HHS Public Access

Author manuscript

*Nurs Res*. Author manuscript; available in PMC 2020 May 01.

Published in final edited form as:

*Nurs Res*. 2019 ; 68(3): 200–209. doi:10.1097/NNR.0000000000000350.

# Comparison of Measures to Predict Mortality and Length of Stay in Hospitalized Patients

**Jianfang Liu, PhD [Assistant Professor]**,
School of Nursing, Columbia University, New York, NY

**Elaine Larson, RN, PhD, FAAN, CIC**,
Associate Dean for Research and Anna C. Maxwell Professor of Nursing Research, School of Nursing, Professor of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY

**Amanda Hessels, PhD, MPH, RN, CIC, CPHQ, FAPIC [Assistant Professor]**,
School of nursing, Columbia University, New York, NY

Nurse Scientist, Hackensack Meridian Health, Neptune, NJ

**Bevin Cohen, PhD, MPH, RN [Associate Research Scientist]**,
School of Nursing, Columbia University, New York, NY

**Philip Zachariah, MD, MS [Assistant Professor]**,
Columbia University Medical Center & NewYork-Presbyterian Morgan Stanley Children's Hospital, New York, NY

**David Caplan, BS [Senior Technical Specialist]**, and
Division of Quality Analytics, NewYork-Presbyterian Hospital, New York, NY

**Jingjing Shang, PhD, RN [Associate Professor]**
School of Nursing, Columbia University, New York, NY

## Abstract

**Background:** Patient risk adjustment is critical for hospital benchmarking and allocation of healthcare resources. However, considerable heterogeneity exists among measures.

Corresponding author: Jianfang Liu, 560 West 168th Street, New York, NY, 10032; jl4029@cumc.columbia.edu.
**Jianfang Liu, PhD**, is Assistant Professor, School of Nursing, Columbia University, New York, NY.
**Elaine Larson, RN, PhD, FAAN, CIC**, is Associate Dean for Research and Anna C. Maxwell Professor of Nursing Research, School of Nursing, Columbia University, New York, NY.
**Amanda Hessels, PhD, MPH, RN, CIC, CPHQ, FAPIC**, is Assistant Professor, School of Nursing, Columbia University, New York, NY.
**Bevin Cohen, PhD, MPH, RN**, is Associate Research Scientist, School of Nursing, Columbia University, New York, NY.
**Jingjing Shang, PhD, RN**, is Associate Professor, School of Nursing, Columbia University, New York, NY.
**Elaine Larson**, is Professor of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY.
**Amanda Hessels**, is Nurse Scientist, Hackensack Meridian Health, Neptune, NJ.
**Philip Zachariah, MD, MS**, is Assistant Professor, Columbia University Medical Center & New York-Presbyterian Morgan Stanley Children's Hospital, New York, NY.
**David Caplan, BS**, is Senior Technical Specialist, Division of Quality Analytics, NewYork-Presbyterian Hospital, New York, NY.

The authors have no conflicts of interest to report.

**Objectives:** The performance of five measures was compared to predict mortality and length of stay (LOS) in hospitalized adults using claims data; these include three comorbidity composite scores (Charlson/Deyo age-comorbidity score, V W Elixhauser comorbidity score, and V W Elixhauser age-comorbidity score), 3M risk of mortality (3M ROM), and 3M severity of illness (3M SOI) subclasses.

**Method:** Binary logistic and zero-truncated negative binomial regression models were applied to a two-year retrospective dataset (2013–2014) with 123,641 adult inpatient admissions from a large hospital system in New York City.

**Results:** All five measures demonstrated good to strong model fit for predicting in-hospital mortality with C-statistics of 0.74 (95% CI [0.74–0.75]), 0.80 (95% CI [0.80–0.81]), 0.81(95% CI [0.81–0.82]), 0.94 (95% CI [0.93–0.94]), and 0.90 (95% CI [0.90–0.91]) for Charlson/Deyo age-comorbidity score, V W Elixhauser comorbidity score, V W Elixhauser age-comorbidity score, 3M ROM, and 3M SOI, respectively. The model fit statistics to predict hospital LOS measured by the Likelihood Ratio Index were 0.3%, 1.2%, 1.1%, 6.2%, and 4.3%, respectively.

**Discussion:** The measures tested in this study can guide nurse managers in the assignment of nursing care and coordination of needed patient services and administrators to effectively and efficiently support optimal nursing care.

## Keywords

3M subclasses; Charlson/Deyo; Elixhauser; length of stay; mortality; nursing care; patient risk adjustment

Accounting for patient-associated risk for adverse outcomes is critical for hospital administrators to allocate hospital resources and prioritize the care. Among multidimensional factors related to patient outcomes, comorbidity is of major importance since patients with more comorbidities have higher risks of death, complications, and use more hospital resources (Fortin, Soubhi, Hudon, Bayliss, & van den Akker, 2007; Ritchie, 2007). Understanding patient comorbidities is critical for nursing care planning. By separating modifiable patient and provider characteristics from those that are attributable to intrinsic, unmodifiable patient risk (Hessels, Liu, Cohen, Shang, & Larson, 2018; Olson et al., 2013), hospital administrators can appropriately allocate nursing staffing and other resources and provide the most cost-effective care. With accurate assessment of patients' risks, nurses and other clinicians at the point of care will be better informed to integrate patient health data with current evidence and treatment protocols to assure that the most effective care is rendered.

