Published in final edited form as:

J Biomed Inform. 2019 March ; 91: 103111. doi:10.1016/j.jbi.2019.103111.

A Method For Analyzing Inpatient Care Variability Through Physicians' Orders

Matthew C. Lenert, MS,

Dept. of Biomedical Informatics | Vanderbilt University | 2525 West End Ave. Suite 1475
Nashville, TN, USA 37203 | matthew.c.lenert@vanderbilt.edu | 1-404-984-1964

Randolph A. Miller, MD,

Dept. of Biomedical Informatics | Vanderbilt University Medical Center | Nashville, TN USA

Yevgeniy Vorobeychik, PhD, and

Dept. of Computer Science and Engineering | Washington University | St. Louis, MO USA

Colin G. Walsh, MD MA

Dept. of Biomedical Informatics | Vanderbilt University Medical Center | Nashville, TN USA

Abstract

Objective: Administrators assess care variability through chart review or cost variability to inform care standardization efforts. Chart review is costly and cost variability is imprecise. This study explores the potential of physician orders as an alternative measure of care variability.

Materials & Methods: The authors constructed an order variability metric from adult Vanderbilt University Hospital patients treated between 2013–2016. The study compared how well a cost variability model predicts variability in the length of stay compared to an order variability model. Both models adjusted for covariates such as severity of illness, comorbidities, and hospital transfers.

Results: The order variability model significantly minimized the Akaike information criterion (superior outcome) compared to the cost variability model. This result also held when excluding patients who received intensive care.

Conclusion: Order variability can potentially typify care variability better than cost variability. Order variability is a scalable metric, calculable during the course of care.

Graphical Abstract

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Care Variability; Physician Orders; Cost Variability

1. INTRODUCTION

The inpatient care two seemingly identical patients receive for the same initial condition can vary significantly within and across hospitals¹⁻⁴. Variation affects patient outcomes and often increases the cost of care⁵⁻⁷. The authors define care variability as the deviation of clinical practice from the best locally available, evidenced-based, targeted approaches. Factors beyond patients' chief concerns affect care decisions within clinical encounters, such as disease severity, comorbid conditions, social factors, availability of local specialty services, and treatment accessibility⁸⁻¹⁶. Patient preferences and values, as well, can influence this process through shared decision-making. Healthcare administrators often assess care variability through the laborious process of chart review by expert panels of clinicians¹⁷. In a landscape of evolving medical knowledge and differing patient preferences, identifying care variability at any scale remains challenging.

The literature documents two prevalent surrogate measures of care variability: nursing intensity variability and cost variability¹⁸⁻²². Nursing intensity variability metrics for a patient in a given Diagnosis Related Group (DRG) capture quality, quantity, and intensity of nursing care rendered to those specific patients¹⁹. Regular quantitative measurement of nursing intensity on an ongoing patient-by-patient basis is not feasible for busy care providers because it requires trained observers^{23,24}. The latter makes nursing intensity difficult to measure through automated means. Nursing intensity remains impractical to employ as an everyday, widespread care variability surrogate.

The current standard for measuring care variability compares the costs of inpatient stays across patients billed for the same DRG. Patients with differences in treatments presumably have corresponding differences in costs²⁵. Similar to nursing intensity, costs vary with the severity of illness^{21,22}. Variation in severity makes it difficult to predict the variation in costs for a particular set of patients. Many hospitals proficiently track the supplies, medications, and services that individual patients receive during their stay. Nevertheless, most hospitals struggle to attribute employee time costs (clinician or otherwise) to individual patient stays²⁴. Ambiguity related to personnel costs reduces the fidelity of overall cost figures. For hospitals that do not attempt to model costs at the patient level, the amount billed (charges) for a patient's encounter provides a reasonable substitute²⁶. The ubiquity of charge and/or cost data makes cost/charge variability a practical surrogate for measuring care variability.

Physicians' orders directly determine the care provided to inpatients. They convert clinicians' plans into actions. Orders comprise the primary means of delegating and communicating work in the hospital²⁷. Physician orders convey work plans asynchronously to other clinicians on a patient's care team, such as nurses. The authors hypothesized that analyzing orders could provide unique insights into patient care variability.

Most hospitals in the United States have adopted electronic order entry as part of broader electronic health record systems, making the set of orders generated for an inpatient easily accessible²⁸. In 1994–1995, faculty members in Vanderbilt's Department of Biomedical Informatics developed WizOrder, an inpatient care provider order entry (CPOE) system with integrated decision support^{29–32}. From 1995–1998, Vanderbilt University Hospital (VUH) units incrementally adopted WizOrder ward-by-ward (except the Emergency Department and Neonatal Intensive Care Unit, implemented by 2002). After subsequent customization and evolution of the system, Vanderbilt licensed WizOrder to McKesson Corporation for commercial release as Horizon Expert Orders in 2001. The commercial version was “back-installed” at VUMC in 2005–06. At VUH, clinicians used WizOrder/Horizon Expert Orders (henceforth “WizOrder”) until November 1, 2017, when a different CPOE platform replaced it. The Wizorder system did not include a poly-hierarchy to categorize orders within itself or within a complementary knowledge management system³³.

