Electronic Trace Data and Legal Outcomes: The Effect of Electronic Medical Records on Malpractice Claim Resolution Time

Sam Ransbotham
Boston College, Chestnut Hill, MA 02467
sam.ransbotham@bc.edu

Eric M. Overby
Georgia Institute of Technology, Atlanta, GA 30332
eric.overby@scheller.gatech.edu

Michael C. Jernigan
Massachusetts General Hospital, Boston, MA 02114
Harvard Medical School, Boston, MA 02115
mjernigan@partners.org

Information systems generate detailed "trace" data about what individuals do and when they do it. Trace data may affect the resolution of lawsuits by, for example, changing the time needed for legal discovery. Trace data might speed resolution by clarifying what events happened when, or they might slow resolution by generating new and potentially irrelevant data that must be analyzed. To investigate this, we analyze the effect of electronic medical records (EMRs) on malpractice claim resolution time. Hospitals' use of basic EMRs at the time of the alleged malpractice is associated with a two month (8%) reduction in resolution time, while use of advanced EMRs is associated with a three month (13%) reduction. Because unresolved malpractice claims impose substantial costs on providers and patients, our finding that EMRs speed resolution has important welfare implications. Our study also contributes to the understanding of the effect of trace data on legal outcomes.

Key words: electronic trace data; lawsuit resolution; electronic medical records; medical malpractice claims; electronic discovery

History: Working paper

---

1. Introduction

Information systems generate and store large amounts of "trace" data: detailed data about what individuals and organizations do and when they do it. This includes not only data that individuals generate in the form of documents, database entries, and social media updates (i.e., the "what"), but also metadata about system use, including when individuals log in and out, where individuals are located, and what device they use for access (i.e., the "when, where, and how"). Indeed, "'data exhaust' — the trail of clicks that internet users leave behind from which value can be extracted — is becoming a mainstay of the internet economy" (Cukier 2010). Increasing digitization and the proliferation of trace data create opportunities for organizations but also present challenges (e.g., Bharadwaj et al. 2013). Opportunities include using trace data to build profiles to personalize marketing messages (e.g., Bleier and Eisenbeiss 2015), to improve matching between parties such as buyers and sellers or people on dating web sites (e.g., Hitsch et al. 2010), to determine credit-worthiness (e.g., Wei et al. 2016), etc. This can create better information, improve service, and/or reduce waste. However, trace data also present challenges, including inability to analyze the data properly due to a shortage of skills, escalating demands on IT infrastructure to store and manage the data, and relentless security risks associated with keeping the data safe (e.g., Hsu et al. 2015, August et al. 2014, Cukier 2010).

Another important but relatively under-explored implication of electronic trace data is their effect on the resolution of lawsuits by, for example, changing the time needed for legal discovery. Discovery is the process through which each party involved in a lawsuit obtains evidence from the other. Electronic trace data may provide detailed evidence about the chain of events leading up to an outcome. Sharing and interpreting this (and other) electronically-derived evidence among parties is referred to as electronic discovery. Concerns about electronic discovery affect organizational behavior; e.g., organizations can be reluctant to implement information systems due to the legal liability they may create, including increasing the likelihood of lawsuits and/or the costs associated with settling them (Miller and Tucker 2014). We extend this line of research by focusing on whether information systems and the trace data they generate increase or decrease the time to resolve lawsuits. This is not clear a priori and warrants empirical analysis. On one hand, trace data may speed resolution by providing a clear and (mostly) incontrovertible account of what events happened when. This can help all parties in the lawsuit move more quickly through the process and reach resolution. On the other hand, the sheer volume of trace data, including mountains of digital records that would not otherwise be available or searchable, may lengthen the process and thereby slow resolution.

We use the healthcare industry as the context in which to examine the tension between the potential clarity provided by trace data and the burden of analyzing it. Specifically, we analyze the relationship between hospitals' use of electronic medical records (EMRs) and the time it takes to resolve medical malpractice claims originating in those hospitals. Healthcare is an important context for IT in general (e.g., Dranove et al. 2014, Angst et al. 2010, Miller and Tucker 2009, Menon et al. 2000), partly due to increasing use of EMRs and their potential effects on healthcare quality and cost. We focus on the effect of EMRs on malpractice claim resolution time for three reasons. First, EMRs generate detailed trace data of the care administered (or not) to a patient, including when procedures were administered, who administered them, etc. As suggested above, these data could either speed or slow malpractice claim resolution, depending on whether they clarify or complicate the information needed for resolution. By examining this empirically, we extend our understanding of the effects of trace data on the resolution of lawsuits. Second, unresolved malpractice claims impose substantial economic and emotional costs on both providers and patients (Mello et al. 2010). Given this, policies or systems (such as EMRs) that affect claim resolution time have welfare implications. Indeed, multiple tort reform initiatives aim to expedite claim resolution (Kessler 2011). Given the limitations of tort reform, we take a fresh look at this issue by studying the potential of EMRs to expedite (or hinder) claim resolution. Last, hospitals are increasingly implementing EMRs, spurred by incentive payments and the promise of better patient care. However, it is critical for hospitals to understand how EMRs affect legal outcomes; indeed, malpractice "lawyers smell blood in electronic medical records" (Mearian 2015).

How does hospitals' use of EMRs affect the resolution time of malpractice claims? To study this, we combined data on hospital EMR use with data on malpractice claim resolution. We analyzed 13,503 resolved malpractice claims (both paid and unpaid) originating in 148 hospitals in Florida between 1999 and 2006. We collected data through June 2015, but we restricted analysis to claims originating no later than 2006 to allow sufficient time for them to resolve (as is necessary for any study of claim outcomes). We implemented a variety of models to estimate the relationship between hospitals' EMR use and claim resolution time, controlling for potential confounders such as claim severity, hospital characteristics (e.g., size, case mix, and location), and time trends. Hospitals' use of basic EMR functionality at the time of the alleged malpractice is associated with a more than two month (8%) reduction in claim resolution time, while use of advanced EMR functionality is associated with a more than three month (13%) reduction. These results hold after accounting for potential endogeneity in a hospital's decision to use EMRs. This suggests that the benefits of the trace data recorded by EMRs outweigh the costs in terms of speeding resolution, at least for this setting. Given the high emotional and economic costs of malpractice claims, faster resolution is likely to generate significant benefits for both patients and healthcare providers.

2. Literature Review and Motivation

Our analysis informs the growing body of research about the implications of electronic trace data for organizations and individuals. We focus on the implications of trace data for legal outcomes by studying how EMRs influence the resolution time of malpractice claims. We review these (increasingly narrow) research streams below.

2.1. Implications of Electronic Trace Data for Organizations and Individuals

Many organizations use electronic trace data to construct profiles about individuals that reflect their behaviors and interests. One use of these profiles is to personalize marketing messages. A key aspect of the business models for organizations such as Facebook and Google is to profile users based on their online behaviors (i.e., the "data exhaust" they generate as they use online services) and then charge marketers to show messages targeted to a given profile (Bleier and Eisenbeiss

2015, Lambrecht and Tucker 2013). Recommender systems provide another method to personalize marketing messages. Systems such as those used by Amazon and Netflix use profiles constructed based on individuals' online behaviors to recommend other products that they might like (e.g. Adomavicius and Tuzhilin 2011). A second use of profiles constructed from electronic trace data is to improve matching in markets. For example, online dating services use profiles to match users to one another (Hitsch et al. 2010). These profiles are based not only on users' stated preferences (e.g., what body type, hair color, etc. they prefer) but also on their revealed preferences reflected in how they use the service, including which other users' profiles they view and who they message (Gelles 2011). A third use of profiles is to assess trustworthiness (Hatton 2015). This includes using trace data from social networks (e.g., who users are "friends" with) to help underwrite loans and issue credit (Armour 2014). Using this type of trace data to assess creditworthiness has implications for whom individuals "friend" in online social networks, with the potential for friends who lower individuals' creditworthiness to be shunned (Wei et al. 2016). A fourth use of profiles generated from trace and other types of data is to engage in price discrimination (Phillips et al. 2015).