Comorbidity measurement is critical to hospital performance comparisons and benchmarking. It is important to adjust for potential confounding when describing and comparing patient outcomes between facilities (van Walraven, Austin, Jennings, Quan, & Forster, 2009). If differences among patients in their initial conditions and prognostic expectations are not appropriately identified and accounted for, subsequent patient outcomes may be erroneously attributed to variations in treatment.

For effective nursing care planning and hospital resource allocation, as well as to adjust for potential confounding for hospital comparisons or benchmarking, it is necessary to describe and compare patient populations whenever chronic disease conditions are associated with outcomes such as length of stay (LOS) or death. This becomes especially important for healthcare organizations participating in the Bundled Payment for Care Improvement Initiative. Under this initiative, hospitals have to absorb the financial burden in the event of an unplanned prolonged LOS. Choosing the right risk adjustment measure can help the organization identify patient factors related to increased costs and allocate resources to avoid incurring additional costs.

Among the several risk-adjustment methods available, the Charlson/Deyo and the Elixhauser methods are the most widely used (Ladha et al., 2015). The Charlson Comorbidity Index was developed in 1987 using 19 medical diagnosis from clinical data to predict one-year mortality for patients with multiple conditions (Charlson, Pompei, Ales, & MacKenzie, 1987). The coding algorithm has been modified several times since its original version. In 1992, a coding algorithm including 17 diagnoses was adapted by Deyo using ICD-9-CM codes from administrative data (Deyo, Cherkin, & Ciol, 1992). The Charlson/Deyo measure has been the most commonly used adaptation of the original Charlson method. More recently, Quan and colleagues (2005) modified and enhanced the ICD-9-CM coding algorithm for each comorbidity and the performance of the enhanced coding algorithm in predicting in-hospital mortality was tested and verified.

The Elixhauser Comorbidity Index is another commonly used method to predict hospital resource use and in-hospital mortality. First developed using ICD-9-CM codes from hospital administrative data, the original Elixhauser Index included 30 diagnoses (Elixhauser, Steiner, Harris, & Coffey, 1998) and was later modified by Garland (Garland et al., 2012) to include 31 diagnoses. In 2009, a weighted composite Elixhauser score, the V W Elixhauser comorbidity score, was generated and similar discriminating ability of the composite score and 30 individual comorbidities in predicting in-hospital mortality was reported (van Walraven et al., 2009).

In addition to the two risk-adjustment methods, the All Patient Refined Diagnosis Related Group (APR-DRG) is a widely adopted commercial patient classification system that uses administrative data (Romano & Chan, 2000). The APR-DRG methodology that provides patient classification systems was developed by the New York State Department of Health in collaboration with 3M Health Information System for the purposes of analyzing and comparing of outcomes for a given diagnostic group (Iezzoni et al., 1995). The APR-DRG expanded the basic DRG structure and added two sets of subclasses to each base APR-DRG: severity of illness (SOI), defined as the extent of physiologic decompensation or organ system loss of function; and risk of mortality (ROM), estimating the likelihood of death. In APR-DRG calculations, patients are first assigned to a base APR-DRG based on primary diagnosis, and then separately assigned to one of the four possible levels (minor, moderate, major, extreme) of the distinct SOI and ROM subclasses; assignment is based on secondary diagnoses and the interaction between diagnoses, age, and selected procedures (Averill et al., 1997; Solutions, 2018). The 3M ROM subclass was tested in intensive care units and found to have good performance (Baram et al., 2008). The APR-DRG has been adopted by several

states to compare hospital performance (Romano & Chan, 2000) and was used in various studies to predict patients' hospital LOS and in-hospital mortality (Hansen et al., 2017; Nante, De Marco, Balzi, Addari, & Buiatti, 2000).

Researchers have reported that using 30 or 31 individual comorbidities from the Elixhauser method improves discriminating ability to predict mortality (Lieffers, Baracos, Winget, & Fassbender, 2011) or in-hospital mortality (Gutacker, Bloor, & Cookson, 2015; Southern, Quan, & Ghali, 2004; Stukenborg, Wagner, & Connors, 2001) over the Charlson/Deyo method, which uses only 17 individual comorbidities. Researchers have also compared APR-DRG with the Elixhauser method using individual comorbidities to predict mortality among patients hospitalized for hip replacement (Messina et al., 2017) and found that the 3M ROM slightly out-performed the Elixhauser method for both in-hospital and 30-day mortalities.

However, using 30 or 31 individual diagnosis to describe or model a patient population could be cumbersome. Austin, Wong, Uzzo, Beck, and Egleston (2015, p.1) suggested that once a valid comorbidity composite score is calculated, "no other information about the comorbidity variables used to create the score is generally needed." In a study that included individuals diagnosed with localized kidney cancer (age ≥ 65), similar performance of individual comorbidities and composite scores in predicting mortality was reported using either the Charlson/Deyo method or the Elixhauser method (Austin et al., 2015). When describing a population and/or controlling for potential confounding, there are at least three appealing advantages of using composite scores rather than the individual comorbidities:

1. being able to provide an overall description of the population;
2. reducing the risk of overfitting in small datasets; and
3. alleviating computational requirements in large datasets (Thompson et al., 2015; van Walraven et al., 2009).

Hence, it is important to compare various composite scores of risk adjustment measures to identify the most relevant, feasible, and valid one for predicting outcomes in adult hospitalized patients.

