.4%	1.85%	0.65%	<0.001
Psychosocial: Time Limited, Minor	372,459	3.5%	1.37%	0.79%	<0.001
Psychosocial: Recurrent or Persistent, Stable	2,456,634	23.4%	1.15%	0.70%	<0.001
Psychosocial: Recurrent or Persistent, Unstable	464,373	4.4%	5.49%	0.59%	<0.001
Signs/Symptoms: Minor	3,154,412	30.0%	1.51%	0.51%	<0.001
Signs/Symptoms: Uncertain	4,601,563	43.8%	1.31%	0.42%	<0.001
Signs/Symptoms: Major	2,740,140	26.1%	1.55%	0.55%	<0.001

ADG	Total number of subjects in ADG	Percent of cohort in ADG (N=10,498,413)	1-year mortality among those with ADG	1-year mortality among those without ADG	P-value
Discretionary	1,542,641	14.7%	1.17%	0.75%	<0.001
See and Reassure	56,204	0.5%	2.52%	0.80%	<0.001
Prevention/Administrative	3,780,119	36.0%	0.89%	0.76%	<0.001
Malignancy	604,408	5.8%	4.56%	0.58%	<0.001
Pregnancy	389,306	3.7%	0.08%	0.84%	<0.001
Dental	160,194	1.5%	1.04%	0.81%	<0.001

**Table 2**

C-statistic (area under the ROC curve) for different models for predicting 1-year mortality in a population cohort.

Regression model	Sample	Predictive accuracy – derivation sample	Predictive accuracy – validation sample
Full logistic regression model (age, sex, and 32 ADGs)	Full population sample (ages 20 years to 100 years)	0.917	0.917
Age and sex	Full population sample (ages 20 years to 100 years)	0.883	0.883
32 ADGs	Full population sample (ages 20 years to 100 years)	0.866	0.864
Final logistic regression model	Full population sample (ages 20 years to 100 years)	0.917	0.917
Charlson comorbidity index (in addition to age and sex)	Full population sample (ages 20 years to 100 years)	0.906	0.906
Elixhauser comorbidities (in addition to age and sex)	Full population sample (ages 20 years to 100 years)	0.909	0.909
Final logistic regression model	Age < 65 years	0.824	0.819
Final logistic regression model	Age ≥ 65 years	0.816	0.814
Final logistic regression model	Subjects alive 365 days after index date (excluded subjects who died within 365 days of index date)	0.905	0.905



**Table 3**

Adjusted association between age, sex, ADGs and 1-year mortality in the final regression model in the derivation sample.

Predictor variable	Odds Ratio	95% Confidence Interval
Age (per year increase in age)	1.084	(1.083, 1.084)
Male subject	1.316	(1.289, 1.344)
Time Limited: Minor	0.946	(0.924, 0.968)
Time Limited: Minor-Primary Infections	1.114	(1.090, 1.139)
Time Limited: Major	1.652	(1.608, 1.697)
Time Limited: Major-Primary Infections	1.666	(1.622, 1.712)
Allergies	0.676	(0.641, 0.712)
Asthma	1.128	(1.085, 1.174)
Likely to Recur: Progressive	1.606	(1.561, 1.651)
Chronic Medical: Stable	0.827	(0.806, 0.849)
Chronic Medical: Unstable	1.849	(1.804, 1.894)
Chronic Specialty: Stable-Orthopedic	0.795	(0.745, 0.848)
Chronic Specialty: Stable-Ear, Nose, Throat	0.763	(0.727, 0.802)
Chronic Specialty: Stable-Eye	0.800	(0.778, 0.822)
Chronic Specialty: Unstable-Orthopedic	0.845	(0.799, 0.894)
Chronic Specialty: Unstable-Ear, Nose, Throat	0.656	(0.582, 0.739)
Chronic Specialty: Unstable-Eye	0.841	(0.814, 0.868)
Dermatologic	0.674	(0.653, 0.694)
Injuries/Adverse Effects: Major	1.159	(1.131, 1.187)
Psychosocial: Time Limited, Minor	1.166	(1.115, 1.219)
Psychosocial: Recurrent or Persistent, Stable	1.046	(1.022, 1.070)
Psychosocial: Recurrent or Persistent, Unstable	2.880	(2.809, 2.953)
Signs/Symptoms: Minor	1.248	(1.220, 1.277)
Signs/Symptoms: Uncertain	1.123	(1.095, 1.152)
Signs/Symptoms: Major	1.248	(1.219, 1.277)
Discretionary	0.835	(0.814, 0.857)
Prevention/Administrative	0.872	(0.853, 0.891)
Malignancy	2.397	(2.341, 2.454)
Pregnancy	0.732	(0.622, 0.862)
Dental	1.198	(1.113, 1.291)

Note: The effect of each variable listed in Table 3 is adjusted for all other variables in Table 3.