The current study evaluated performance of order-derived metrics as a substitute measure for care variability. It compared order variability results to costs - the standard surrogate for measuring care variability³⁴. Improved care variability metrics could enable hospital administrators to more accurately monitor service lines for deviation from expected outcomes, in order to design more effective interventions^{35–37}.

2. MATERIALS AND METHODS

2.1. Study Environment and Sources of Physician Order and Cost Data

The authors conducted a retrospective cohort study at a single academic institution—Vanderbilt University Medical Center (VUMC)—located in Nashville, TN. VUMC's Institutional Review Board approved this study under IRB 151156. The cohort was comprised of adult (age 18–64) patients admitted to VUH from July 1st of 2013 through December 31st of 2016, who survived their encounter, and had their encounter primarily attributed to a service within the department of internal medicine. From July 2013 through December 2016, VUH admitted 206,908 patients to the hospital, with 13,597 of that total who had a primary service of internal medicine and survived their inpatient stay.

2.2. Order Pre-processing Methods

The authors downloaded all orders for all patient encounters in the cohort from the VUMC electronic data warehouse. The fields downloaded for each order included: medical record number, encounter number, order number, start date, stop date, daily frequency, duration, order status, entry date, entry clinician ID, cosigning clinician ID, parameter text responses, user comments, service department, hold start date, hold end date, administration route, administration dosage, and an order description. “Panel” orders packaging a group of

individual orders under a single name were retained as composites. For example, the Basic Metabolic Panel (BMP) included serum sodium, potassium, creatinine, glucose, and other items. Subsequent data processing collapsed functionally similar orders into groupings related to an over-arching UMLS terminology concept. For example, the orders “Elevate HoB” and “Adjust Head of the Bed to 45 degrees” both fell under the UMLS concept “Elevation of head of bed C1827414”. Many order (e.g., nursing orders) were of the form “DEPARTMENT: [free text comment]”. The study mapped the free-text comment fields of the latter type of orders to UMLS concepts, while retaining the service department (clinical domain) of the order. Mapping free-text orders into UMLS concepts facilitated their handling, by replacing disparately named redundant orders with a single, uniform category name. Authors had attempted to use the then-current version of NLM’s MetaMap to automatically compress order concepts via natural language processing³⁸. The authors found that MetaMap could not uniformly compress redundant free-text orders without the development of heuristics to unify concepts to the same term within the UMLS hierarchy. Navigation of the UMLS over hundreds of thousands of concepts is not trivial and was beyond the scope of this study. As a result, authors used customized regular expressions and, when necessary, direct manual mapping to match fine-grained order descriptions to higher-level UMLS terminology concepts. Appendix A describes the different mapping strategies used. It also describes how authors validated the mapping of orders to concepts. For study purposes, the shallow categorizations in the homegrown WizOrder system required authors to undertake grouping of semantically redundant orders. Such an effort would not be necessary for health systems that have hierarchically categorized physician orderables, as was the case, for example, with the MED at Columbia³⁹.

The current study used different potential features to make up a variability metric. It grouped orders by encounter, DRG, and service department. Table 1 enumerates the different order features derived. The study used the unique order count to describe the variety of orders used across all encounters in a DRG/service-department group. The study used different central moments of the distribution (variance (2nd) and kurtosis (4th)) of the requested number of order completions/day to characterize differences between encounters within each group. The common order count and common order ratio both quantify how many orders were held in common across encounters within a group. To derive the common order count, the authors derived the set of unique orders issued for each encounter. Next, the authors determined which orders were present in more than half of all the encounters coded for the same DRG.

2.3. Order and Cost Variability Metric Construction

To select which order features comprised the order variability measure, the authors chose the features with the greatest positive Pearson correlation to the outcome. The correlation was initially assessed qualitatively by scatter plot analysis. This analysis involved plotting a single order feature on the X-axis and the outcome on the Y-axis. This project combined the order features into one order variability measure using an averaging method described below. A single metric permits a fairer comparison to cost variability across DRGs. In statistical models, multiple predictors have a natural advantage over one predictor in terms of model fit. The regression methodology minimizes prediction error, which means that

adding predictors cannot decrease the fit of a model. Additional predictors do not decrease model fit, because uncorrelated predictors can always be weighted to have zero effect. To weight each order-based feature equally, the authors centered and standardized each order-based feature before averaging them into a singular DRG-based measure of order variability.

The project integrated clinical outcomes with cohort-related cost data from the VUMC Finance Department. The VUMC Finance office uses the bottom up cost accounting method to determine the fixed and variable costs for each procedure and service⁴⁰. The Finance Department uses the cost data operationally, suggesting rigorous provenance. The total per-patient cost formed the basis of the study's cost variability metric.