Since health claims data and other electronic sources are increasingly used for evaluating patient outcomes (Blumenthal, 2009; Cohen et al., 2015), it is important to recognize the strengths and limitations of these measures for predicting patient outcomes. The aim of this study was to compare the performance of measures commonly used for patient risk adjustment for predicting in-hospital mortality and hospital LOS.

## Methods

### Theoretical Framework

This study was guided by the Quality Health Outcomes Model (Mitchell, Ferketich, & Jennings, 1998) that delineates relations among four components—system, client, interventions, and outcomes. In the model, *system* refers to organizational structure; *intervention* refers to processes and activities that clinicians perform; *outcomes* represent

health indicators such as mortality, LOS, etc., and client refers to patient's individual characteristics such as demographics, health status, and comorbidity. In the Quality Health Outcome Model, the effect of medical interventions on outcomes is indirect and mediated by system and client components.

## Data Source

This was a secondary data analysis that included data from all adult (≥ 18 years) inpatient hospital discharges between January 1, 2013 and December 31, 2014 from a large hospital system in the New York City metropolitan area. The hospital system includes a community hospital and two tertiary/quaternary care hospitals with > 2,000 beds and over 100,000 patient admissions annually. All study data, including patients' demographic information, admission date, discharge date and status, and International Classification of Disease, 9th revision, Clinical Modification (ICD-9-CM) codes were extracted retrospectively from the institution's clinical data warehouse server to an institution encrypted network drive using standardized query language by the data manager.

## Risk Adjustment Measures Tested

Table 1 provides a summary of the five measures used in this study.

**Charlson/Deyo method: Charlson/Deyo age-comorbidity score.—**The validated Charlson/Deyo age-comorbidity score represents a patient's clinical comorbidity score. It is measured as a weighted summation of 17 categories of comorbid illnesses and the patient's age point: beginning at age 40, one point is added for each 10-year increment (Charlson, Szatrowski, Peterson, & Gold, 1994). The modified and enhanced version of the ICD-9-CM coding algorithm proposed by Quan et al. (2005) was used. Appendix A lists the ICD-9-CM codes and assigned points for each comorbidity and age range. The Charlson/Deyo age-comorbidity score ranges from 0 to 38. Present on admission (POA) indicator is collected in New York State discharge data (Coffey, Milenkovic, & Andrews, 2006) to distinguish POA comorbidities from complications that occurred during the admission (Garrett, 2009). Only comorbidities identified by the POA indicator were used to calculate the Charlson/Deyo score as well as the two composite scores from the Elixhauser method as described below.

**V W Elixhauser method: V W Elixhauser comorbidity score and V W Elixhauser age-comorbidity score.—**Using POA as the indicator to identify comorbidities, we calculated two Exlihauser comorbidity composite scores. More recently, Thompson et al. (2015) proposed a similar weighted composite score that slightly outperformed the V W Elixhaurser score. However, the Thompson score uses diagnosis-related group codes to identify comorbidities rather than using POA indicator, and it is not clear if the Thompson score is superior in other datasets. We first used the V W Exlihauser comorbidity score, which is a weighted composite score of 30 individual comorbidities based on the association between each comorbidity and hospital death proposed by van Walraven et al. (2009). The points assigned to each comorbidity ranged from −7 (for drug abuse) to 12 (for metastatic cancer) (Appendix A). The V W Exlihauser comorbidity score ranges from −19 to 89. It does not include age but is highly correlated with a patient's age. Next, we created a V W Exlihauser age-comorbidity score including patient age points based

on the same age-point algorithm as the Charlson/Deyo method. As with the Charlson/Deyo method, the modified and enhanced ICD-9 coding algorithm was used (Quan et al., 2005).

**APR-DRG method: 3M risk of mortality and 3M severity of illness.—**The ROM and SOI subclasses were extracted directly from the hospital database as generated by the 3M™ Core Grouping Software, which is licensed and maintained by the hospital system. Annual updates to the Grouping Software programs are required to accommodate annual revisions to the ICD diagnosis and procedure codes. In the United States, these revisions become effective October 1 of each year, and thus the Grouping Software versions applicable to our study period are V30, V31, and V32 (3M™ Core Grouping Software). APR-DRGs are proprietary products of the 3M Health Information System. Notably, ROM and SOI subclasses are determined by both in-hospital complications and POA comorbidities; therefore, all diagnoses—POA and non-POA—were used (Romano & Chan, 2000).

## Outcomes and Other Variables

Patient characteristics including age, gender, admitting hospital, LOS, ICD-9-CM codes, and discharge status were extracted from hospital discharge data. The outcome variables in this study were in-hospital mortality and LOS.

## Analysis

Descriptive statistics were used to summarize the admission characteristics in our sample. To assess the association between each of the five risk-adjustment measures and in-hospital mortality, we built separate logistic regression models for each measure. As hospital LOS is a typical example of zero-truncated count data with evidence of overdispersion (Zuur, Ieno, Walker, Saveliev, & Smith, 2009), a zero-truncated negative binomial regression model was built to assess the association between each of the five measures and hospital LOS.

To assess the performance of each logistic regression model, we calculated the C-statistic and associated 95% confidence intervals (CIs) to determine the absolute fit. The C-statistic ranges from 0.5 to 1, where a 0.5 indicates that the model is no better than a random chance in predicting the outcome and a 1 indicates that the model has perfect discriminating ability. Models are typically considered good when the C-statistic is greater than 0.7 and strong when it is greater than 0.8 (Hosmer & Lemeshow, 2000).