Electronic trace data have other uses beyond development of individual profiles, including being used as a signaling mechanism. For example, potential contributors to a crowdfunding campaign may interpret trace data showing who else has funded the campaign (and how much they contributed) as a signal of the campaign's quality. Given this, crowdfunding platforms have implemented tools to let contributors conceal their actions (Burtch et al. 2015). As another example, organizations such as Google (Page 2001) and Facebook (Lunt et al. 2010) incorporate trace data as a signal of content usefulness.

Another use of electronic trace data that is particularly relevant to our study is to improve organizational control and manage risk. For example, organizations can mine trace data for unusual behavioral patterns that might indicate theft or reduced productivity. One recent study showed a 22% decrease in theft by restaurant servers after an employee monitoring system was put in place (Pierce et al. 2015). This reflects the long-standing idea that the data generated by organizational information systems can identify process improvements and assess whether resources are being utilized optimally (e.g., Zuboff 1988). As we explore further below, it is likely that the trace data created by and stored within EMRs affects hospitals' ability to understand and audit the actions of their workers. This has implications for how hospitals manage the risk created by malpractice claims, with our focus being on how long it takes for these claims to be resolved.

Despite the opportunities created by electronic trace data for organizations, there are also challenges (Bharadwaj et al. 2013). One is developing and maintaining the necessary IT infrastructure and workforce to manage and analyze the data properly. Several industry reports note the supply/demand imbalance in workers with data analysis skills (e.g., Morris 2013). A related concern is keeping the data secure, particularly given that it can contain private and sensitive information about individuals' background, interests, financial situation, etc. (e.g, Hsu et al. 2015, Mitra and Ransbotham 2015). If trace data is valuable to organizations, then it is valuable to attackers as well.

Much of the trace data referenced above is generated by online behaviors, such as what links users click and who they interact via e-mail and on social networks. However, some trace data reflect offline behaviors that are recorded in information systems. This includes what dishes and drinks were ordered (and when) in the restaurant theft study noted above. It also includes what care was administered (or not) to healthcare patients, which we discuss further below.

2.2. Electronic Trace Data, Electronic Discovery, and Lawsuit Resolution

Electronic trace data also have implications for lawsuit resolution by, for example, affecting the time needed for discovery. Discovery is a process through which each party involved in a lawsuit obtains evidence from the other(s). This can involve depositions and requests for documents related to the lawsuit. In many jurisdictions, courts compel parties involved in the lawsuit, as well as others who are not directly involved but may have relevant information, to provide information. The speed of the discovery process has a large effect on how quickly a lawsuit is resolved, as "the discovery phase is easily the most time consuming portion of most cases..." (HG.org Legal Resources 2016).

Electronic discovery refers to the aspect of the discovery process in which evidence is generated from electronic data. This includes electronically-stored documents, database entries, and other types of electronic trace data. For example, defense attorneys in a New York City lawsuit used trace data from a murder suspect's public transit card (which showed where and when the suspect had used subways and buses) to generate evidence that the suspect was too far away from the crime scene to have been guilty (Chan 2008). As information systems and sensors (e.g., the Internet of Things) have become more pervasive, electronic discovery has become an increasingly large component of the overall discovery process.

Whether electronic data increase or decrease the time to resolve a lawsuit (e.g., through changes in the discovery process) is unclear, as illustrated in Table 1. On one hand, electronic data may slow discovery and resolution time. First, information systems often generate massive amounts of data. If too much of these data are irrelevant to the lawsuit, then they may obscure the relevant information. Combing through these data requires time and resources that may introduce delays (Beisner 2010). To use an analogy, electronic data enlarge the size of the "haystack" (i.e., the overall amount of data), making it harder to find the "needles" (i.e., the relevant data). Second, relevant data may be spread across multiple information systems. Cross-checking and/or integrating data across multiple systems creates an analytical burden that may create delay. Third, substantial (and potentially idiosyncratic) technical and data expertise can be required to process information from each system (Luoma 2006). This expertise may be difficult to find (Ransbotham et al. 2015), may be occupied elsewhere, and/or may "seriously disrupt IS and over-burden [IS] staff" (Volonino 2003, p. 3), resulting in delay. Fourth, electronic storage and dissemination of evidence create data security and privacy considerations (Kim 2006). Handling these may add time to the process. Overall, electronic discovery may absorb substantial time and money (Pace and Zakaras 2012), and many firms may be largely unprepared for the implications of electronic discovery (Swartz 2007).

On the other hand, electronic data may speed discovery and resolution time. First, electronic data, including trace data, may provide a detailed account of what happened when. This may allow

Table 1 Conflicting Effects of Electronic Data on Lawsuit Resolution

Concept	Description	Possible Effect
Creation	Information systems create large amounts of data that may either illuminate or obfuscate	May speed or slow resolution, depending on data relevance
Data distribution / redundancy	Multiple systems may contain relevant information	May speed or slow resolution, depending on data integration / consistency
Analysis	Analysis of electronic data may require scarce or overburdened technical expertise	May slow resolution
Data security and privacy	Electronic data require additional measures to ensure security and privacy	May slow resolution
Codification	Electronic data are often already codified as part of the primary use of the information system	May speed resolution
Communication and workflow	Electronic data can be stored, copied, and communicated quickly to improve workflow efficiency	May speed resolution
Processing	Information review systems can automate processing of electronic data and/or electronic data can be distributed for parallel processing by humans	May speed resolution

the most relevant evidence to be collected quickly and reliably, thereby speeding resolution. Second, electronic data often require less codification than do non-electronic data (because the former are often coded when input into the information system), which reduces subsequent processing costs and can speed interpretation (Cohendet and Meyer-Krahmer 2001). Third, data can be stored, copied, and transferred more quickly and cost-effectively when electronic. This increase in communication efficiency has long been touted as a benefit of information technology (Mendelson and Pillai 1998). Fourth, IT tools such as document review systems facilitate searching through electronic data, often automating much of the process. For aspects of the analysis that cannot be automated effectively, electronic records can be distributed for parallel human processing if desired.

2.3. Electronic Medical Records and Malpractice Claim Resolution Time

We examine the effect of electronic data, including trace data, on lawsuit resolution in the healthcare industry. The effect of electronic data — and information technology in general — on healthcare is an important and increasingly well-studied phenomenon. Information technology (including EMRs) can improve the care given to patients by increasing healthcare worker productivity, standardizing care practices, helping to ensure that these practices are followed, and reducing medical errors (Balas 2001, Devaraj and Kohli 2000, Dexter et al. 2001, Hersh 2004, Koppel et al. 2005, Menon and Kohli 2013, Menon et al. 2000, Kuperman et al. 2007, Gray and Goldmann 2004). However, challenges associated with successful use of IT in healthcare have slowed adoption (Jha et al. 2009). For example, the sensitive personal nature of the information within healthcare information systems exacerbates privacy concerns (Miller and Tucker 2009, Angst and Agarwal 2009), many providers complain that EMRs are cumbersome to use (McNickle 2011), and the productivity benefits of EMRs are unclear (Bhargava and Mishra 2014). Healthcare IT can also facilitate practices such as aggressive billing that may increase healthcare costs that are already considered too high (Abelson et al. 2012, Curfman et al. 2013, Sidorov 2006, Soumerai and Koppel 2012). For reviews of the effects of IT on healthcare quality, cost, and related outcomes, see Agarwal et al. (2010) and Kellermann and Jones (2013).

