Impact of Health IT on Nature and Severity of Malpractice Claims

Completed Research

Deepti Singh
Information Systems Decision Sciences,
Muma College of Business,
University of South Florida
deeptisingh@mail.usf.edu

Shivendu Shivendu
Information Systems Decision Sciences,
Muma College of Business,
University of South Florida
shivendu@usf.edu

Abstract

Healthcare costs in the US are rising at an alarming rate and medical liability claims are significant contributors. along with financial losses, medical malpractice compromises patient safety and provider reputation. Several federal and state initiatives like Meaningful Use work towards Health IT adoption and use. We investigate the impact of Health IT on nature and severity of malpractice claims reported in Florida from 2011-2016. We provide a novel method to estimate information failures in malpractice claims using machine learning. We found the relative risk of malpractice claims with high level of medical error get reduced by 87-89% with each successful attestation to MU-Stage 2 and by 52-62% for MU-Stage 1. The information failure in malpractice claims shows a reduction with MU-Stage 1 attestation but shows increase with MU-Stage 2. We provide evidence of direct positive value creation by Health IT for to Healthcare Providers and discuss its policy implications.

Keywords

Health Information Technology, Malpractice claims, Meaningful use of IT, Information failure

Introduction

Burgeoning national healthcare costs, estimated to rise to 19.9% share of GDP by 2025, have been a major concern for the healthcare industry as well as policy makers (Sahadi, 2018). The medical liability system has been shown to be a major contributor to these costs because substantially high medical liability risks give rise to a culture of defensive medicine and overtreatment (Mello et al., 2010). The medical malpractice claims against service providers are a key part of the medial liability system and are often based on probable medical errors that not only create significant costs for healthcare providers, but also hurt patients. Although the major impact of any medical error in terms of adverse health outcomes is borne by the patient, healthcare providers also suffer financial costs and reputation loss when malpractice cases are filed. The financial risk can often be mitigated through malpractice insurance, but risk to reputation persists. Moreover, medical errors resulting in malpractice claims negatively affect quality of care. Prior research has shown that the malpractice laws in themselves are inadequate in promoting healthcare quality in absence of technological support systems (Minami et al. 2017).

With the advances in Health Information Technology (HIT), there has been significant interest in understanding the mechanisms through which IT can produce better outcomes and value for the patients, healthcare providers and their organizations (Devaraj and Kohli 2003). Although HIT promises benefits, physician resistance along with the up-front financial investments, lack of technical expertise, clinical workflow disruption, and reduced quality time with patients, were reported as some common barriers in early 2000s (Boonstra and Broekhuis 2010). In recent years, despite drastic improvements in HIT adoption rates, the healthcare industry struggles to derive value from HIT systems. Consequently, healthcare providers do not perceive benefits of HIT as offsetting the potential costs and thereby lack motivation to focus on technological advancements in the clinical and administrative processes.

Federal initiatives like Meaningful-Use incentive program were introduced by the Centers for Medicare/Medicaid Services (CMS) and the Office of the National Coordinator for Health IT (ONC) to overcome preliminary barriers in order to promote adoption and meaningful use of HIT. While these initiatives provide incentives towards meaningful use of HIT, they are not met by overwhelming enthusiasm by healthcare providers, raising the question of how else the benefits of HIT might be conceptualized in terms of administrative and clinical outcomes. The question is whether meaningful use of HIT benefits healthcare providers in ways never realized before. If so, then how? What are the mechanisms through which this effect is observed?

We posit that HIT impacts patient information acquisition and processing and provides an effective decision support system to the healthcare providers. One of the key variable that can be impacted by the efficiency and effectiveness of patient information processing is the likelihood of a malpractice claim against the provider. The textual analysis of the description of the basis for the claim can also provide insights to determine if the alleged medical error was due to information failure. In other words, HIT may impact the likelihood of a medical error and thereby may directly impact the provider's potential costs due to malpractice claims. This motivates us to investigate the following two research questions: (1) How does the Health IT capability of healthcare providers impact the nature and likelihood of malpractice claims against them? and (2) How does Health IT capability impact the probability of information-failures reported in the malpractice claims?