2.4. Evaluation of Order Variability and Cost Variability

This study used a Huber-White variance-corrected regression framework using Stata for its evaluation of order variability as a measure for care variability⁴¹. The variance of the length of stay (LOS) ratio was the outcome of interest and each observation signified a DRG. The LOS ratio is a patient's actual length of stay divided by the Center for Medicare & Medicaid Services (CMS) calculated geometric average length of stay (for the same DRG). The study made use of the geometric average because the study cohort includes persons less than 65 years of age. The geometric average is always less than or equal to the arithmetic average. The authors repeated the analysis using the CMS arithmetic average to test the sensitivity of the results to the exact values of the expected length of stay. The CMS average length of stay for a DRG is also known as the expected length of stay⁴². To ensure reasonable precision for estimates of variability, DRGs were eliminated from consideration if they did not have at least ten patient encounters in the study cohort. The authors varied the minimum encounter threshold to twenty encounters and thirty encounters to evaluate the sensitivity of the results to this parameter. This study compared a robust regression model with an order variability metric to a robust regression model with cost variability. Each model used the same covariates apart from the care variability surrogate. The authors compared the fit of the models based on the Akaike information criterion (AIC) of both models. This criterion describes the information loss of a model based on the data. The model with the smallest AIC is relatively better than the other models. A difference in AIC of six or more suggests there is less than a 5% chance that the greater AIC model minimizes information loss better than the lesser AIC model. The study also reported the Student's T-Test of significance on the variable coefficients. The analysis was done at the DRG scale to develop a unified framework for evaluating care variability that could be used at the health system level by healthcare administrators or quality officers. In this framework, one DRG was compared to another. Comparing the LOS ratio between DRGs requires some form of case-mix adjustment, because the length of stay is sensitive to patient specific factors. The authors performed a sub-analysis, eliminating encounters involving intensive/critical care to determine if the order metric could predict variability in more homogenous encounter groupings.

The authors chose covariates for their care variability analysis from a review of the literature. Each model was adjusted for the intra-DRG variance of: age, sex, race, admission season, a weekday admission indicator, the Van Walraven modification of the Elixhauser

Comorbidity Index^{43,44}, admission service, an intensive care utilization indicator, an indicator for outside hospital transfers, and the Emergency Patient Severity Index (EPSI)⁴⁵. Any missing data was imputed using multiple imputation implemented in the `aregImpute` function from the `Hmisc` package in the R programming language⁴⁶.

3. Results

The study began with an initial cohort of 13,597 admissions. The authors downloaded 3,455,292 orders (1.1 GB) for the 13,597 patient encounters. Figure 1 illustrates how the selection criteria affected the total cohort size and the sub-analysis sample size. The sub-analysis excluding ICU patients further reduced the cohort by nearly a quarter.

The characteristics of the full study cohort, and all covariates appear in Table 2. The cohort encounters can include multiple temporally separate instances for a given patient.

3.1. Order & Cost Variability Construction Results

The two features with the greatest positive correlation with the variance in the LOS ratio were the number of commonly executed lab procedures (0.255) and the total number of unique rehab orders (0.241). This project scaled the features to the standard deviation scale and centered the order features at zero, so that they could be averaged to form order variability. Table 3 shows differences among DRG-related LOS ratios, ranked by order variability and cost variability.

After the authors removed patients who received ICU care, the different variability measures aligned more closely. The order variability metric tended to favor DRGs for severe infections, while cost variability tended to favor gastrointestinal and vascular DRGs (See Table 4).

3.2. Evaluation Results for Order Variability and Cost Variability

Table 5 presents the AIC, adjusted R^2 , coefficient sign, and P-value for order variability metric (order var) as well as cost variability (cost var) across encounter thresholds. Order variability consistently produced better-fit models than cost variability across all thresholds for the minimum number of encounters. This result held in the sub-group without ICU patients. The authors found this results also held after repeating the analysis using an LOS ratio incorporating the CMS arithmetic average as the expected LOS. The CMS arithmetic average results are presented in Appendix B.

4. DISCUSSION

This study analyzed the ability of a novel order variability metric to predict the variance of the LOS ratio across DRGs. It compared the new metric's performance on this task to the de facto standard, cost variability. In this preliminary study, order variability outperformed cost variability's ability to explain the variance in the LOS ratio. This result held throughout a sensitivity analysis of sample size, and in a sub-analysis excluding ICU patients.

The coefficient of order variability remained consistently positive throughout all these analyses. The coefficient for cost variability did not consistently keep the same sign. It switched signs from positive to negative when the minimum number of encounters was increased to thirty in the full cohort analysis. This inconsistency suggests that cost variability may have an inconsistent correlation with variability in patient outcomes. The interpretation of the order variability coefficient was that greater order variability contributed to greater variance in the LOS ratio after accounting for the selected confounders. Order variability may be a useful surrogate for the variance of the LOS ratio, because order variability can be derived during the admission, while the LOS ratio cannot.

The analysis was done at the level of an entire health system and across a wide variety of DRGs. This scale comprises a strength of the work. The authors' method for creating an order variability metric is applicable to all DRGs. This study used novel feature sets to attempt to identify variability in care. When modeling care variability, the current study corrected for many covariates (such as severity of illness, admission service, transferred patients, and an indicator for weekday admissions) that are often left out of analyses. One might improve the correlation of cost variability to outcome variability by taking a risk/severity-weighted variance of patient costs. The best severity of illness and risk models are generally disease process specific, leading to a large model development and maintenance costs at scale⁴⁷⁻⁴⁹. Order variability might itself benefit from risk/severity-weighted order measures. Exploring the benefits and complexity costs of weight adjustments using risk or severity is an avenue of future research.