We contribute to this stream of research — and also explore the implications of electronic trace data for lawsuit resolution — by examining the effect of EMRs on how long it takes to resolve medical malpractice claims. EMRs may affect malpractice claims in several ways, including not only how long it takes to resolve a claim but also whether a claim is filed at all and whether and to what degree the provider was liable. Regarding the likelihood of a claim being filed and the provider's subsequent liability, if EMRs improve quality of care, then they lower the likelihood that malpractice occurs and a claim is filed (Studdert et al. 2006). However, EMRs may increase the discoverability of the care given to (or withheld from) patients by creating trace data that serves as an electronic paper trail, including the exact timing and sequence of care procedures (Mangalmurti et al. 2010). Although this can be helpful in defending malpractice claims if care was properly administered (Miller and Glusko 2003), it can serve as a "smoking gun" if it was not (Korin and Quattrone 2007). This could lead to increased claim likelihood and liability for providers. Indeed, there is evidence that hospitals may be deferring IT investments due to the legal liability associated with electronic discovery of these trace data (e.g., Miller and Tucker 2014). This has not gone unnoticed by plaintiffs' attorneys, who "smell blood in electronic medical records"

(Mearian 2015). Given this concern, a small number of studies have examined the link between EMR use and the likelihood of a malpractice claim being filed. Two companion studies showed a negative correlation between EMR use and malpractice claims (Quinn et al. 2012, Virapongse et al. 2008), while another showed no correlation (Victoroff et al. 2012).

Although malpractice outcomes related to claim likelihood and liability are important, we focus our analysis on the effect of EMRs on claim resolution time. We do so for three reasons.

First, focusing on resolution time allows us to examine the tension between the potential clarity provided by electronic trace data and the burden of analyzing it. Malpractice claims take a long time to resolve for several reasons, including a long discovery process in which the parties must determine whether the provider adhered to the appropriate standard of care as well as who was involved in the care that led to the injury (Seabury et al. 2013). As noted above, the data created by and stored within EMRs might speed claim resolution because the data represent a legible, electronic paper trail of the timing and sequence of care administered (or not) to a patient, along with which health care professional administered the care (assuming the integrity of these data is protected) (Hoffman and Podgurski 2009, Hoffman 2010). This might help the parties involved in the claim determine its merits more quickly, leading to faster resolution. However, the sheer volume of data stored within EMRs, potential conflicts with other data sources, and difficulty extracting and analyzing the relevant data might slow discovery and claim resolution (e.g., Degnan 2011).

Second, medical malpractice claims create substantial monetary and emotional costs for providers and patients (Studdert et al. 2004, Mello et al. 2010). An important and often overlooked factor contributing to these costs is the length of time required to resolve claims. A recent study estimated that the average physician spends almost 11% of his/her career with an open, unresolved claim pending against him/her (Seabury et al. 2013). The long resolution time negatively affects health-care providers by increasing stress, distracting them from the practice of medicine, and delaying their ability to implement changes to prevent future medical errors (Studdert et al. 2004, Seabury et al. 2013, Sage 2004). A 2015 survey showed that nearly half of physicians sued for malpractice had claims last 3 years or longer, with one physician complaining about "years of agonizing about the potential for a catastrophic outcome, loss of license, practice, etc." (Peckham 2015). The same survey showed that over 95% of physicians sued for malpractice found the experience "unpleasant", "upsetting", "very bad", or "horrible". Long resolution times also increase the uncertainty faced by malpractice insurers about their risk exposure, leading to fluctuations in insurance premiums that further disadvantage providers (Government Accountability Office 2003). Long resolution times negatively affect patients through increased anxiety, lack of closure with the provider, and delays in receiving appropriate compensation (Hobgood et al. 2005, Gallagher and Levinson 2005). Although many analysts agree on the negative consequences of the malpractice system, most focus on tort reform as the method to ameliorate these issues (Thorpe 2004). However, because there is significant pessimism about the potential for tort reform to provide relief (Mello et al. 2003), it is important to identify other methods. Accordingly, we study the potential of EMRs to expedite (or hinder) claim resolution. EMRs may help improve the malpractice system because "excellent documentation is the backbone of defensive medicine" (Gart 2008), and one of the key pieces of advice from physicians who have been sued is to "document, document, document" (Peckham 2015).

Third, hospital adoption of EMRs is increasing steadily, and much research has been devoted to their effects on quality of patient care, healthcare costs, physician and patient satisfaction, etc. However, relatively little research has focused on the effect of EMRs on healthcare litigation (see above), and none has focused on the effect on malpractice claim resolution time. Given the potentially large welfare implications associated with speeding up malpractice claim resolution, understanding whether EMRs decrease or increase claim resolution time is an important policy issue.

3. Data and Empirical Approach

Because malpractice claims take years to resolve, assessing the effects of EMRs on claim resolution time has only recently become feasible, which may explain the lack of empirical research on this topic. In this study, we analyzed longitudinal data on the use of EMRs (basic and advanced) and malpractice claims originating in hospitals in Florida. We used a difference-in-differences model (using both the full sample and the sub-sample of hospitals that implemented EMRs during the sample period) to estimate the effect of EMRs on claim resolution time, controlling for claim characteristics such as injury severity; hospital characteristics such as size, case mix, and location; and a time trend.

3.1. Data Sources

Despite the negative consequences of long claim resolution times, little empirical research investigates the factors that contribute to that length, including the use of healthcare IT such as EMRs (Seabury et al. 2013). One reason is the absence of a consolidated, multi-provider data repository suitable for analyzing relationships (Institute of Medicine 2011). To overcome this and to advance research on this important issue, we constructed a new data set suitable for examining the relationship between EMRs and malpractice claim resolution time by consolidating data from the sources described below. We focus on the state of Florida due to the availability of detailed claim data. Claims are only reported when they are resolved; the Florida reporting regulations do not require reporting of unresolved claims. Therefore, we collected data through June 2015 but we restricted the analysis to claims resulting from injuries occurring from 1999 to 2006 to allow sufficient time for claims to be resolved and reported, as is necessary in any study of claim outcomes (Quinn et al. 2012, Virapongse et al. 2008). Matching data across the sources yields a sample of 13,503 resolved claims (both paid and unpaid) originating in 148 hospitals.

3.1.1. EMRs: The Healthcare Information and Management Systems Society (HIMSS) conducts annual surveys of hospital chief information officers and information systems managers about the use of information technology such as EMRs (Healthcare Information and Management Systems Society 2013). These established measures of EMR use predate the emerging Meaningful Use standard (Angst et al. 2011, 2010, Dranove et al. 2014). As shown in Table 2, HIMSS considers EMRs to include multiple component systems and considers each hospital to be in one of

Table 2 HIMSS U.S. EMR Adoption Model

Stage	System capabilities (cumulative)
Stage 7	Complete EMR; Continuity of Care Document transactions to share data; Data warehousing; Data continuity with Emergency Department, ambulatory, outpatient
Stage 6	Physician documentation (structured templates), full Clinical Decision Support System (variance & compliance), full Radiology Picture Archiving and Communication System
Stage 5	Closed loop medication administration
Stage 4	Computerized Practitioner Order Entry, Clinical Decision Support (clinical protocols)
Stage 3	Nursing/clinical documentation (flow sheets), Clinical Decision Support System (error checking), Picture Archiving and Communication System available outside Radiology
Stage 2	Clinical Data Repository, Controlled Medical Vocabulary, Clinical Decision Support, may have Document Imaging; Health Information Exchange capable
Stage 1	Ancillaries (laboratory, radiology, pharmacy) all installed
Stage 0	All three stage 1 ancillaries (laboratory, radiology, pharmacy) not installed

Source: Healthcare Information and Management Systems Society (2015)

seven stages of EMR use each year based on the component systems at the hospital (Healthcare Information and Management Systems Society 2015).