Prior research on impact of HIT on health outcomes has been inconclusive (Black et al. 2011). While some researchers provide positive support (Hersh et al., 2014), others find these studies unconvincing due to limited sample size or weak methodology. Moreover, some critical gaps in the previous research that examines the impact of HIT have also been highlighted. With an extensive focus on adoption of the technology rather than value derived from HIT systems, researchers have faced issues related to improper measurements (Barnett et al., 2016). Certain other groups of researchers have been limited in focus covering only a subsection of patient population and single hospitals, losing adequate representation and generalizability (Appari et al., 2013).

A recent study in Pennsylvania hospitals looks at the moderating effect HIT and the impact of advanced Electronic Medical Records (EMRs) on patient safety indicators, Malpractice Insurance Premiums (MIP) as well as patient quality of care (Menon and Kohli 2013). Lately, researchers used data from hospitals to examine the impact of HIT on patient outcomes like length of stay (Wani and Malhotra 2018). While these studies marked considerable improvement over prior methodological and data related limitations, all of them focus on the impact of HIT on patient outcomes at the hospital-level. Our study is different in two key aspects. First, we use a large data set at the level of individual medical practioners for the state of Florida, and therefore, our research focuses on the physicians. Second, we directly take into considertaion the impact of HIT on medical error and its direct consequences for the physicans. Therefore, we bridge the gap in existing literature by investigating business value creation by HIT for physicians and service providers. Our study provides critical support for healthcare organizations to justify their investments and organizational focus. Our study is neither limited to any special category of patient population, nor any specialty of healthcare providers. This overcomes the biases associated with limited focus and applicability in prior research.

We employ a topic modeling methodology from computational linguistic research to extract thematic content from malpractice claim reports to examine information-based failures. We further investigate how HIT impacts these failures closely linked to risks associated with patient safety and quality of service in malpractice claims. We extend the methodology used by (Huang et al., 2014) to Information Sciences (IS) in healthcare management domain by utilizing improved methods to analyze case details and medical records through unsupervised machine learning.

We examine the impact of HIT on service providers by analyzing medical malpractice claims reported against licensed medical professionals in Florida between 2011 and 2016. When licensed health providers attested for MU Stage-1, we found the relative risk of a malpractice claim being reported and not settled reduced by 62.32% while the relative risk of a claim being reported and settled reduced by 52.21%. Similarly, when licensed health providers attested for MU Stage-2, we found the relative risk of a malpractice claim being reported and not settled reduced by 89.12% while the relative risk of a claim being reported and settled reduced by 87.38%. The two groups of claims differ as claim settlements indicate substantial evidence discovery of probable medical error by the insurance-carrier of the defendant. In both cases, the likelihood of malpractice claims, and probability of medical errors is reduced with an increase in demonstrated HIT capability measured through meaningful use attestation. Therefore, we find a positive impact of HIT on the healthcare providers; this positive impact increases with increased HIT capability. Our study provides evidence that HIT capability directly benefits service providers by lowering the likelihood of a malpractice claim. In other words, we recommend that the healthcare providers should embrace HIT to realize administrative and clinical improvements.

Meaningful-Use Stages

The Meaningful-Use (MU) program promotes usage of HIT in a meaningful way for improving healthcare outcomes. CMS grants an incentive payment to Eligible Professionals (EPs) or Eligible Hospitals (EHs), who can attest that they have engaged in efforts to adopt, implement or upgrade certified EHR technology (CEHRT). The MU Program is rolled out in a phased approach, which is divided into three stages: 2011 (data capture and sharing), 2013 (advanced clinical processes) and 2015 (improved outcomes). The incentive payments range from $44,000 over 5 years for Medicare providers and $63,750 over 6 years for Medicaid providers (starting in 2011). Participation in the MU program is voluntary, however if EPs or EHs fail to join in by 2015, there will be negative adjustments to their Medicare/Medicaid fees starting at a 1% reduction and escalating to a 3% reduction by 2017 and beyond.

The Medicare and Medicaid MU Programs designed to measure the use of CEHRT has three stages:

- Stage 1 defined requirements for the electronic capture of clinical data
- Stage 2 expanded upon the Stage 1 criteria with a goal of enhancing clinical processes through encouraged use of CEHRT for continuous quality improvement at the point of care and the exchange of health information in standardized structured format.
- Stage 3 focused on using CEHRT to improve healthcare outcomes.

Malpractice Claims

A medical malpractice claim originates when the patient believes the quality of care provided was compromised either with a diagnosable, avoidable and/or preventable medical error, or medical negligence. Medical facilities and insurance carriers thoroughly evaluate risks associated with each malpractice claims proceeding to trial.