The study had several limitations. First and foremost, the order variability metric is not prescriptive. The metric cannot forecast which interventions might improve a service line or the course of care for an individual patient. The metric may simply help healthcare administrators to identify and prioritize DRGs where services require improvement, because patients are experiencing a higher amount of unexplained variance in outcomes. Chart review and observational studies should thus still play a role in devising implementable enhancements. The study's use of data from a single academic institution formed a second limitation. Studies at other sites should determine the generalizability of the authors' findings. A further limitation involved how the study grouped patients by DRG. This approach took advantage of pre-existing CMS calculations for the expected length of stay by DRG. Other means of grouping patients exist and may be superior. This study used the expected length of stay for patients 65 years old and older. The arithmetic average LOS for younger adults is likely shorter than that for seniors, while the arithmetic average length of stay for younger adults is generally greater than the geometric average for seniors. Use of CMS data thus can generally bound the true expected average of the cohort from above and below, based on the results in Table B.2. The authors found the results were invariant to the form of the expected length of stay. This study's use of the EPSI was a crude surrogate for severity of illness; more accurate results might have followed from calculating severity using one of the Acute Physiology And Chronic Health Evaluation (APACHE) scoring systems. Nevertheless, doing so might have adversely affected the degrees of freedom of the statistical model⁵⁰. The fit of the models suggested that some covariates were missing. Because the model was under-fit, the statistical importance of order variability may have been overstated. Model diagnostics such as residual versus fitted value plots, however, did

not suggest bias or violations of key regression assumptions. Furthermore, the study only explored features the authors believed would be explanatory for order variability. The variables were not exhaustively inclusive of the possibilities but focused on one particular task. A final limitation was that the study did not differentiate the intended number of times an order should be completed (i.e., what was ordered) from the actual number of times an order was completed (i.e., what was done). Future work could overcome this by incorporating data sources such as the medical administration record (MAR) along with results from lab and imaging systems to bridge this gap.

5. CONCLUSIONS

This project examined the variability of care among inpatients with different DRGs from a novel point of view—orders. Previous research focused on cost, case-mix, and utilization differences between DRGs^{12,22,51}. Orders have the potential to elucidate care that current cost-capture mechanisms do not handle well, such as nursing utilization.

In the current small observational analysis, the authors' order metric outperformed the current standard metric (cost variability) in modeling LOS variability by DRG. Additional, larger studies should independently validate this result. The new metric cannot directly identify the root causes of care variability; it merely indicates when variability occurs.

Acknowledgements

The authors acknowledge the contributions of Jon Wolfe and the VUMC Finance Office for sharing costing data at the patient level.

Funding Sources

Matthew Lenert's effort was supported by National Library of Medicine (NLM) training grant 4T15 LM007450–15.

Appendix A:: Order Mapping Results

The authors applied a variety of concept mapping strategies, based on each order's clinical domain and other characteristics of the orders. Most of the Pharmacy orders (before downloading the EDW dataset) were mapped to the RxNorm terminology by Vanderbilt's Research Derivative Team using the Medex system⁵². Author MCL manually mapped 87,408 order descriptions to UMLS concepts. Of the total 87,174 manually mapped order descriptions, only 10,234 were pharmacy orders. The remaining manually mapped orders included all of the unique order descriptions from fourteen different service departments. Nursing orders had the greatest variety, making them the best candidate for regular expression-based mapping. Table A.1 presents the number of unique order descriptions and the terminologies and strategies applied.

A.1 :**Terminologies and Strategies Applied to Order Service Departments**

Service Department	Number of Unique Orders	Terminology Applied	Mapping Strategy
Pharmacy	678,820	RxNorm ⁵⁷	Medex ⁵⁶ , Manual Mapping
Nursing	485,858	SNOMED-CT ⁵⁸ , MEDCIN ⁵⁹ , HCPCS ⁶⁰	Regular Expressions
All Others	77,174	SNOMED-CT ⁵⁸ , MEDCIN ⁵⁹ , HCPCS ⁶⁰	Manual Mapping

For the nursing orders, the authors assigned a noun concept and an optional verb concept to describe the action. The authors identified 305 unique noun UMLS concepts and 52 unique verb/adjective UMLS concepts. The authors coded 695 regular expressions for the 305 noun concepts and 171 regular expressions for the 52 action concepts. Author MCL carried out the mapping (manually and algorithmically) of nursing orders to noun and verb concepts. Physician-authors CGW and RAM manually reviewed the mappings to validate them. The project retained only those mappings on which both physician reviewers concurred were accurate. The sampling and evaluation statistics for the order concept mapping done by direct mapping are presented in Table A.2.