Due to the limited number of hospitals in each of the seven stages and to be consistent with other research (e.g., Dranove et al. 2014), we dichotomize hospitals as having either basic (stage 1 through stage 3) or advanced (stage 4 through stage 7) EMR functionality. Basic EMR functionality includes a clinical data repository fed by ancillary clinical systems, flowsheets, and basic clinical decision support. Advanced EMR functionality includes computerized order entry, advanced decision support such as variance/compliance tracking, and physician documentation systems. Indicator variables designate whether each hospital had basic or advanced EMR functionality in each year. During our sample period, 43 of the 148 hospitals always had EMRs, 55 never had EMRs, and 50 adopted EMRs. Of the 93 hospitals with EMRs during the period, 1 always had advanced EMRs and 17 migrated from basic to advanced EMRs.

3.1.2. Hospitals: The state of Florida requires healthcare facilities within the state to report information annually to the Florida Agency for Healthcare Information. We used the reports published by this agency to collect annual data on hospital size, patient population, operations, and finances. We collected data on hospital adherence to accepted care processes from the Hospital Compare databases compiled by the U.S. Department of Health and Human Services (DHHS Centers for Medicare and Medicaid Services 2013).

3.1.3. Malpractice claims: The state of Florida requires licensed medical malpractice insurance providers to report on resolved malpractice claims, pursuant to reporting statute Chapter 627.912, F.S. These data are public record and are available through the Florida Office of Insurance Regulation (`https://apps.fldfs.com/PLCR/Search/MPLClaim.aspx`). We downloaded data for the claims originating in the hospitals in our sample through the end of June 2015, including the date the alleged malpractice occurred, the date the claim was filed, the severity of the injury, the date the claim was resolved, and whether the claim went to court. Because claims are not reported until resolved, we limited our analysis to claims filed between 1999 and 2006. This allows at least 8.5 years for claim resolution, which is a longer (i.e., more conservative) window than that typically used in studies of claim outcomes (Quinn et al. 2012, Virapongse et al. 2008). We also conduct robustness checks to mitigate the possibility that truncation of exceptionally slow-to-resolve claims biases our results.

3.2. Variables

3.2.1. Claim resolution time: The dependent variable is the length of time required to resolve the claim (*Resolution Time*), measured as the number of days from the date the claim was filed until it was resolved. We also consider an alternative measure of the number of days from the date the injury occurred until the claim was resolved. Figure 1 depicts the distribution of resolution times of the claims in the sample.

3.2.2. Claim characteristics: Aspects of the claim itself may affect the resolution time. *Court* indicates whether the claim went to court or was settled beforehand. *Severity* describes the extent of the patient's injury. In escalating order of severity, these are Slight, Minor, Significant, Major, Grave, and Death. Table 3 provides descriptive statistics for the claims in the sample; Table 5 provides correlations.

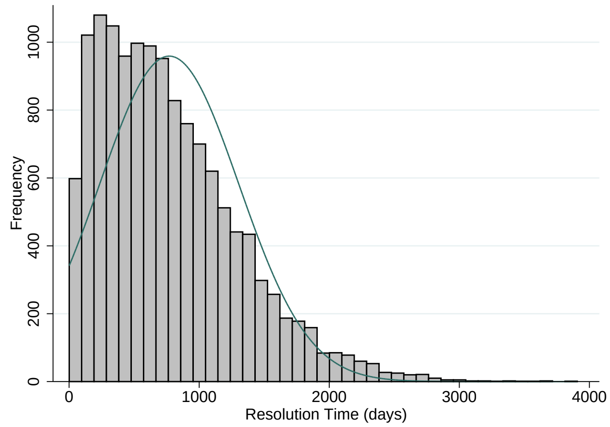

Figure 1 **Histogram of Resolution Time of Malpractice Claims (1999–2006)**

Table 3 **Descriptive Statistics for Medical Malpractice Claims**

Variable	Units	Min	Max	Mean	Median	Std. Dev.
Resolution Time	(days)	1	3907	769.813	674	630.526
Court	(1 if went to court)	0	1	0.159	0	0.366
Severity	(1 if Slight)	0	1	0.070	0	0.255
	(1 if Minor)	0	1	0.306	0	0.461
	(1 if Significant)	0	1	0.082	0	0.275
	(1 if Major)	0	1	0.171	0	0.377
	(1 if Grave)	0	1	0.037	0	0.188
	(1 if Death)	0	1	0.334	0	0.472

The unit of analysis is the malpractice claim; 13,503 observations.

3.2.3. Hospital characteristics: We control for characteristics of the hospitals that also may affect resolution time. We control for size using *Employees*, measured as the number of full-time equivalent (FTE) employees at the hospital in each year. *Employees* is highly and positively correlated with other measures of hospital size, such as number of physicians, number of beds, number of staffed beds, number of visits, etc. Due to the collinearity among these variables, we include only *Employees* in the models; however, the results are robust to alternative measures of size. *Working Hours per Patient* controls for (the inverse of) the utilization level of the hospital staff. We measure this as the total number of working hours available from the hospital employees divided by the patient load (measured in adjusted patient days) in each year. *Salary per FTE* is the average compensation per full-time equivalent employee at the hospital in each year; this variable partially controls for the mix of staff (e.g., physicians, technicians) at a hospital. *Net Revenue*

Table 4 Descriptive Statistics for Hospitals

Variable	Units	Min	Max	Mean	Median	Std. Dev.
Basic EMR	(1 if in use)	0	1	0.437	0.000	0.496
Advanced EMR	(1 if in use)	0	1	0.056	0.000	0.229
Employees	(full-time equivalents)	5.926	12170.900	1191.596	746.750	1439.437
Working Hours per Patient	(hours / patient days)	11.060	65.100	24.989	23.635	6.036
Salary per FTE	(US$K)	18.784	63.185	41.870	41.667	6.583
Net Revenue per Patient	(US$K / patient days)	0.024	3.042	1.338	1.296	0.371
Occupancy	(percentage)	0.010	1.060	0.560	0.570	0.178
Geographic Index	(index)	13.840	109.370	99.790	100.185	5.225
Case Mix Index	(index)	0.150	2.290	1.267	1.230	0.234
Use of Accepted Practices	(index)	0.330	0.970	0.774	0.780	0.075

The unit of analysis is hospital / year, e.g., hospital A in 1999, hospital A in 2000, hospital B in 1999, hospital B in 2000, etc. All dollar amounts adjusted to January 2000 equivalents using the Consumer Price Index as reported by the U.S. Bureau of Labor Statistics. 1,310 observations.

per Patient is the net operating revenue of the hospital divided by the patient load (measured in adjusted patient days) in each year. This variable controls for aspects of the hospital's financial condition. *Occupancy* is the average percentage of beds that were in use throughout the year and helps control for utilization level of the hospital. *Geographic Index* measures the variation in cost and services attributable to market conditions in a region. An index below 100 indicates that a hospital's charges for services are lower than the state average in a given year. This variable helps control for aspects of the hospital's cost structure. *Case Mix Index* is a diagnosis-weighted average of the patients that the hospital treats in each year and controls for the underlying population served by the hospital. *Use of Accepted Practices* is the average score of the items in the Hospital Compare database that relate to medical procedures administered in the hospital; this controls for the degree to which the hospital adheres to accepted care practices in each year. These data are only available for the last years of our sample (2004–2006). We use the 2004 data for observations prior to 2004, although results are consistent if we drop this variable from the analysis. All dollar amounts are adjusted to January 2000 equivalents using the Consumer Price Index as reported by the U.S. Bureau of Labor Statistics. Table 4 summarizes these variables for the hospitals in the sample; Table 5 provides correlations.