It is not common to report R-square measures of model fit for count data (Cameron & Windmeijer, 1996). However, in order to provide a goodness of fit measure that can be compared with other or future studies, we assessed the performance of each zero-truncated negative binomial regression model by calculating the likelihood ratio index for each model. Likelihood ratio index measures the proportionate reduction in the log-likelihood due to inclusion of regressors and is calculated as one minus the ratio of the full-model log-likelihood to the intercept-only log-likelihood (Cameron & Windmeijer, 1996). Likelihood ratio index is sometimes proposed as a general pseudo R-square measure (Cameron & Windmeijer, 1997) with advantages of being reported by statistical packages (Cameron &

Windmeijer, 1996), and the ease of calculation based on results provided by general statistical packages for exponential family regression models.

Odds ratios (*OR*s) from logistic regression and incidence rate ratios (*IRR*s), calculated as exponentiated regression coefficients from zero-truncated negative binomial regression, are also reported for the size and direction of effects. All statistical analyses were performed using SAS version 9.3 (SAS Institue, 2012) or R (R core team, 2013).

## Results

Characteristics of the sample and summaries of the risk adjustment measures and outcomes are shown in Table 2. The sample included 123,641 adult admissions, of whom 63,689 (51.5%) were female. The average age was 62.2 (*SD* = 18.5) years. The in-hospital mortality rate was 2.8%. The median hospital LOS was 5 with interquartile range of 3–9 days.

Table 3 shows the model performance for the binary logistic regression models assessing the relationship between each of the risk adjustment measures and in-hospital mortality. All five measures demonstrated good to strong model fit for predicting in-hospital mortality by C-statistics of Charlson/Deyo age-comorbidity score (0.74, 95% CI [0.74–0.75]), V W Elixhauser comorbidity score (0.80, 95% CI [0.80–0.81]), V W Elixhauser age-comorbidity score (0.81, 95% CI [0.81–0.82]), 3M ROM (0.94, 95% CI [0.93–0.94]), and 3M SOI (0.90, 95% CI [0.90–0.91]). The two V W Elixhauser composite scores out-performed the Charlson/Deyo age-comorbidity score with higher C-statistics. The V W Elixhauser age-comorbidity score had similar model performance to the V W Elixhauser comorbidity score with overlapping 95% CIs of C-statistics. Both 3M ROM and 3M SOI subclasses provided statistically superior predictions on in-hospital mortality with higher C-statistics, and 3M ROM had the best performance.

Higher Charlson age-comorbidity scores were associated with higher risk of in-hospital mortality: With a one-unit increase in the Charlson age-comorbidity score, the odds of in-hospital mortality increased by 28% (*OR* = 1.28, 95% CI [1.27–1.29]). Higher V W Elixhauser comorbidity scores (OR = 1.12, 95% CI [1.12–1.12]) and higher V W Elixhauser age-comorbidity scores (OR = 1.12, 95% CI [1.12–1.12]) were also associated with higher risk of in-hospital mortality. Greater odds of mortality were associated with higher risk category measured by 3M ROM. Results of in-hospital mortality rate by risk category of 3M subclasses are presented in Table 4. Inpatients who scored minor risk of mortality had the minimal in-hospital mortality rate (0.02%), followed by those who scored moderate risk of mortality (0.14%), and those who scored major risk of mortality (2.45%). Inpatients who scored extreme risk of morality had the highest in-hospital mortality rate (28.5%). The same pattern of in-hospital mortality rate was also found in the 3M SOI.

Table 5 shows the model performance for the zero-truncated negative binomial regression models assessing the relationship between each of the risk adjustment measures and hospital LOS. The Charlson/Deyo age-comorbidity score, V W Elixhauser comorbidity score, and V W Elixhauser age-comorbidity score were weak predictors of hospital LOS (Likelihood ratio index = 0.3%, 1.2%, and 1.1%, respectively). The 3M SOI and 3M ROM were better in

predicting hospital LOS (Likelihood ratio index = 6.2% and 4.3%, respectively). Among the five measures, 3M SOI had the best predictive performance.

Higher Charlson age-comorbidity scores (*IRR* = 1.042, 95% CI [1.041–1.045]) were associated with longer hospital stays: With a one-unit increase in the Charlson age-comorbidity score, the hospital LOS increased by 4.2%. Higher V W Elixhauser comorbidity scores (*IRR*, 1.029, 95% CI [1.028–1.030]) and higher V W Elixhauser age-comorbidity scores (*IRR* = 1.025, 95% CI, [1.025–1.026]) were also associated with longer hospital stays. Longer hospital stays were associated with a greater risk category of 3M SOI. As compared to inpatients who scored minor severity of illness, inpatients with extreme severity of illness had the longest hospital stay (*IRR* = 5.30, 95% CI [5.22–5.39]), followed by those with a major severity of illness (*IRR* = 2.30, 95% CI [2.27–2.33]) and those with moderate severity of illness (*IRR* = 1.44, 95% CI [1.42–1.46]). The same pattern was also found in 3M ROM.

## Discussion

We compared the performance of five measures developed from three methods to predict inpatients' in-hospital mortality and hospital LOS using large scale electronic healthcare data. Previous researchers have reported the Elixhauser method outperformed the Charslon/Deyo method to predict mortality (Lieffers et al., 2011) or in-hospital mortality by using individual comorbidities (Gutacker et al., 2015; Southern et al., 2004; Stukenborg et al., 2001). However, to our knowledge, no previous research has compared the performance of the weighted composite scores from these two methods in adult inpatients with different medical conditions and various age ranges in in-hospital mortality or LOS. We found using composite scores that the Elixhauser method still out-performed the Charslon/Deyo method.