Table A.2:**Expert Validation of Directly Mapped and String Parsed Orders**

Service Department	Mapping Strategy	Number of Orders Mapped	Sample Size	Accuracy
Pharmacy	Direct Mapping	10,234	200 (2%)	99%
All Others (Excludes Nursing)	Direct Mapping	77,174	200 (0.3%)	99%

The authors sampled the nursing orders by concept. All the order descriptions within the same concept needed to be approved by both physician reviewers for that concept to be labeled as accurate. The authors also recorded the accuracy based on the individual orders, instead of by concept. The validation and sampling statistics for concept mapping done by regular expressions is shown in Table A.3. Ultimately, none of the nursing order features were selected to be part of the order variability metric. The results below had no bearing on the conclusions of the study.

Table A.3:**Expert Validation of Nursing Orders Mapped by Regular Expressions**

Sample #	Number of Concepts Sampled	Accurately Mapped Concepts	Number of Orders in Sample	Order Mapping Accuracy
1	36/276 (13%)	23/36 (64%)	2,945/485,858 (1%)	93%
2	42/313 (13%)	28/42 (67%)	3,319/485,858 (1%)	97%
3	20/305 (7%)	14/20 (55%)	48/485,858 (0.1%)	85%

Appendix B:: Order Variability Analysis Using CMS Arithmetic Average LOS

Table B.1 demonstrates that the results are largely invariant to differences in the expected LOS. The majority of the same relationships hold when using the geometric average LOS or the arithmetic average LOS from CMS.

Table B.1:

Care Variability Analysis with Variance of the LOS Ratio as the Outcome

	DRGs with 10+ Encounters	DRGs with 20+ Encounters	DRGs with 30+ Encounters
Full Cohort			
# of Encounters	12,598	11,391	10,574
# of DRGs	226	138	104
Adjusted-R ²			
Order Var	0.246	0.128	0.330
Cost Var	0.183	0.097	0.284
AIC			
Order Var	-64.38 *	-76.88	-116.67 *
Cost Var	-46.51	-71.97	-109.90
Coefficient Sign			
Order Var	+	+	+
Cost Var	+	+	+
Coefficient P-Value			
Order Var	0.001 *	0.069	0.063
Cost Var	0.375	0.494	0.705

	DRGs with 10+ Encounters	DRGs with 20+ Encounters	DRGs with 30+ Encounters
Non-ICU Cohort			
# of Encounters	8,893	7,913	7,048
# of DRGs	188	115	79
Adjusted-R ²			
Order Var	0.152	0.165	0.231
Cost Var	0.103	0.123	0.186
AIC			
Order Var	-46.34 *	-72.36	-83.79
Cost Var	-35.81	-66.70	-81.37
Coefficient Sign			
Order Var	+	+	+
Cost Var	+	+	-
Coefficient P-Value			
Order Var	0.006 *	0.023 *	0.173
Cost Var	0.940	0.182	0.792

* signifies statistical significance

+ signifies that the LOS ratio variance is positively associated with that variable

⁻ signifies that the LOS ratio variance is negatively associated with that variable

Table B.2 shows how the expected length of stay varies for the most common DRGs in the dataset. The cohort average (avg) was generally greater than the CMS geometric average, and less than the CMS arithmetic average.

Table B.2:

Most Prevalent DRGs and Associated Average Length of Stays

Variability Ranked DRGs (Full Cohort)	Number of Encounters	CMS Geometric Avg	CMS Arithmetic Avg	Cohort Avg
1. Septicemia or severe sepsis w/ complications	802	4.9	6.4	6.0
2. Cellulitis w/o complications	568	3.3	4.0	3.1
3. Esophagitis, gastroenterologic & miscellaneous digestive disorders w/o complications	496	2.7	3.3	3.2
4. Septicemia or severe sepsis w/o complications	421	3.7	4.5	3.6
5. Diabetes w/ complications	344	3.0	3.6	3.0
6. Alcohol/drug abuse or dependence w/o rehabilitation s	295	3.4	4.3	2.9
7. Kidney & urinary tract infections w/o complications	279	3.0	3.6	3.0
8. Poisoning and toxic effects of drugs w/ complications	271	3.5	4.8	3.8
9. Poisoning and toxic effects of drugs w/o complications	253	2.3	3.0	2.3
10. Renal failure w/ complications	245	3.0	3.6	2.6

References

1. Kime C Patterns of Inpatient Care for Newly Diagnosed Immune Thrombocytopenia in US Children's Hospitals. *Pediatrics* 2013;131:880. doi:10.1542/peds.2012-2021. [PubMed: 23569091]
2. Macias CG, Mansbach JM, Fisher ES, et al. Variability in inpatient management of children hospitalized with bronchiolitis. *Acad Pediatr* 2015;15(1):69–76. doi:10.1016/j.acap.2014.07.005. [PubMed: 25444654]
3. Rand CS, Powe NR, Wu AW, Wilson MH. Why Don't Physicians Follow Clinical Practice Guidelines? *JAMA J Am Med Assoc* 1999;Vol 282(15):1458–1465. doi:10.1001/jama.282.15.1458.
4. Francke A, Smit M, de Veer A, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: a systematic meta-review. *BMC med Inf decis mak* 2008;8:38. doi:10.1186/1472-6947-8-38.
5. Rozich JD, Howard RJ, Justeson JM, Macken PD, Lindsay ME, Resar RK. Standardization as a mechanism to improve safety in health care. *Jt Comm J Qual Saf* 2004;30(1):5–14. doi:10.1016/S1549-3741(04)30001-8. [PubMed: 14738031]