3.2.4. EMRs: Our focal variable is whether the hospital used EMRs at the time the injury occurred. *Basic EMR* indicates whether a hospital used systems that provide basic EMR functionality at the time of the injury. *Advanced EMR* indicates whether a hospital used systems that

Table 5 Correlations

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
1	Resolution Time	1.00											
2	Year	−0.34	1.00										
3	Severity	0.16	−0.03	1.00									
4	Court	0.23	−0.12	0.05	1.00								
5	Employees (ln)	0.01	0.04	0.01	0.03	1.00							
6	Working Hours per Patient	0.05	−0.10	0.01	0.04	0.64	1.00						
7	Salary per FTE	−0.25	0.74	−0.00	−0.09	0.13	−0.16	1.00					
8	Net Revenue per Patient	−0.17	0.44	−0.04	−0.06	0.31	0.26	0.42	1.00				
9	Occupancy	−0.03	0.13	0.01	−0.00	0.31	−0.04	0.17	0.26	1.00			
10	Geographic Index	0.06	−0.12	0.03	0.05	0.19	−0.08	0.25	−0.01	0.06	1.00		
11	Case Mix Index	0.00	−0.07	−0.02	−0.01	0.44	0.34	0.08	0.54	0.16	0.05	1.00	
12	Use of Accepted Practices	0.07	−0.19	0.02	0.02	0.02	0.11	−0.12	0.04	−0.11	0.18	0.10	1.00
13	EMR	−0.13	0.32	−0.01	−0.04	−0.04	−0.08	0.25	0.23	0.13	−0.15	−0.01	−0.11

Pearson product moment correlation using 13,503 observations. To help assess collinearity, ordinal integers replace ordered indicators. EMR is 0 for none, 1 for basic and 2 for advanced. The Severity indicator variables are represented as an integer ordered by increasing severity.

provide advanced EMR functionality at the time of the injury. These variables change across years for hospitals as they adopt EMRs. Section 3.1.1 describes how we coded basic versus advanced EMR functionality.

3.3. Empirical Approach

We investigate whether use of basic and advanced EMRs within a hospital influenced the resolution time for malpractice claims originating in that hospital. We conducted the analysis for claims originating in all hospitals in the data, which we refer to as the "all hospitals" analysis ($n = 13,503$), as well as for claims originating in only those hospitals that adopted EMRs during the sample period, which we refer to as the "adopting hospitals" analysis ($n = 5,479$). The "adopting hospitals" sub-sample excludes hospitals that never had EMRs during the sample period as well as those that always had EMRs.

We begin with an aggregate estimation of the effect of EMR adoption through a comparison of means and medians. Next, we use multiple regression to control for factors other than EMR use that prior research indicates might influence claim resolution time. We use a difference-in-differences approach where we take advantage of variation in which hospitals adopt as well as when they adopt. We include year indicator variables (i.e., fixed effects) to account for macro-level changes in the healthcare and legal environments that might influence resolution times. For example, the state of Florida enacted statute 766.118 in 2003 to limit non-economic damages that could be awarded to a patient; the indicator variables for 2003, 2004, 2005, and 2006 capture the effect that the passage of this law might have on claim resolution times. Year indicator variables are a critical part of our identification strategy, as they account for many unobserved factors that manifest over time. We used ordinary least squares regression because the dependent variable, *Resolution Time*, although an integer, has a mean much greater than zero and an approximately normal distribution (Figure 1). Later robustness tests consider alternative approaches.

Our focal model is

$$t_{jky} = \beta_0 + \beta_b Basic_{jy} + \beta_a Advanced_{jy} + \beta_{h1} Hospital_j + \beta_{h2} Hospital_{jy} + \beta_c Claim_k + \beta_y Year_y + \epsilon_{jky} \tag{1}$$

where t_{jky} is the *Resolution Time* for claim k originating in hospital j in year y; $Basic_{jy}$ ($Advanced_{jy}$) indicates if basic (advanced) EMRs were in use at hospital j in year y; $Hospital_j$ are hospital fixed effects that control for all time-invariant hospital characteristics; $Hospital_{jy}$ is a vector of covariates for hospital j that vary over years y; $Claim_k$ is a vector of covariates for claim k; $Year_y$ are yearly indicator variables (i.e., fixed effects); ϵ_{jky} is an error term clustered by hospital so that we avoid understating the standard errors (Bertrand et al. 2004); and β's are estimated intercepts and coefficients.

Hospitals do not randomly use EMRs; use is a strategic choice made by each hospital. If we ignore the potential endogeneity this creates, then our estimate of the relationship between EMRs and claim resolution time may reflect the possibility that hospitals that choose to use EMRs have characteristics (such as size or case mix) that inherently lead to shorter (or longer) claim resolution time, rather than capturing a more direct relationship. We mitigate this potential bias in five ways.

First, we explicitly controlled for hospital characteristics that might influence a hospital's choice to use EMRs and claim resolution time. We did this by including in our regressions a rich set of time-variant hospital covariates from the data sources described above as well as hospital fixed effects that control for time-invariant hospital characteristics, such as profit/non-profit and location. Second, it is possible that hospitals that adopted EMRs during the sample period are systematically different from those that did not in ways that lead to faster or slower claim resolution. In light of this, we replicated our analyses using only the "adopting hospitals" sub-sample. As shown below, results are similar, thereby mitigating this concern. Third, we measure the differential effects of both basic and advanced EMRs, and we find a stronger effect for advanced EMRs. This nuanced relationship between level of EMR functionality and claim resolution time is consistent with a causal relationship and harder to account for with alternative explanations. (The year fixed effects in our regressions separate the differential effects of basic and advanced EMRs from a general time trend, which is important because advanced EMR functionality became more prevalent over time.) Fourth, we used instrumental variables (IV) regression to account for possible endogeneity in whether hospitals use EMRs. As shown below, tests suggest that the instruments are valid, and the IV results are consistent with the main results. Fifth, we used propensity scoring methods to match claims originating in hospitals with basic EMRs to those in hospitals without EMRs. These results are consistent with the main results, and we also use them to conduct a Rosenbaum sensitivity test (Rosenbaum 2002, 2005) to assess how large an effect any unobserved confounding variables would need to have to overturn our conclusions.

We conducted several other robustness checks. These include implementing a Cox proportional hazards model to analyze factors that affect resolution time; testing a multi-level model that includes random intercepts for each hospital; using a negative binomial count model; mitigating the possibility that the 1999–2006 study period truncates slow-to-resolve claims by analyzing only those claims resolved within 5 or 6 years and (separately) by re-running the analysis for claims filed between 1999 and 2004 or between 1999 and 2005; and mitigating the potential effect of outliers by removing claims originating in hospitals with the largest claim volumes.

4. Results

4.1. Preliminary, Model-Free Evidence

First, we explored the effect of EMRs on claim resolution time through a basic comparison of means and medians (Table 6). Mean comparisons are based on t-tests and median comparisons use continuity-corrected Pearson's χ^2. On average, claims originating in hospitals that did not use EMRs in the year of the alleged malpractice took 831 days to resolve (median = 740). In hospitals that had basic EMR functionality in the year of the alleged malpractice, the mean days to resolution was 704 (median = 600). This represents a 15%–19% speed improvement and is a statistically significant difference ($p < 0.001$). Resolution time was faster still in hospitals with advanced EMR functionality in the year of the alleged malpractice: mean = 614 and median = 507, which is 26%–31% faster than "no EMR" ($p < 0.001$) and 13%-16% faster than "basic EMR" ($p < 0.001$). We replicated this analysis using only the "adopting hospitals" sub-sample (lower panel of Table 6). Results are consistent with those obtained from the full sample, with the estimated effect sizes somewhat larger. This suggests that hospitals' claim resolution times decrease after they adopt EMRs, with greater levels of EMR functionality leading to greater decreases.