Settlement is a favorable option compared to trial, since it allows the client and the law firm to intelligently participate in a non-binding negotiation process. Sometimes the negotiations are handled between counsel, but often mediation is utilized. If a sum is achieved during these negotiations that fully and fairly compensates the plaintiff for their injuries, and the plaintiff is willing to accept that sum, then the settlement successfully concludes the case. If the settlement process fails with direct negotiations, mediation, and arbitration, trial takes place. The results at trial are neither certain nor guaranteed.

With the involvement of medical facilities and insurance carriers, settlement is the most common outcome. Most often, settlement happens after the Medical Malpractice Tribunal requirement is satisfied for the case to proceed to litigation. In our study, we categorize claims based on the malpractice insurance-carrier's treatment. The insurance-carrier's willingness to settle is directly related to the strength of evidence discovered against a medical error or negligence during their own assessment of claim risk. The insurance-carrier with ample defensible evidence against any medical error or negligence in the case, generally proceeds to trial and does not settle. This brings us to two categories of medical claims based on their outcome – claims that settle before trial and claims that proceed to trial. Insurance-carrier's perceived medical error is high for the claims that settle before trial and low for those which proceed to trial.

Literature Review

Our research connects to three streams of literature, namely impact of health IT on quality of care, application of machine learning methods in information systems research, and medical errors and malpractice claims.

Impact of Health IT on Quality of Care

Extant literature describes the impact of HIT on healthcare outcomes and quality of care with inconclusive evidence (Appari et al. 2013; Menon and Kohli 2013). Positive impact of HIT is expected to reduce likelihood and severity of malpractice claims. On the contrary, use of HIT can be a non-trivial overhead which interferes with the clinical decision making process and increases the likelihood of medical errors (Vartak et al., 2009). Previous studies provide insight into the impact of HIT on hospital-level patient outcomes with inadequate focus on the provider-based outcomes. Our study bridges this critical gap in analysis.

Machine-learning Methods and Application

Following the work of (Huang et al. 2014) and (Bao and Datta 2014), we quantify our empirical measure of information-failures in malpractice claims records by extracting meaningful topics from the detailed records and by using a topic modeling method called Latent Dirichlet Allocation (LDA) (Blei et al., 2003). We estimate the optimum number of topics by using log likelihood and perplexity scores. Our topic numbers range from 2 to 120 for each topic model while we keep the α and β parameters at 0.1 and 0.01 in accordance to previous literature (Huang et al. 2014; Steyvers and Griffiths 2006). We contribute to the literature on application of machine learning methods in healthcare management by employing topic modeling approach for extracting relevant information from medical malpractice claim records, which are otherwise difficult to read and process.

Medical Errors and Malpratice Claims

Impact of HIT has been previously examined on quality of patient care and risk associated with malpractice insurance premiums (Menon & Kohli, 2013). Also, studies have established association of HIT with a decrease in inappropriate clinical practice variability and ineffective therapies in order to reduce medical errors (Akenroye et al., 2014). Prior studies related to the economics of medical errors and adverse drug events focus on hospital performance or patient health outcomes (David et al., 2013). We contribute to this body of literature with investigation of impact of HIT on licensed medical professionals and their outcomes. We essentially examine probable medical errors and information-failures in reported medical malpractice events.

Hypotheses Development

First, in order to estimate the impact of meaningful use of HIT on healthcare providers in terms of the likelihood of facing a malpractice claim, we investigate the following first set of hypotheses.

Hypothesis 1a: Practitioners attested for MU Stage-1 are associated with lower level of medical error in reported malpractice claims as compared to those who have not attested to any MU Stage.

Hypothesis 1b: Practitioners attested for MU Stage-2 are associated with lower level of medical error in reported malpractice claims as compared to those who have not attested to any MU Stage.

Hypothesis 1c: Practitioners attested for MU Stage-2 are associated with lower level of medical error in reported malpractice claims as compared to those who have attested for MU Stage-1.

Prior research has provided evidence that use and assimilation of HIT reduces information-failures leading to medical errors including wrong diagnosis and medication as well as possible complications during care (Menon & Kohli, 2013). This leads us to examine second set of hypotheses relating to information-failures and meaningful use of IT.

Hypothesis 2a: Practitioners attested for MU Stage-1 are associated with lower probability of information-failures reported in malpractice claims as compared to those who have not attested to any MU Stage.