6. Bates DDW, Boyle DL, Vander Vliet MB, Schneider J, Leape L Relationship between medication errors and adverse drug events. *J Gen Intern Med* 1995;10(4):199–205. doi:10.1007/BF02600255. [PubMed: 7790981]
7. Ross MA, Hockenberry JM, Mutter R, Barrett M, Wheatley M, Pitts SR. Protocol-driven emergency department observation units offer savings, shorter stays, and reduced admissions. *Health Aff* 2013;32(12):2149–2156. doi:10.1377/hlthaff.2013.0662.
8. Jacke CO, Albert US, Kalder M. The adherence paradox: guideline deviations contribute to the increased 5-year survival of breast cancer patients. *BMC Cancer* 2015;15:734. doi:10.1186/s12885-015-1765-0. [PubMed: 26481452]
9. Britton DJ, Bloch RB, Strout TD, Baumann MR. Impact of a computerized order set on adherence to centers for disease control guidelines for the treatment of victims of sexual assault. *J Emerg Med* 2013;44(2):528–535. doi:10.1016/j.jemermed.2012.06.016. [PubMed: 22921854]
10. Cisse B, Moore L, Kuimi BLB, et al. Impact of socio-economic status on unplanned readmission following injury: A multicenter cohort study. *Injury* 2016;47(5):1083–1090. doi:10.1016/j.injury.2015.11.034. [PubMed: 26746984]
11. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;306(15):1688–1698. doi:10.1001/jama.2011.1515. [PubMed: 22009101]
12. Williams L Modified DRGs as Evidence for Variability in Patient Severity. *Med Care* 1988;26(1):53–61. [PubMed: 3121959]
13. Donzé J, Lipsitz S, Bates DW, Schnipper JL. Causes and patterns of readmissions in patients with common comorbidities: retrospective cohort study. *BMJ* 2013;347(dec16_4):f7171. doi:10.1136/bmj.f7171. [PubMed: 24342737]
14. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. *Inst Med* 2001;(March):1–8. doi:10.17226/10027.
15. Maas J, Verheij RA, Groenewegen PP, de Vries S, Spreeuwenberg P. Green space, urbanity, and health: how strong is the relation? *J Epidemiol Community Heal* 2006;60(7):587–592. doi:10.1136/jech.2005.043125.
16. Betihavas V, Davidson PM, Newton PJ, Frost SA, Macdonald PS, Stewart S. What are the factors in risk prediction models for rehospitalisation for adults with chronic heart failure? *Aust Crit Care* 2012;25(1):31–40. doi:10.1016/j.aucc.2011.07.004. [PubMed: 21889893]
17. Shahar Y, Musen M. Plan Recognition and Revision in Support of Guideline-Based Care. *Work notes AAAI Spring Symp ...* 1995:118–126.
18. Williams L Variation in Resource Use within Diagnosis-Related Groups: The Effect of Severity of Illness and Physician Practice. *Med Care* 1986;24(5):388–397. [PubMed: 3084888]
19. Thompson JD. The measurement of nursing intensity. *Health Care Financ Rev* 1984;Suppl(Suppl):47–55. doi:hcf-84-supp-047 [pii]. [PubMed: 10311076]
20. Pirson M, Delo C, Di Pierdomenico L, Laport N, Biloque V, Leclercq P. Variability of nursing care by APR-DRG and by severity of illness in a sample of nine Belgian hospitals. *BMC Nurs* 2013;12(1):26. doi:10.1186/1472-6955-12-26. [PubMed: 24112381]
21. Horn SD, Sharkey PD, Chambers AF, Horn RA. Severity of illness within DRGs: Impact on prospective payment. *Am J Public Health* 1985;75(10):1195–1199. doi:10.1097/00005650-198603000-00005. [PubMed: 3929632]
22. Smits HL, Fetter RB, McMahon LF. Variation in resource use within diagnosis-related groups: The severity issue. *Health Care Financ Rev* 1984;1984(Suppl):71–78. doi:hcf-84-supp-071 [pii].
23. Pirson M, Delo C, Di Pierdomenico L, Laport N, Biloque V, Leclercq P. Variability of nursing care by APR-DRG and by severity of illness in a sample of nine Belgian hospitals. *BMC Nurs* 2013;12(1):26. doi:10.1186/1472-6955-12-26. [PubMed: 24112381]
24. Welton JM, Fischer MH, DeGrace S, Zone-Smith L. Hospital nursing costs, billing, and reimbursement. *Nurs Econ* 2006;24(5):227,239–245,262 <http://www.ncbi.nlm.nih.gov/pubmed/17131615> <http://graphics.tx.ovid.com/ovftpdfs/FPDDNCDCCDDCAN00/fs046/ovft/live/gv023/00006073/00006073-200609000-00004.pdf>. [PubMed: 17131615]