4.2. Focal Analysis

Table 7 describes the results of the multiple regression analysis (equation 1) using the "all hospitals" sample and the "adopting hospitals" sub-sample. In all models, EMRs have a negative and significant effect on claim resolution time. For the "all hospitals" analysis, basic EMRs reduce resolution time by 65 days ($\beta = -65.43$, $p < 0.01$), which represents an 8% reduction. Advanced EMRs reduce claim resolution time by 110 days on average (i.e., $\beta = -109.92$, $p < 0.001$), which represents a 13% reduction. The advanced EMR effect is statistically different from the basic EMR effect ($p < 0.05$). Results from the "adopting hospitals" analysis are similar: $\beta = -87.13$ ($p < 0.01$) for basic EMRs and $\beta = -133.44$ ($p < 0.001$) for advanced EMRs. These results are consistent with the model-free results, although the effect sizes are smaller.

A plausible explanation for these results is that EMRs produce an electronic record of what care was administered (or withheld) from the patient, when, and by whom. This leads to faster

Table 6 EMRs and Malpractice Claim Resolution Time – Comparison of Means and Medians

All Hospitals

Variables	n	Mean	Std. Dev.	Median	
No EMR	7420	831.40	552.36	740	
Basic EMR	5466	703.81	509.40	600	
Advanced EMR	617	613.91	436.82	507	
Differences		Δ Mean	t	Δ Median	χ^2
Basic EMR minus No EMR		-127.58^{***}	13.39	-140.00^{***}	162.32
Advanced EMR minus Basic EMR		-89.91^{***}	4.21	-93.00^{***}	14.51
Advanced EMR minus No EMR		-217.49^{***}	9.53	-233.00^{***}	76.91

Adopting Hospitals

Variables	n	Mean	Std. Dev.	Median	
No EMR	3917	869.80	557.83	801	
Basic EMR	1118	617.03	432.72	520	
Advanced EMR	444	564.96	391.57	480	
Differences		Δ Mean	t	Δ Median	χ^2
Basic EMR minus No EMR		-252.76^{***}	14.00	-281.00^{***}	160.73
Advanced EMR minus Basic EMR		-52.07^{*}	2.20	-40.00^{*}	2.10
Advanced EMR minus No EMR		-304.83^{***}	11.21	-321.00^{***}	100.80

Comparison of claim resolution time by degree of EMR use: none, basic, or advanced. Mean comparisons are based on t-tests; median comparisons use continuity-corrected Pearson's χ^2 tests. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

claim resolution by helping the hospital and the claimant gather and analyze data relevant to the claim more quickly than if the data were stored non-electronically (or not at all). Advanced EMRs produce a more complete electronic record than do basic EMRs, thereby leading to even faster resolution. This differential effect of advanced EMRs supports a causal link between EMRs and claim resolution time. If our results were simply due to hospitals that use EMRs having inherent (and unobserved) characteristics that lead to faster claim resolution, then the effects of basic and advanced EMRs would be less likely to differ.

To further account for bias due to potential endogeneity in hospitals' decision to use EMRs, and to provide additional evidence for a causal link between EMRs and claim resolution time, we used an instrumental variable approach. To simplify instrumentation, we combined the basic and advanced EMR indicators into a single indicator, labeled *Any EMR*. We used two instruments for *Any EMR*: *Document Management* and *Medical Terms*. *Document Management* is an indicator variable for whether hospitals used administrative document management systems to automate workflow

Table 7 EMRs and Malpractice Claim Resolution Time

Variables	Control Model 0	All Hospitals Model 1	Adopting Hospitals Model 2
Hospital fixed effects	yes	yes	yes
Year indicators	yes	yes	yes
Constant	29.297	75.837	1,239.687
	(425.324)	(425.229)	(781.771)
Severity: Minor	132.283***	131.781***	134.659***
	(17.448)	(17.438)	(26.734)
Severity: Significant	239.938***	238.962***	244.239***
	(21.485)	(21.474)	(33.734)
Severity: Major	189.666***	189.467***	239.773***
	(18.764)	(18.755)	(29.269)
Severity: Grave	213.919***	214.730***	257.829***
	(26.864)	(26.855)	(41.938)
Severity: Death	280.918***	280.271***	285.014***
	(17.379)	(17.371)	(26.581)
Court	267.219***	267.060***	237.496***
	(11.576)	(11.570)	(17.768)
Employees (ln)	40.070	50.742	−96.676
	(41.427)	(41.507)	(72.587)
Working Hours per Patient (std)	−64.534***	−59.019**	−46.889
	(19.041)	(19.119)	(28.162)
Salary per FTE (std)	19.476	19.532	34.031
	(19.478)	(19.471)	(28.285)
Net Revenue per Patient (std)	8.149	15.754	22.623
	(16.296)	(16.395)	(22.433)
Occupancy (std)	4.256	11.920	21.508
	(20.283)	(20.352)	(31.849)
Geographic Index	5.730*	5.517*	5.769
	(2.791)	(2.791)	(5.093)
Case Mix Index	−91.803	−147.352*	−229.467*
	(72.963)	(74.105)	(95.918)
Use of Accepted Practices (std)	−3.501	−3.915	29.243
	(12.356)	(12.350)	(21.199)
Basic EMR		−65.426**	−87.133**
		(21.617)	(28.007)
Advanced EMR		−109.923***	−133.436***
		(29.461)	(36.525)
Observations	13,503	13,503	5,479
R^2	0.175	0.176	0.181
F	134.5	123.8	52.07

OLS regression on claim resolution time (measured in number of days since claim filed); Standard errors (clustered by hospital) in parentheses; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; (std) indicates standardized variables.

of documents such as invoices and purchase orders. *Medical Terms* is an indicator variable for whether hospitals adopted the use of controlled medical vocabularies. These variables are recorded in the HIMSS database (Section 3.1.1) and may change for a hospital during our focal time period, depending on when the hospital adopts. We used these instruments for the following reasons. They are likely to be correlated with *Any EMR* because EMRs and other systems are often implemented together as part of IT and process improvement initiatives (i.e., the instruments are likely to be relevant and not weak). Furthermore, the instruments are unlikely to affect malpractice claim resolution time directly (i.e., after the effect of EMRs is accounted for), because any relevant information they might provide will be encapsulated in the EMRs (i.e., the instruments are likely to be exogenous). We tested the validity of these instruments as follows. First, the instruments are significantly correlated with *Any EMR* and are not weak. We tested for instrument weakness using the Kleibergen-Paap rk Wald F statistic (Kleibergen and Paap 2006), which we used instead of the Cragg-Donald Wald F statistic because our error terms are clustered by hospital. This statistic ($F = 21.26$) exceeds the $\alpha = 0.05$ critical value (Stock and Yogo 2005), allowing us to reject the null hypothesis that the instruments are weak and generating evidence that the instruments are relevant. Second, because the system is overidentified (i.e., we have more instruments than potentially endogenous variables), we were able to use the Hansen test to assess whether the instruments were exogenous. The null hypothesis in the Hansen test is that the instruments are exogenous (technically, that they are not correlated with the residuals of the second stage of the instrumental variables regression). The Hansen test yielded a p-value of 0.34, thereby failing to reject the null hypothesis and suggesting that the instruments are exogenous. The instrumental variables regression results (Table 8, Model A1) are consistent with our earlier results. EMRs significantly reduce claim resolution time ($\beta = -157.66$, $p < 0.05$).