Hypothesis 2b: Practitioners attested for MU Stage-2 are associated with lower probability of information-failures reported in malpractice claims as compared to those who have not attested to any MU Stage.

Empirical Methods

Our study examines the impact of changes in demonstrated Health IT capability on the likelihood and severity of medical error during malpractice event among licensed medical professionals benefitting from Medicare/Medicaid insurance programs in Florida.

Datasets

Our dataset has been merged together from five independent sources. The first source of data is Florida Office of Insurance Regulation's Professional Liability Claims Reporting Database (PLCR) which provided information about Medical Professional Liability (MPL) claims reported during the time period of interest 2011-2016. The second source of data is the Agency for Health Care Administrations (AHCA) Medicaid Program Meaningful-Use Stage attestation records for 2011-2016. The third source is Centers for Medicare and Medicaid (CMS) public-use data files for Meaningful-Use attestation data for practitioners benefiting from the Medicare program. The fourth dataset is sourced from the Florida Department of Health's Practitioner Profile Dataset. This profile data is filed by the practitioners with the Division of Medical Quality Assurance as required by law. The fifth dataset is sourced from CMS's National Plan and Provider Enumeration System (NPPES) public-use datafiles to match the license numbers of practitioners with their individual NPIs. All five datasets are merged together for the time period 2011-2016 capturing all the licensed medical practitioners in Florida with reported medical malpractice claims along with their demonstrated HIT capability.

Probabilistic Information-Failure

We derive our information-failure empirical measure by extracting meaningful topics from claim-misdiagnosis details based on the prior research method of using probabilistic topic modeling. The extracted details of claims get textually preprocessed through stemming and exclusion of stop words. We then use perplexity scores and log likelihood to determine optimum number of topics within the range of 2 to 15. The results of four metrics produced, using minimization (Arun et al. 2010; Cao et al. 2009) and maximization (Deveaud et al. 2014; Griffiths et al. 2004) of perplexity scores to find the extremum, are depicted in the graph in Figure 1. The optimal number of topics was extracted with individual term distributions as 5.

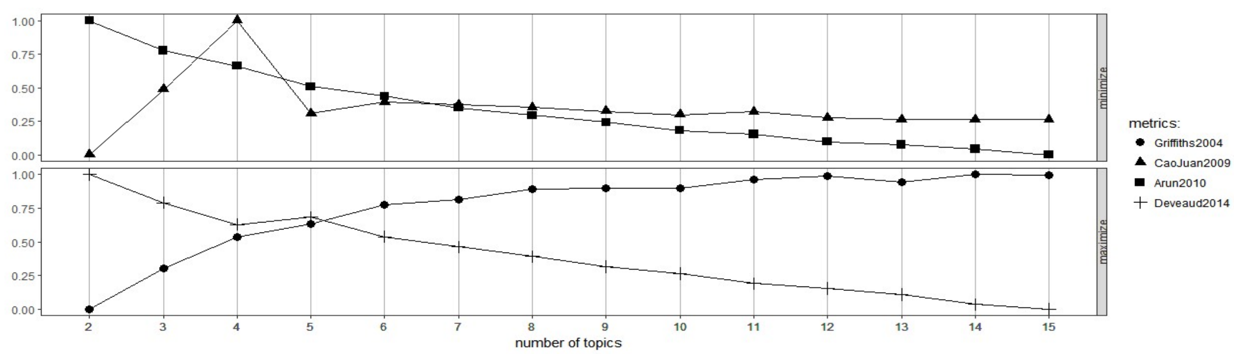

Figure 1: Number of Topics Selection through perplexity scores and log likelihood

The term distribution in extracted topics were manually analyzed to determine if the topic relates to information-failures resulting in medical errors and malpractice claims. Of the five topics extracted, terms of topic numbers 1, 4, and 5 indicated information-failures and then probabilities of these topics in each claim document were added together to get the probabilistic estimate of information-failure content of medical errors in claim details. This measure was used in quantifying information-failure in medical malpractice reports in our second econometric model.

Key Variables

We develop a multinomial ordered response variable based on whether a settlement was reached between the parties, in order to correctly estimate insurance-carrier's perceived medical error in a malpractice claim.

The claims settled before trial indicate a high level of perceived medical error discovered by the insurance carrier associating higher risk proceeding to trial. Whereas claims not settled indicate a dispute where the insurance carrier possesses credible evidence to defend the medical provider with low level of perceived medical error or negligence. Our ordinal response variable PME takes values None, Low, High in increasing order corresponding situations of no-claim-reported, claim-reported-not-settled, and claim-reported-and-settled respectively. Essentially, we are map underlying, naturally ordered insurance-carrier's risk assessment scale to a discrete ordered observed outcome of malpractice claim's final settlement.