The performance of the Elixhauser method may be because the Elixhauser's ICD-9-CM coding algorithm captured more comorbidities on more patients as compared to the Charlson/Deyo method (Southern et al., 2004). However, it is important to note that the differences in predictive power between the Charlson/Deyo and the two Elixhauser composite scores were small; the C-statistic for predicting in-hospital mortality were 0.74, 0.80, and 0.81 for the Charlson/Deyo, V W Elixhauser comorbidity score, and V W Elixhauser age-comorbidity score, respectively. Other researchers also reported C-statistics of 0.7 – 0.8 for predicting in-hospital mortality using these methods, with differences of < 0.1 between the Charlson/Deyo and Elixhauser methods (Southern et al., 2004; Stukenborg et al., 2001). More recently, researchers in Taiwan reported C-statistics of 0.76 in predicting post-illness mortality in patients with thoracic empyema using Charlson Method (Wu, Liu, Lee, Kuo, & Hsieh, 2018). In addition, in alignment with the Charlson/Deyo age-comorbidity score, we created the new V W Elixhauser age-comorbidity score by adding age points to the Elixhauser composite score; the new score had similar accuracy to the V W Elixhauser comorbidity score in predicting in-hospital mortality and LOS.

Our study further extends the current literature by testing both the 3M ROM and SOI subclasses in adult inpatients with different medical conditions and by comparing those two subclasses with the Charlson/Deyo and Elixhauser methods using composite scores rather

than individual comorbidities. The high predictive power of both the 3M ROM and SOI subclasses was verified in our study with C-statistics of 0.94 and 0.90 for predicting in-hospital mortality. Previous researchers found that the 3M ROM slightly outperformed the Elixhauser method using individual comorbidities for both in-hospital and 30-day mortalities (Messina et al., 2017). Ours is the first study to compare the Charlson/Deyo and Elixhauser weighted composite scores with the 3M ROM and 3M SOI subclasses for predicting in-hospital mortality and LOS. We found that both 3M subclasses outperformed Charlson/Deyo and Elixhauser methods in predicting these two outcomes.

We also found that the 3M ROM had the best predictive power for in-hospital mortality while the 3M SOI was best in predicting hospital LOS. This is not surprising because the 3M SOI subclass was structured for evaluating resource use, while the 3M ROM was developed primarily for evaluating patient mortality (Averill et al., 1997). However, the high predictive performance of the two APR-DRG subclasses may be because they include all diagnoses, both POA and non-POA (Romano & Chan, 2000). When the 3M ROM subclass was tested by including only POA diagnoses, the predictive power was significantly less; the C-statistic dropped from 0.83 – 0.85 to 0.74 (Romano & Chan, 2000). Further, complications developed after admission are also directly affected by the care patients receive during the hospitalization. Previous researchers using 2005 California discharge data and vital statistics death files found that substantial effects on hospital rankings of performance were reported when incorporating POA indicators into risk-adjusted models (Goldman, Chu, Bacchetti, Kruger, & Bindman, 2015). Thus, APR-DRG methods may not be suitable for comparisons of hospital performance or for clinical research studies in which it is important to control for severity of illness at admission.

Our findings have significant clinical practice implications. Hospitals are now mandated to submit standardized outcome measures. Risk adjustment is critical to conduct and fairly interpret benchmarking of organizations and provider performance. Our findings suggest that the Elixhauser method is slightly better than the Charlson/Deyo method in predicting in-hospital mortality and LOS, and caution should be applied when using APR-DRG to ensure the POA indicator is factored into the calculation. However, since the differences between the risk adjustment methods are small, hospital administrators may choose one based on the best available sources. The risk adjustment is also important in plan of care. Risk profiles generated by these approaches can help clinicians prioritize tasks and develop personalized care based on individual patient risk score and needs. Nurse managers and administrators can also use this information to assign nursing care, distribute direct-care resources, and provide needed support services. However, it is important to note that these comorbidity measures have been developed primarily to predict death and hospital LOS, taking only medical diagnoses into account. None of these measures consider functional capacity, which may have significant effects on patient outcomes and recovery. Future research should include these important factors and should also assess their predictive value for other outcomes such as rehospitalization or cost.

It is important to note that these study findings are not generalizable to pediatric populations, as the diagnoses included in these indices such as chronic pulmonary disease and congestive heart failure are not commonly found in children. It is also important to note that these

measures were developed for use in acute-care settings. Caution needs to be applied when they are used in other healthcare settings such as long-term care, nursing homes, or home healthcare.

## Conclusion

In conclusion, all of the measures examined in this study had good discrimination for in-hospital mortality. The composite scores from the Elixhauser method slightly outperformed the Charlson/Deyo method. Between the two composite scores from the Elixhauser method, the age-comorbidity score might be a better choice when the purpose is to identify the most parsimonious model—particularly with small datasets. The advantage of the V W Elixhauser comorbidity score is that the effects of the age and comorbidity composite score can be examined separately.