Despite this (and other) analysis, it remains possible that unobserved hospital characteristics are correlated with both faster claim resolution and EMR use. If that were the case, then faster resolution might be attributable to these unobserved confounders and not to EMRs. A Rosenbaum sensitivity analysis quantifies how influential these unobserved confounders would have to be to alter our conclusion (Rosenbaum 2002). To conduct this analysis, we first needed to generate a matched sample of claims originating in hospitals with basic EMR functionality (i.e., "treated"

Table 8 EMRs and Malpractice Claim Resolution Time – Alternative Models

Variables	Instrument Model A1	Hazard Model A2	Mixed Model A3	Truncated Model A4	Neg. Binomial Model A5
Hospital fixed effects	yes	see below	yes	yes	yes
Year indicators	yes	yes	yes	yes	yes
Constant			682.701^{*}	$-1,832.596^{*}$	4.976^{***}
			(283.750)	(875.379)	(0.988)
Severity: Minor	131.311^{***}	-0.287^{***}	134.979^{***}	290.111^{***}	0.223^{***}
	(24.218)	(0.038)	(25.883)	(38.715)	(0.048)
Severity: Significant	238.531^{***}	-0.515^{***}	243.577^{***}	474.218^{***}	0.390^{***}
	(28.400)	(0.046)	(29.578)	(43.661)	(0.051)
Severity: Major	188.564^{***}	-0.402^{***}	192.057^{***}	394.518^{***}	0.304^{***}
	(31.345)	(0.040)	(32.899)	(40.574)	(0.056)
Severity: Grave	213.848^{***}	-0.465^{***}	219.382^{***}	436.114^{***}	0.364^{***}
	(36.284)	(0.058)	(37.544)	(49.596)	(0.058)
Severity: Death	279.3249^{***}	-0.588^{***}	283.757^{***}	542.257^{***}	0.431^{***}
	(26.829)	(0.038)	(28.381)	(38.441)	(0.050)
Court	266.765^{***}	-0.439^{***}	269.277^{***}	390.534^{***}	0.308^{***}
	(16.502)	(0.025)	(16.574)	(16.994)	(0.019)
Employees (ln)	54.938	−0.051	29.126	100.407	0.071
	(56.519)	(0.088)	(16.035)	(88.835)	(0.098)
Working Hours per Patient (std)	-58.284^{*}	0.148^{***}	−30.170	-85.286^{**}	-0.086^{*}
	(27.864)	(0.043)	(19.527)	(32.950)	(0.039)
Salary per FTE (std)	18.438	−0.009	22.317	21.827	0.018
	(27.317)	(0.043)	(18.219)	(33.037)	(0.037)
Net Revenue per Patient (std)	23.210	−0.011	2.991	17.172	0.007
	(22.112)	(0.036)	(17.246)	(28.784)	(0.026)
Occupancy (std)	17.375	−0.032	8.662	28.729	0.020
	(21.286)	(0.045)	(11.844)	(35.076)	(0.030)
Geographic Index	5.624	−0.011	0.441	10.389^{*}	0.007
	(3.696)	(0.006)	(2.462)	(4.840)	(0.005)
Case Mix Index	−186.971	0.232	−102.063	-271.255^{*}	−0.209
	(132.472)	(0.165)	(59.639)	(129.895)	(0.148)
Use of Accepted Practices (std)	−4.749	0.022	−5.659	−17.435	−0.015
	(15.948)	(0.027)	(12.518)	(20.764)	(0.021)
Basic EMR		0.153^{**}	-51.024^{**}	-107.747^{**}	-0.078^{*}
		(0.047)	(16.499)	(38.817)	(0.034)
Advanced EMR		0.196^{**}	-82.622^{**}	-156.193^{**}	-0.109^{*}
		(0.065)	(31.778)	(52.026)	(0.049)
Any EMR	-157.661^{*}				
	(75.428)				
Observations	13,501	13,503	13,503	13,503	13,503

Instrumental variables regression using Document Management and Medical Terms as instruments (A1); Proportional hazard model on risk of resolution (A2); Multilevel mixed-effects linear regression (A3); Truncated linear regression (A4); Truncated negative binomial regression (A5). Standard errors (clustered by hospitals) in parentheses; hospital fixed effects accounted for in model A2 via stratification by hospitals; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; (std) indicates standardized variables.

claims) and those with no EMR functionality (i.e., "control" claims). We did this using propensity scoring (Leuven and Sianesi 2003). We matched claims based on *Year*, *Severity*, *Court*, *Employees*, *Working Hours per Patient*, *Salary per FTE*, *Net Revenue per Patient*, *Occupancy*, *Geographic Index*, *Case Mix Index*, and *Use of Accepted Practices.* After matching, we compared *Resolution Time* between treated and control claims. Resolution Time for treated claims was 122 days faster ($p < 0.001$). The objective of Rosenbaum sensitivity analysis is to quantity how much of an effect unobserved confounders would have to have to overturn our conclusion of a significant difference between treated and control claims. We used Wilcoxon's signed rank test for matched pairs to perform the sensitivity analysis (Rosenbaum 2005). We concluded that in order to attribute the faster resolution time to unobserved confounders, they would need to: a) be highly correlated with faster resolution, and b) increase the odds of the claimed-against hospital having basic EMRs by a factor of 1.42 (i.e., $\Gamma = 1.42$ in sensitivity analysis notation). Although there is no consensus about the appropriate size for Γ in social science research, $\Gamma = 1.5$ indicates substantial insensitivity and $\Gamma = 1.2$ is around average (Sen 2014). Figure 2 illustrates this result by showing that the 95% confidence interval for the estimated effect of basic EMRs does not include 0 until $\Gamma = 1.42$.

4.3. Alternative Analyses and Robustness Checks

To reduce the possibility that our results are due to our choice of empirical approach, we conducted several robustness checks. These address issues of a) model choice, b) potential truncation of claim resolution time, c) alternative measurement, and d) sample composition.

4.3.1. Model Choice: To ensure that our results are not specific to our choice of model specification, we used alternative specifications, the results of which are shown in Table 8.

First, we used a Cox proportional hazard model — stratified by hospital — to model how EMRs and other covariates affect claim resolution time (Table 8, Model A2). We find that the hazard of claim resolution is significantly increased with basic ($\beta = 0.15$, $p < 0.01$) and advanced ($\beta = 0.20$, $p < 0.01$) EMRs; i.e., both basic and advanced EMRs are associated with faster claim resolution. Figure 3 illustrates this by showing the Kaplan-Meier survivor curves estimated from this model.