Variable	**Description**
PME	Multinomial ordered response variable for medical error perceived by the insurance carrier during evidence discovery (None < Low < High)
IncentiveProgram	Incentive programs run by CMS and AHCA in Florida (Medicare, Medicaid)
MUStage	Meaningful-Use Stage the practitioner was successfully attested for in that year (Stage-None, Stage-1, Stage-2)
AHCADisciplinary	Number of AHCA Disciplinary actions invoked against the practitioner
JobCertifications	Number of specialty certifications received by the practitioner
CriminalOffenses	Number of criminal offenses charged against the practitioner
FacultyAppointments	Number of faculty appointments served by the practitioner
Awards	Number of awards received by the practitioner
Degrees	Number of professional and other degrees earned by the practitioner
InfoFailureProb	Probability of information-failure as a leading cause for the claim ($0 < p < 1$; values between 0 and 1)

Table 1: Description of Variables

The independent variable MU stage is a categorial variable that captures demonstrated level of HIT capability of a healthcare provider. The control variables include practitioner's personal and professional characteristics like awards, criminal offenses, and disciplinary actions. We drop any observations with missing data on the control variables. These variables of interest are described in the Table 1. We address multicollinearity issue in our model, by analyzing correlation coefficients between the variables. Quite expectedly, we found high positive correlation between Degrees and JobCertifications. As a result, we drop variable Degrees from analysis.

Econometric Models

Since the practitioners achieve successful MU-Stage attestations for different stages in different years, we perform a time-invariant analysis with the attested stages as treatments. Since the malpractice claims are rare events and the meaningful-use attestation for different stages overlaps for practitioners over the years 2011-2016, we use pooled logit model estimation for ordinal multinomial outcome (Cameron and Trivedi 2009).

Pooled Logistic Regression (PLR) model provides estimates of conditional odds ratios for having an event in an interval. While there can be a concern regarding multiple records within an individual contributing to several time intervals exhibiting dependence, researchers have shown that there is no inflation of test statistics in such case with lack of independence. This property holds intuitively because the likelihood factors into a distinct term for each interval. Also, researchers have informed that the odds in PLR describe the conditional probability of event occurrence, where the conditioning depends upon the individual characteristics until that particular time period. This allows all records within our dataset to be considered as conditionally independent. Conditional logistic regression model and pooled logistic regression models are equivalent when the length of time interval tends towards zero. Also, when the time interval is short, and the event is rare, the logistic regression estimates, and their standard errors approximate those from the proportional hazard mode.

We consider ordered multinomial logit extension to traditional PLR model to exploit the ordinal nature of our response variable with multinomial outcomes. An alternative consideration could be multinomial logit (MNL) model with nominal outcomes. Although an ordered estimator makes more assumptions than an MNL estimator, the ordered estimator is more efficient than the MNL estimator if these additional assumptions are true (Cameron and Trivedi 2009). The ordinal logit model makes additional assumption of proportional odds which means the distance between each category is equivalent. We perform Brant test to check if proportionality assumption holds (Long and Freese 2006). The Brant Test results in chi2 of 12.33 for 8 degrees of freedom with p-value of 0.137 and insignificant test statistics for variables included. This establishes the proportionality assumption holds true in our case.

We empirically model insurance-carrier's *perceived medical error (PME)* in a claim reported against the practitioner *i* in the year of observation *t* as a function of demonstrated HIT capability, and practitioner profile characteristics. Our base empirical specification is given below:

$$PME_{it}^{*} = \alpha_1 + \alpha_2 MUStage_{it} + \beta X_{it} + \varepsilon_{it} \tag{1}$$

In this equation, represents the latent multinomial outcome of perceived medical error in malpractice claim for practitioner *i* in time *t*, represents the treatment variable for demonstrated HIT capability of practitioner *i* at time *t*, and refers to profile characteristics of the practitioner *i* in time *t*. is the intercept of logistic regression while provides odds ratio for having the claim reported and settled against the practitioner depending upon the attested stage of meaningful-use by the practitioner *i* in time *t*.