25. Abbass I Variability in the Initial Costs of Care and One-Year Outcomes of Observation Services. *West J Emerg Med* 2015;16(3):395–400. doi:10.5811/westjem.2015.2.24281. [PubMed: 25987913]
26. Shwartz M, Young DW, Siegrist R. The Ratio of Costs to Charges: How Good a Basis for Estimating Costs? *Inquiry* 1995;32(4):476–481. [PubMed: 8567084]
27. Berger RG, Kichak JP. Computerized physician order entry: helpful or harmful? *J Am Med Informatics Assoc* 2004;11:100–103. doi:10.1197/jamia.M1411.In.
28. Adler-Milstein J, DesRoches CM, Kralovec P, et al. Electronic Health Record Adoption In US Hospitals: Progress Continues, But Challenges Persist. *Health Aff* 2015;34(12):2174–2180. doi: 10.1377/hlthaff.2015.0992.
29. Geissbuhler A, Miller RA. WizOrder, a user-friendly interface for order entry and clinical decision support tools. In: *Proceedings of the Annual Symposium on Computer Application in Medical Care*. American Medical Informatics Association; 1995:1002.
30. Geissbühler A, Miller RA. A new approach to the implementation of direct care-provider order entry. In: *Proceedings of the AMIA Annual Fall Symposium*. American Medical Informatics Association; 1996:689.
31. Geissbuhler A, Miller RA. Distributing knowledge maintenance for clinical decision-support systems: the “knowledge library” model. In: *AMIA Annual Symposium*; 1999:770–774. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2232510&tool=pmcentrez&rendertype=abstract>.
32. Starmer JM, Talbert DA, Miller RA. Experience using a programmable rules engine to implement a complex medical protocol during order entry. In: *Proceedings of the AMIA Symposium*. American Medical Informatics Association; 2000:829.
33. Cimino James J. “Desiderata for controlled medical vocabularies in the twenty-first century.” *Methods of information in medicine* 374–5 (1998): 394. [PubMed: 9865037]
34. Teich JM, Merchia PR, Schmiz JL, Kuperman GJ, Spurr CD, Bates DW. Effects of computerized physician order entry on prescribing practices. *Arch Intern Med* 2000;160:2741–2747. doi: 10.1001/archinte.160.18.2741. [PubMed: 11025783]
35. Neilson EG, Johnson KB, Rosenbloom ST, et al. The impact of peer management on test-ordering behavior. *Ann Intern Med* 2004;141(3):196–204. [PubMed: 15289216]
36. Schwartz AL, Chernew ME, Landon BE, McWilliams JM. Changes in Low-Value Services in Year 1 of the Medicare Pioneer Accountable Care Organization Program. *JAMA Intern Med* 2015;02115(11):1–11. doi:10.1001/jamainternmed.2015.4525.
37. James BC, Savitz LA. How Intermountain Trimmed Health Care Costs Through Robust Quality Improvement Efforts. *Health Aff* 2011;30(6):1185–1191. doi:10.1377/hlthaff.2011.0358.
38. Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In: *Proceedings of the AMIA Symposium*. American Medical Informatics Association; 2001:17.
39. Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based Approaches to the Maintenance of a Large Controlled Medical Terminology. *J Am Med Inf Assoc* 1994;1(1):35–50. doi:10.1080/02693799408901993.
40. Azoulay A, Doris NM, Filion KB, Caron J, Pilote L, Eisenberg MJ. The use of the transition cost accounting system in health services research. *Cost Eff Resour Alloc* 2007;5:1–9. doi: 10.1186/1478-7547-5-11. [PubMed: 17233893]
41. StataCorp. 2017 Stata Statistical Software: Release 15 College Station, TX: StataCorp LLC.
42. CMS. Department of Health and Human Services Regulations. *Fed Regist* 2013;78(184).
43. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity Measures for Use with Administrative Data. *Med Care* 1998;36(1):8–27. [PubMed: 9431328]
44. Van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *CMAJ* 2010;182(6):551–557. doi:10.1503/cmaj.091117. [PubMed: 20194559]
45. Wuerz RC, Milne LW, Eitel DR, Travers D, Gilboy N. Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 2000; 7:236–42.
46. Harrell JFE. Package ‘Hmisc’ 2018.

47. Shindo Y, Sato S, Maruyama E, et al. Comparison of severity scoring systems A-DROP and CURB-65 for community-acquired pneumonia. *Respirology* 2008;13(5):731–735. doi:10.1111/j.1440-1843.2008.01329.x. [PubMed: 18713094]
48. Walsh CG, Sharman K, Hripesak G. Beyond discrimination: A comparison of calibration methods and clinical usefulness of predictive models of readmission risk. *J Biomed Inform* 2017;76(October):9–18. doi:10.1016/j.jbi.2017.10.008. [PubMed: 29079501]
49. Macias CG, Mansbach JM, Fisher ES, et al. Variability in inpatient management of children hospitalized with bronchiolitis. *Acad Pediatr* 2015;15(1):69–76. doi:10.1016/j.acap.2014.07.005. [PubMed: 25444654]
50. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13(10):818–829. <http://europepmc.org/abstract/MED/3928249>. [PubMed: 3928249]
51. Schreyögg J, Tiemann O, Busse R. Cost accounting to determine prices: How well do prices reflect costs in the German DRG-system? *Health Care Manag Sci* 2006;9:269–279. doi:10.1007/s10729-006-9094-0. [PubMed: 17016933]
52. Xu H, Stenner SP, Doan S, Johnson KB, Waitman LR, Denny JC. MedEx: a medication information extraction system for clinical narratives. *J Am Med Informatics Assoc* 2010;17(1):19–24.