Although the APR-DRG method performed better in predicting both in-hospital mortality and hospital LOS, the 3M ROM and SOI are usually calculated after hospital discharge and therefore are not appropriate for use at admission for planning care. Further, the inclusion of complications developed after hospital admission in the calculation of APR-DRG makes it less appealing than the Charlson/Deyo and Elixhauser methods, which also have the advantage of being readily available and free of cost. Health services researchers and hospital administrators must choose the best measure based on the available datasets, sources, and needs.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgement:

This work was supported by a grant from the Agency for Healthcare Research and Quality (AHRQ) R01 HS02491501.

**Ethical Conduct of Research:** This study was approved by Columbia University Medical Center's institutional review board (AAK4050).

## References

3M™ Core Grouping Software. Retrieved December 19, 2017, from https://www.3m.com/3M/en_US/company-us/all-3m-products/~/3M-Core-Grouping-Software/?N=5002385+3290603312&rt=rud

Austin SR, Wong YN, Uzzo RG, Beck JR, & Egleston BL (2015). Why summary comorbidity measures such as the Charlson comorbidity index and Elixhauser score work. Medical Care, 53, e65–e72. doi:10.1097/MLR.0b013e318297429c [PubMed: 23703645]

Averill R, Goldfield N, Steinbeck B, Grant T, Muldoon J, Brough J, & Gay J (1997). Development of the All Patient Refined DRGs (APR-DRGs). St. Paul, MN: 3M Retrieved March 12, 2017, from https://pdfs.semanticscholar.org/d26f/de2af6c9f56d4460367b4bb6694adbbe9a57.pdf

Baram D, Daroowalla F, Garcia R, Zhang G, Chen JJ, Healy E, … Richman P (2008). Use of the All Patient Refined-Diagnosis Related Group (APR-DRG) risk of mortality score as a severity adjustor in the Medical ICU. Clinical Medicine: Circulatory, Respiratory and Pulmonary Medicine, 2, CCRPM-S544. doi:10.4137/ccrpm.s544

Blumenthal D (2009). Stimulating the adoption of health information technology. New England Journal of Medicine, 360, 1477–1479. doi:10.1056/NEJMp0901592 [PubMed: 19321856]

Cameron AC, & Windmeijer FA (1996). R-squared measures for count data regression models with applications to health-care utilization. Journal of Business & Economic Statistics, 14, 209–220. doi:10.2307/1392433

Cameron AC, & Windmeijer FA (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. Journal of Econometrics, 77, 329–342. doi:10.1016/s0304-4076(96)01818-0

Charlson M, Szatrowski TP, Peterson J, & Gold J (1994). Validation of a combined comorbidity index. Journal of Clinical Epidemiology, 47, 1245–1251. doi:10.1016/0895-4356(94)90129-5 [PubMed: 7722560]

Charlson ME, Pompei P, Ales KL, & MacKenzie CR (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Clinical Epidemiology, 40, 373–383. doi:10.1016/0021-9681(87)90171-8

Coffey RM, Milenkovic M, & Andrews RM (2006). The Case for the Present-on-Admission (POA) Indicator. 2006 HCUP Methods Series Report # 2006-01. Online. June 26, 2006. U.S. Agency for Healthcare Research and Quality Retrieved March 12, 2017, from https://www.hcup-us.ahrq.gov/reports/methods/2006_1.pdf

Cohen B, Vawdrey DK, Liu J, Caplan D, Furuya EY, Mis FW, & Larson E (2015). Challenges associated with using large data sets for quality assessment and research in clinical settings. Policy, Politics, & Nursing Practice, 16, 117–124. doi:10.1177/1527154415603358

Deyo RA, Cherkin DC, & Ciol MA (1992). Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. Journal of Clinical Epidemiology, 45, 613–619. doi:10.1016/0895-4356(92)90133-8 [PubMed: 1607900]

Elixhauser A, Steiner C, Harris DR, & Coffey RM (1998). Comorbidity measures for use with administrative data. Medical Care, 36, 8–27. doi:10.1097/00005650-199801000-00004 [PubMed: 9431328]

Fortin M, Soubhi H, Hudon C, Bayliss EA, & Van den Akker M (2007). Multimorbidity's many challenges. BMJ, 334, 1016. doi:10.1136/bmj.39201.463819.2c [PubMed: 17510108]

Garland A, Olafson K, Ramsey C, Yogendren M, Chateau D, & McGowan K (2012). The epidemiology and outcomes of critical illness in Manitoba. Winnipeg, MB: Manitoba Centre for Health Policy.

Garrett GS (2009). Present on admission: Where we are now. Journal of the American Health Information Management Association, 80, 22–26.

Goldman LE, Chu PW, Bacchetti P, Kruger J, Bindman A (2015). Effect of present-on-admission (POA) reporting accuracy on hospital performance assessments using risk-adjusted mortality. Health Services Research, 50, 922–938. doi:10.1111/1475-6773.12239 [PubMed: 25285372]

Gutacker N, Bloor K, & Cookson R (2015). Comparing the performance of the Charlon/Deyo and Elixhauser comorbidity measures across five European countries and three conditions. European Journal of Public Health, 25, 15–20. doi:10.1093/eurpub/cku221

Hansen RN, Pham AT, Böing EA, Lovelace B, Wan GJ, & Miller TE (2017). Comparative analysis of length of stay, hospitalization costs, opioid use, and discharge status among spine surgery patients with postoperative pain management including intravenous versus oral acetaminophen. Current Medical Research and Opinion, 33, 943–948. doi:10.1080/03007995.2017.1297702 [PubMed: 28276273]

Hessels AJ, Liu J, Cohen B, Shang J, & Larson EL (2018). Severity of illness measures for pediatric inpatients. Journal of Healthcare Quality, 40, e77–e89. doi:10.1097/jhq.0000000000000135

Hosmer DW & Lemeshow S (2000). Applied logistic regression (2nd ed.). New York, NY: John Wiley & Sons.

Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, & Mackiernan YD (1995). Predicting who dies depends on how severity is measured: Implications for evaluating patient outcomes. Annals of Internal Medicine, 123, 763–770. doi:10.7326/0003-4819-123-10-199511150-00004 [PubMed: 7574194]

Ladha KS, Zhao K, Quraishi SA, Kurth T, Eikermann M, Kaafarani HMA, … Lee J (2015). The Deyo-Charlson and Elixhauser-van Walraven Comorbidity Indices as predictors of mortality in critically ill patient. BMJ Open, 5, e008990. doi:10.1136/bmjopen-2015-008990

Lieffers JR, Baracos VE, Winget M, & Fassbender K (2011). A comparison of Charlson and Elixhauser comorbidity measures to predict colorectal cancer survival using administrative health data. Cancer, 117, 1957–1965. doi:10.1002/cncr.25653 [PubMed: 21509773]

Messina G, Forni S, Rosadini D, Falcone M, Collini F, & Nante N (2017). Risk adjusted mortality after hip replacement surgery: A retrospective study. Annali dell'Istituto Superiore di Sanita, 53, 40–45. doi:10.4415/ANN_17_01_09

Mitchell PH, Ferketich S, & Jennings BM (1998). Quality health outcomes model. Journal of Nursing Scholarship, 30, 43–46. doi:10.1111/j.1547-5069.1998.tb01234.x

Nante N, De Marco MF, Balzi D, Addari P, & Buiatti E (2000). Prediction of mortality for congestive heart failure patients: Results from different wards of an Italian teaching hospital. European Journal of Epidemiology, 16, 1017–1021. [PubMed: 11421469]

Olson D, Davis NL, Milazi R, Lufesi N, Miller WC, Preidis GA, … McCollum ED (2013). Development of a severity of illness scoring system (inpatient triage, assessment and treatment) for resource-constrained hospitals in developing countries. Tropical Medicine & Internationl Health, 18, 871–878. doi:10.1111/tmi.12137

Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, … Ghali WA (2005). Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical Care, 43, 1130–1139. doi:10.1097/01.mlr.0000182534.19832.83 [PubMed: 16224307]

R Core Team (2013). R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org/

Ritchie C (2007). Health care quality and multimorbidity: The jury is still out. Medical Care, 45, 477–479. doi:10.1097/MLR.0b013e318074d3c1 [PubMed: 17515773]

Romano PS, & Chan BK (2000). Risk-adjusting acute myocardial infarction mortality: Are APR-DRGs the right tool? Health Services Research, 34, 1469–1489. [PubMed: 10737448]

SAS Institute (2012). SAS 9.3 language reference: Concepts (2nd ed.). Cary, NC: SAS Institute.

Solutions T (2018). All patient refined DRGs (APR-DRGs): An overview. Retrieved from https://www.bcbst.com/providers/webinar/APRDRG.pdf

Southern DA, Quan H, & Ghali WA (2004). Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Medical Care, 42, 355–360. doi:10.1097/01.mlr.0000118861.56848.ee [PubMed: 15076812]

Stukenborg GJ, Wagner DP, & Connors AF Jr. (2001). Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. Medical Care, 39, 727–739. doi:10.1097/00005650-200107000-00009 [PubMed: 11458137]

Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, & Griffith SD (2015). A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Medical Care, 53, 374–379. doi:10.1097/mlr.0000000000000326 [PubMed: 25769057]

van Walraven C, Austin PC, Jennings A, Quan H, & Forster AJ (2009). A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical Care, 47, 626–633. doi:10.1097/MLR.0b013e31819432e5 [PubMed: 19433995]

Wu J, Liu C, Lee S, Kuo Y, & Hsieh T (2018). Assessment of the charlson comorbidity index score, CHADS2 and CHA2DS2-VASc scores in predicting death in patients with thoracic empyema. Heart & Lung, 47, 157–161. doi:10.1016/j.hrtlng.2017.12.003 [PubMed: 29370933]

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, & Smith GM (2009). Zero-truncated and zero-inflated models for count data In Mixed effects models and extensions in ecology with R (Statistics for biology and health) (pp.261–293). New York, NY: Springer.

**Table 1**

Summary of risk adjustment measures

| Measure | Author and year | Method | Type | Includes only present-on-admission comorbidities | Number of comorbidities used to create score | Includes age |
|---|---|---|---|---|---|---|
| Charlson age-comorbidity | Charlson et al. (1987), Deyo et al. (1992), & Charlson et al. (1994) | Charlson/Deyo | Non-commercial | Yes | 17 | Yes, beginning at age 40, one point is added for each 10 year increase. |
| V W Elxihauser comorbidity | Elixhauser et al. (1998) van Walraven et al. (2009) | Elxihauser | Non-commercial | Yes | 30 | No. |
| V W Elxihauser age-comorbidity | Modified from van Walraven et al. (2009) | Elxihauser | Non-commercial | Yes | 30 | Yes, beginning at age 40, one point is added for each 10 year increase. |
| APR-DRG 3M severity of illness | 3M™ Core Grouping Software v30, v31, v32 | APR-DRG | Commercial | No | NA | Yes, but algorithm is proprietary and details are not available. |
| APR-DRG 3M risk of mortality | 3M™ Core Grouping Software v30, v31, v32 | APR-DRG | Commercial | No | NA | Yes, but algorithm is proprietary and details are not available. |

*Note*. APR-DRG = All Patient Refined Diagnosis Related Group; NA = not available.