We develop our second econometric model for to estimate the impact of HIT information-failures reported in malpractice claims. We use simple regression Ordinary Least Squares (OLS) model to estimate this impact. The econometric model is as described below:

$$IF_{it} = \gamma_1 + \gamma_2 MUStage_{it} + \omega_{it} \tag{2}$$

In this equation, IF_{ijt} presents the information-failure for claim *j* derived from the probabilistic topic approach described in methods section. The value of this variable is derived from probability distribution of topics. $MUStage_{it}$ represents the treatment variable demonstrated HIT capability by the practitioner *I* at *t*, the time of occurrence of the malpractice event.

Results & Discussion

The main results of our first econometric model are provided in Table 2. We discuss the impact of HIT capability demonstrated through MU-stage attestation, on the relative risk of perceived medical error in malpractice claims reported in our base model. In order to better interpret our results, we report odds ratios or relative risk ratios obtained from the ordered multinomial logit model.

The risk of a malpractice claim being reported and not settled against a practitioner is 62.32% lower when the practitioner attested for MU Stage-1 than when s/he did not attest at all. Interestingly, this risk is 89.12% lower when the practitioner attested for MU Stage-2.

Likewise, the risk of a claim being reported against a practitioner and being settled, indicating substantial evidence of perceived medical error, is 52.21% lower when the practitioner attested for MU Stage-1 than when s/he did not attested at all. This risk is 87.38% lower when the practitioner attested for MU Stage-2.

The ordered logit model enables us to compare the relative odds ratios obtained for MU Stage-1 and MU Stage-2. Irrespective of a malpractice claim being settled, we find the reduction in relative risk ratios is more for MU Stage-2 as compared to MU Stage-1.

After controlling for practitioner's professional and personal characteristics, the results corroborate significant reduction in relative risk of perceived medical error in the claims against practitioners attested to MU Stage-1 and MU Stage-2,. Hence, we find significant evidence in support of the first set of hypotheses H1(a),(b), and (c).

Explanatory Variables	**Relative Risk Ratios: Multinomial Logistic Regression Dependent Variable: PME (Base level = None)**	
	PME = Low (Claim reported and not-settled)	**PME = High (Claim reported and settled)**
MUStage: Stage-1	.3767395**(.1932627)	.4778289***(.1704413)
MUStage: Stage-2	.1087479***(.0589348)	.1261793***(.049468)
IncentiveProgram: Medicare	5.88198***(2.662634)	3.277465***(1.009892)
AHCADisciplinary	.7960792***(.0639731)	.6670127***(.0365491)
JobCertifications	.7993503***(.0179083)	.817801***(.0156949)
CriminalOffenses	.9351046(.0864796)	1.204173(.1749623)
FacultyAppointments	1.067448***(.0273666)	1.01159(.0208024)
Awards	.9781749***(.0102996)	.9950324(.0100482)
_const	.0187977***(.0136722)	.0179984***(.0167599)
Log pseudolikelihood	-3683.4691	
Pseudo-R2[1]	0.0648	
Observations	88695	
Note: Standard errors are in parentheses (). *, **, *** indicate significance at 5%, 1% and 0.1% confidence levels. [1]Pseudo-R2 is a measure for logistic regression analogous to OLS R2 measure calculates as ratio of log-likelihoods.		

Table 2: Base Model Estimation- Multinomial Logistic Regression

Other parameters of interest show relative risk ratios reduced with additional job certifications which can be explained by the practitioner's attitude towards self-improvement and responsible professional attitude driving them towards these certifications. Since the estimators for criminal offense charges is not significant, we cannot infer any effect at this time. A disciplinary action also reduced the relative risk ratios significantly, which can be partially explained by the fact that practice time is lost during a disciplinary action proceeding and that the practitioner becomes more professionally careful and aware as a result. The increase in relative risk ratios with more faculty appointments is surprising and needs further exploration. A possible explanation could be the lost practice time and experience while teaching make a practitioner prone to committing medical errors or negligence. Another explanation may be related to the nature of this position. Usually a teaching faculty is responsible for multiple residents and their clinical decisions. This increases the number of cases they oversee, which also increases the risk of errors and generates a new source of malpractice claims risk.

Dependent Variable: Info-Failure Probability	**Base Model: Linear Regression**
MUStage: Stage-1	-.9341579***(.231307)
MUStage: Stage-2	.6050062**(.2169066)
_const	-5.948458***(.8626493)
R-Squared	0.1039
Observations	88695
Note: Robust standard errors are in parentheses. *, **, *** indicate significance at 5%, 1% and 0.1% confidence level.	