Highlights

- Assessing variability in patient care is hard to do at the scale of a hospital
- Many institutions use cost variability instead to make data-driven decisions
- Order data may quantify care variability better than cost variability

Figure 1:
Cohort Sample Size with and without Sub-Analysis

Table 1:

Order Variability Features and Descriptions

Feature Name	Description
Kurtosis of Completions/Day	Average of the Average Number of Order Completions per Day
Variance of Completions/Day	Variance of the Average Number of Order Completions per Day
Unique Order Count	Count of Unique Terminology Concepts Across All Encounters
Common Order Count	Count of Terminology Concepts Present in > 50% of Encounters for the Group
Common Order Ratio	<u>Common Order Count</u> Unique Order Count

Table 2:

Study Cohort Characteristics

Variable	
Unique encounters (count)	13,597
Unique patients (count)	10,081
Age at admission	44.9 ± 13.1 years
Elixhauser index	1.3 ± 0.1
Female sex	50.6%
Received intensive care	23.2%
Weekday Admission	75.5%
Transfers	3.0%
Race	
White	73.3%
Black	22.6%
Asian	1.1%
Unknown or other	3.0%
Season of Admission	
Spring (Mar-May)	21.9%
Summer (June-Aug)	26.8%
Fall (Sep-Nov)	27.9%
Winter (Dec-Feb)	23.4%
Admission Service	
Internal Medicine	74.6%
Emergency Room	3.3%
Infectious Disease	1.0%
General Surgery	1.0%
Others Not Listed	20.1%

Table 3:

Top 5 DRGs of the Full Cohort When Sorted by Variability

Variability Ranked DRGs (Full Cohort)	Number of Encounters	Variance of LOS Ratio
Order Variability		
1. Tracheostomy w/ ventilator support > 96hr	18	1.53
2. Septicemia or severe sepsis w/ ventilator support > 96hr	41	1.47
3. Septicemia or severe sepsis w/ complications	802	1.51
4. Pulmonary embolism w/ complications	44	1.74
5. Respiratory infections & inflammations w/ complications	97	1.50
Cost Variability		
1. Osteomyelitis with complications	30	1.45
2. Alcohol/drug abuse or dependence w/o rehabilitation	295	0.90
3. Fever	40	1.12
4. Other respiratory system operation w/ complications	37	1.85
5. Intracranial hemorrhage or cerebral infarction w/ complications	23	1.28

Table 4:

Top 5 DRGs of the Non-ICU Cohort When Sorted by Variability

Variability Ranked DRGs (Non-ICU Cohort)	Number of Encounters	Variance of LOS Ratio
Order Variability		
1. Respiratory infections & inflammations w/ complications	75	1.36
2. Infectious & parasitic diseases w/ surgery w/ complications	53	1.23
3. Septicemia or severe sepsis w/ complications	382	1.21
4. Other vascular procedures w/ complications	23	2.04
5. Other digestive system diagnoses w/ complications	54	1.31
Cost Variability		
1. Coagulation disorders	53	1.16
2. Other digestive system diagnoses w/ complications	54	1.31
3. Gastrointestinal hemorrhage	24	1.23
4. Non-extensive surgery unrelated to principal diagnosis w/ complications	21	1.65
5. Other vascular procedures w/ complications	23	2.04

Table 5:

Care Variability Analysis with Variance of the LOS Ratio as the Outcome

	DRGs with 10+ Encounters	DRGs with 20+ Encounters	DRGs with 30+ Encounters
Full Cohort			
# of Encounters	12,598	11,391	10,574
# of DRGs	226	138	104
Adjusted-R²			
Order Var	0.276	0.215	0.423
Cost Var	0.232	0.188	0.387
AIC			
Order Var	179.87 *	76.89	-7.64 *
Cost Var	193.07	81.53	-0.21
Coefficient Sign			
Order Var	+	+	+
Cost Var	+	+	-
Coefficient P-Value			
Order Var	0.001 *	0.084	0.053
Cost Var	0.206	0.373	0.954

	DRGs with 10+ Encounters	DRGs with 20+ Encounters	DRGs with 30+ Encounters
Non-ICU Cohort			
# of Encounters	8,893	7,913	7,048
# of DRGs	188	115	79
Adjusted-R²			
Order Var	0.193	0.277	0.237
Cost Var	0.146	0.238	0.154
AIC			
Order Var	133.17 *	38.86 *	-3.56 *
Cost Var	143.86	44.86	4.67
Coefficient Sign			
Order Var	+	+	+
Cost Var	+	+	+
Coefficient P-Value			
Order Var	0.002 *	0.007 *	0.081
Cost Var	0.381	0.038 *	0.934

* signifies statistical significance

+ signifies that the LOS ratio variance is positively associated with that variable

- signifies that the LOS ratio variance is negatively associated with that variable