**Table 2**

Admission characteristics, severity of illness measures, and outcomes

| ***N*** | **123,641** |
|---|---|
| Characteristics | |
| Sex | |
| Female, *n* (%) | 63,689 (51.5) |
| Male, *n* (%) | 59,948 (48.5) |
| Age in years, mean (*SD*) | 62.2 (18.5) |
| Hospital | |
| Community, *n* (%) | 14,898 (12.1) |
| Tertiary/quaternary 1, *n* (%) | 50,930 (41.2) |
| Tertiary/quaternary 2, *n* (%) | 57,813 (46.8) |
| Risk adjustment measures | |
| Charlson/Deyo age-comorbidity score, mean (*SD*) | 5.0 (3.1) |
| V W Elxihauser comorbidity score, mean (*SD*) | 7.6 (8.6) |
| V W Elxihauser age-comorbidity score, mean (*SD*) | 10.3 (9.2) |
| APR-DRG 3M risk of mortality | |
| Minor, n (%) | 46,890 (37.9) |
| Moderate, n (%) | 39,927 (32.3) |
| Major, n (%) | 27,148 (22.0) |
| Extreme, n (%) | 9,676 (7.8) |
| APR-DRG 3M severity of illness | |
| Minor, n (%) | 23,343 (18.9) |
| Moderate, n (%) | 46,102 (37.3) |
| Major, n (%) | 39,155 (31.7) |
| Extreme, n (%) | 15,041 (12.2) |
| Outcome measures | |
| Hospital mortality, n (%) | 3,493 (2.8) |
| Length of stay in days, median (interquartile range) | 5 (3.0–9.0) |

*Note.* APR-DRG = All Patient Refined Diagnosis Related Group.

**Table 3**

Performance of risk adjustment measures for predicting hospital mortality using binary logistic regression ($N$ = 123,641)

| | Parameter Estimates | Model performance |
|---|---|---|
| **Model** | **Odds Ratio (95% CI)** | **C-statistics (95% CI)** |
| Charlson age-comorbidity | 1.28 [1.27, 1.29] | 0.74 [0.74, 0.75] |
| V W Elxihauser comorbidity | 1.12 [1.12, 1.12] | 0.80 [0.80, 0.81] |
| V W Elxihauser age-comorbidity | 1.12 [1.11, 1.12] | 0.81 [0.81, 0.82] |
| APR-DRG 3M risk of mortality | | |
| Minor | Reference | |
| Moderate | 5.98 [3.13, 11.4] | |
| Major | 107 [59.0, 194.1] | |
| Extreme | > 1,000 [939.3, > 1,000] | 0.94 [0.93, 0.94] |
| APR-DRG 3M severity of illness | | |
| Minor | Reference | |
| Moderate | 10.48 [3.29, 33.3] | |
| Major | 114.5 [36.82, 356.1] | |
| Extreme | > 1,000 [589, > 1,000] | 0.90 [0.90, 0.91] |

*Note.* APR-DRG = All Patient Refined Diagnosis Related Group.

**Table 4**

In-hospital mortality rates and chi-square analysis of APR-DRG 3m subclasses and in-hospital mortality ($N$= 123641)

| | APR-DRG 3M risk of mortality | | | APR-DRG 3M severity of illness | | |
|---|---|---|---|---|---|---|
| **Risk Category** | **Mortality rate (%)** | **$\chi^2$ (df=3)** | ***p* value*** | **Mortality rate (%)** | **$\chi^2$ (*df* = 3)** | ***p* value*** |
| Minor | 0.02 | 25680 | < .0001 | 0.01 | 16,517 | < .0001 |
| Moderate | 0.14 | | | 0.13 | | |
| Major | 2.45 | | | 1.45 | | |
| Extreme | 28.52 | | | 19.01 | | |

*Note.* APR-DRG = All Patient Refined Diagnosis Related Group.

*Results from Chi-square analysis

*Nurs Res*. Author manuscript; available in PMC 2020 May 01.

**Table 5**

Performance of risk adjustment measures for predicting length of stay using zero-truncated negative binomial regression ($N$ = 123,641)

| | Parameter Estimates | Model performance |
|---|---|---|
| **Model** | **Incidence Rate Ratio (95% CI)** | **Likelihood Ratio Index (%)*** |
| Charlson age-comorbidity score | 1.042 [1.041, 1.045] | 0.3 |
| V W Elxihauser comorbidity Score | 1.029 [1.028, 1.030] | 1.2 |
| V W Elxihauser age-comorbidity Score | 1.025 [1.025, 1.026] | 1.1 |
| APR-DRG 3M severity of illness | | |
| Minor | Reference | |
| Moderate | 1.44 [1.42, 1.46] | |
| Major | 2.30 [2.27, 2.33] | |
| Extreme | 5.30 [5.22, 5.39] | 6.2 |
| APR-DRG 3M risk of mortality | | |
| Minor | Reference | |
| Moderate | 1.28 [1.27, 1.30] | |
| Major | 2.11 [2.08, 2.14] | |
| Extreme | 4.16 [4.09, 4.24] | 4.3 |