Table 3: Linear Regression-OLS Estimation

Secondly, we estimate our linear regression model for testing a second set of hypotheses H2(a) and (b). The parameter estimates are significant and provide interesting insight into the impact of HIT in different MU-Stages. We find statistically significant support for H2(a), but not for H2(b). The estimated parameter for Stage-2 appears to increase the information failure probability in the malpractice claim. A possible reason could be information exchange incompatibility issues arising due to the additional requirement of information exchange capability in MU Stage-2. The results are presented in Table 3.

The results suggest that MU Stage-1 is effective in reducing the information-failure probability in medical malpractice claims but MU Stage-2 affects negatively. Although this is an unusual result, it contributes more to our understanding of how HIT impacts health outcomes.

Conclusion & Limitations

The implications of our study provide significant insights to researchers, medical practitioners, and policy makers. We adopt the machine learning method of topic modeling to the healthcare systems information management domain which is both data-intensive and information-deficient. Medical records and data is collected during patient encounters and clinical processes but the utility of all this data to create actionable information is missing. We extend methodology to extract information with machine learning algorithms in order to generate utility.

Medical practitioners are concerned about the risk associated with medical liability claims, but they resist the use of HIT in a meaningful manner due to perceived inconvenience of changing their clinical routines. We provide evidence of how HIT creates business value, including both clinical and administrative outcomes. Our analysis provides evidence of how HIT improves clinical outcomes for both patients and providers by reducing the associated risk of medical errors and negligence.

Although federal initiatives are designed to promote HIT assimilation and use, the programs may not be efficient in doing so. We find MU Stage-2 may not have generated intended outcomes. This may be because the requirement of interoperability was realized through policy before the capability to do so developed. Although MU Stage-1 reduces information-failures, these failures can get complicated with inefficient interoperability. We inform the policy makers to lookout for such caveats while designing a phased program in order to achieve desired results.

Just like any empirical work, there are certain limitations to our analysis. We study data on provider outcomes from state of Florida only. Generalizability to other states and the whole country may be debatable; we argue the sample size in our study is large enough to counter such issues. Future research may be able to extend our work to multiple state data set to address the issue of generalizability of our results. We have not addressed endogeneity concerns in our second econometric model and future research may be able to address this limitation as well.

REFERENCES

Akenroye, A. T., Baskin, M. N., Samnaliev, M., and Stack, A. M. 2014. "Impact of a Bronchiolitis Guideline on ED Resource Use and Cost: A Segmented Time-Series Analysis," *PEDIATRICS*. (https://doi.org/10.1542/peds.2013-1991).

Appari, A., Eric Johnson, M., and Anthony, D. L. 2013. "Meaningful Use of Electronic Health Record Systems and Process Quality of Care: Evidence from a Panel Data Analysis of U.S. Acute-Care Hospitals," *Health Services Research*. (https://doi.org/10.1111/j.1475-6773.2012.01448.x).

Arun, R., Suresh, V., Madhavan, C. E. V., and Murty, M. N. 2010. "On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. (https://doi.org/10.1007/978-3-642-13657-3_43).

Bao, Y., and Datta, A. 2014. "Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures," *Management Science*. (https://doi.org/10.1287/mnsc.2014.1930).

Barnett, M. L., Mehrotra, A., and Jena, A. B. 2016. "Adverse Inpatient Outcomes during the Transition to a New Electronic Health Record System: Observational Study.," *BMJ (Clinical Research Ed.)*. (https://doi.org/10.1136/bmj.i3835).

Black, A. D., Car, J., Pagliari, C., Anandan, C., Cresswell, K., Bokun, T., McKinstry, B., Procter, R., Majeed, A., and Sheikh, A. 2011. "The Impact of Ehealth on the Quality and Safety of Health Care: A Systematic Overview," *PLoS Medicine*. (https://doi.org/10.1371/journal.pmed.1000387).

Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. "Latent Dirichlet Allocation David," *Journal of Machine Learning Research2*. (https://doi.org/10.1162/jmlr.2003.3.4-5.993).

Boonstra, A., and Broekhuis, M. 2010. "Barriers to the Acceptance of Electronic Medical Records by Physicians from Systematic Review to Taxonomy and Interventions," *BMC Health Services Research*. (https://doi.org/10.1186/1472-6963-10-231).

Cameron, A. C., and Trivedi, P. K. 2009. "Microeconometrics Using Stata," *Stata Press Books*. (https://doi.org/10.1016/S0304-4076(00)00050-6).

Cao, J., Xia, T., Li, J., Zhang, Y., and Tang, S. 2009. "A Density-Based Method for Adaptive LDA Model Selection," *Neurocomputing*. (https://doi.org/10.1016/j.neucom.2008.06.011).

David, G., Gunnarsson, C. L., Waters, H. C., Horblyuk, R., and Kaplan, H. S. 2013. "Economic Measurement of Medical Errors Using a Hospital Claims Database," *Value in Health*. (https://doi.org/10.1016/j.jval.2012.11.010).

Devaraj, S., and Kohli, R. 2003. "Performance Impacts of Information Technology: Is Actual Usage the Missing Link?," *Management Science*. (https://doi.org/10.1287/mnsc.49.3.273.12736).

Deveaud, R., SanJuan, E., and Bellot, P. 2014. "Accurate and Effective Latent Concept Modeling for Ad Hoc Information Retrieval," *Document Numérique*. (https://doi.org/10.3166/dn.17.1.61-84).

Griffiths, T. L., Steyvers, M., Blei, B., and Blei, J. 2004. "Finding Scientific Topics A First Step in Identifying the Content of a Document Is Determining Which Topics That Document Addresses. We Describe a Generative Model for Documents, Introduced," *PNAS*. (https://doi.org/10.1073/pnas.0307752101).

Hersh, W. R., Weiner, M. G., Embi, P. J., Logan, J. R., Payne, P. R., Bernstam, E. V, Lehmann, H. P., Hripcsak, G., Hartzog, T. H., Cimino, J. J., and Saltz, J. H. 2014. "Caveats for the Use of Operational Electronic Health Record Data in Comparative Effectiveness Research," *Medical Care*. (https://doi.org/10.1097/MLR.0b013e31829b1dbd.Caveats).

Huang, A., Lehavy, R., Zang, A., and Zheng, R. 2014. "Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach," *SSRN*. (https://doi.org/10.2139/ssrn.2409482).

Long, J. S., and Freese, J. 2006. "Regression Models for Categorical Dependent Variables Using STATA," *Sociology The Journal Of The British Sociological Association*. (https://doi.org/10.12659/MSM.889196).

Mello, M. M., Chandra, A., Gawande, A. A., and Studdert, D. M. 2010. "National Costs of the Medical Liability System.," *Health Affairs (Project Hope)*. (https://doi.org/10.1377/hlthaff.2009.0807).

Menon, N. M., and Kohli, R. 2013. "Blunting Damocles' Sword: A Longitudinal Model of Healthcare It Impact on Malpractice Insurance Premium and Quality of Patient Care," *Information Systems Research*. (https://doi.org/10.1287/isre.2013.0484).

Minami, C. A., Sheils, C. R., Pavey, E., Chung, J. W., Stulberg, J. J., Odell, D. D., Yang, A. D., Bentrem, D. J., and Bilimoria, K. Y. 2017. "Association Between State Medical Malpractice Environment and Postoperative Outcomes in the United States," *Journal of the American College of Surgeons*. (https://doi.org/10.1016/j.jamcollsurg.2016.12.012).

Sahadi, J. 2018. "Warren Buffett Is Right. Health Care Costs Are Swallowing the Economy," *CNN Money*. (https://money.cnn.com/2018/01/30/news/economy/health-care-costs-eating-the-economy/index.html).

Steyvers, M., and Griffiths, T. L. 2006. "Probabilistic Topic Models," in *In Landauer, T., McNamara, D., Dennis, S., and Kintsch, W., Editors, Latent Semantic Analysis: A Road to Meaning*. *Laurence Erlbaum. Tang, Z. and MacLennan, J.* (https://doi.org/10.1109/TKDE.2009.122).

Vartak, S., Crandall, D. K., Brokel, J. M., Wakefield, D. S., and Ward, M. M. 2009. "Transformation of Emergency Department Processes of Care with EHR, CPOE, and ER Event Tracking Systems," *Health Information Management Journal*. (https://doi.org/10.1177/183335830903800204).

Wani, D., and Malhotra, M. 2018. "Does the Meaningful Use of Electronic Health Records Improve Patient Outcomes?," *Journal of Operations Management*. (https://doi.org/10.1016/j.jom.2018.06.003).