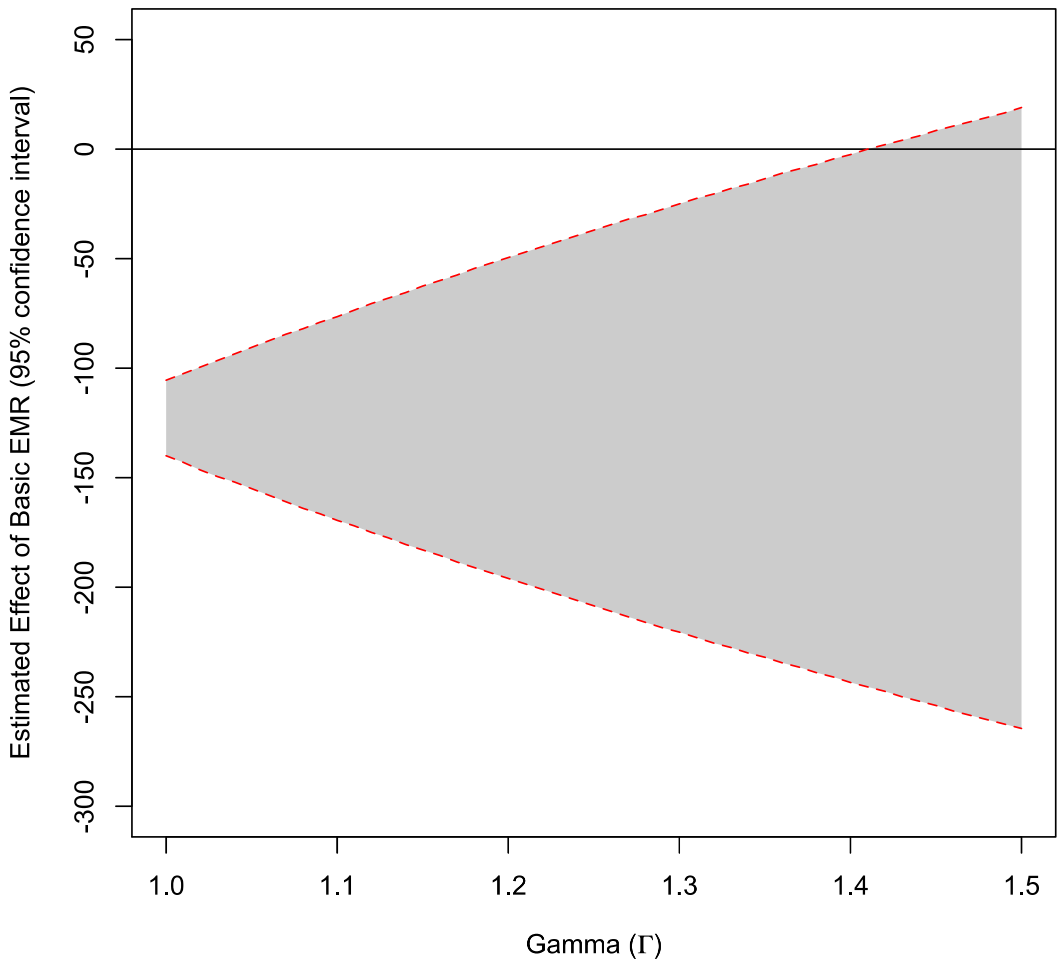

Figure 2 Rosenbaum Sensitivity of Estimated Effect of Basic EMR

While confirmatory, we elected not to focus on this specification because the claims in our analysis are not censored, and the coefficients of ordinary least squares regressions are easier to interpret.

Second, we implemented a multi-level model that mirrors the ordinary least squares model except that we allowed random intercepts for each hospital (Table 8, Model A3). We find that claim resolution time is significantly reduced with both basic ($\beta = -51.02$, $p < 0.01$) and advanced ($\beta = -82.62$, $p < 0.01$) EMRs.

Third, because claims cannot be resolved before they occur (i.e., the minimum of *Resolution Time* is 0), the assumption of normal distribution of error terms in ordinary least squares regression

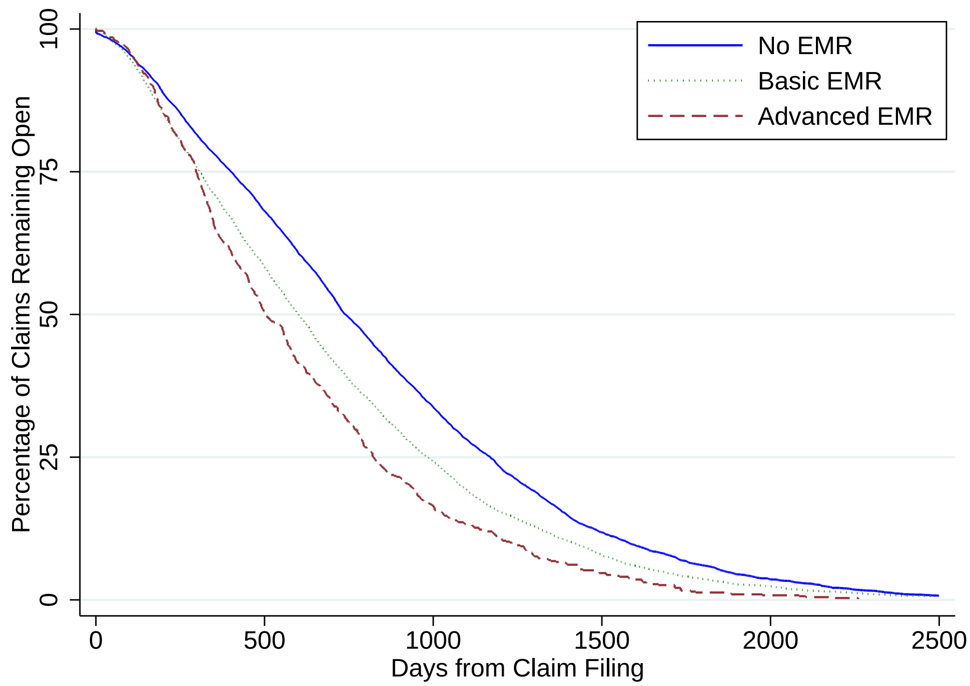

Figure 3 **Kaplan-Meier Survivor Curves for Malpractice Claims, by Degree of Hospital EMR use at the time of the Injury**

could be violated. To account for this, we estimated a truncated regression model (Table 8, Model A4). Again, we find that claim resolution time is significantly reduced with both basic ($\beta = -107.75$, $p < 0.01$) and advanced ($\beta = -156.19$, $p < 0.01$) EMRs. Also, because *Resolution Time* is an integer count of the number of days, we estimated a truncated negative binomial regression model (Table 8, Model A5). Again, we find that claim resolution time is significantly reduced with both basic ($\beta = -0.08$, $p < 0.05$) and advanced ($\beta = -0.11$, $p < 0.05$) EMRs.

4.3.2. Potential Truncation of Claim Resolution Time: Because of reporting requirements for the state of Florida, we do not observe claims until they are resolved. Although we collected data through June 2015, we limited our analysis to claims filed from 1999 to 2006 to allow adequate time for claims to be resolved and reported. This is consistent with prior analysis of malpractice claims (Quinn et al. 2012, Virapongse et al. 2008). However, to consider potential bias from truncation, we analyzed two additional models that allow additional years for claims to reach resolution. In the first model (Table 9, Model W1), we restricted our sample to claims filed from 1999 to 2005. We find that claim resolution time is significantly reduced with both basic ($\beta = -58.00$, $p < 0.05$) and advanced ($\beta = -117.24$, $p < 0.001$) EMRs. In the second model (Table 9, Model W2), we further restricted our analysis to the years 1999 to 2004 and find that claim resolution time is significantly reduced with both basic ($\beta = -74.94$, $p < 0.01$) and advanced ($\beta = -125.00$, $p < 0.001$) EMRs. Twelve of the 13,503 claims in our sample have resolution times longer than 8.5 years. If they had been reported the last day of 2006, they would have not had time to resolve within the time period of our focal analysis. Reducing the focal period to 1999 to 2004 (Table 9, Model W2) would have allowed all but one to resolve even if it was reported the final day of 2004.

To further limit potential problems due to potential truncation, we restricted our analysis to include only claims resolved within five years (Table 9, Model W3). We find that claim resolution time is significantly reduced with both basic ($\beta = -41.65$, $p < 0.05$) and advanced ($\beta = -82.17$, $p < 0.01$) EMRs. We similarly restrict the analysis to claims resolved within 6 years (Table 9, Model W4). Again, we find that claim resolution time is significantly reduced with both basic ($\beta = -51.44$, $p < 0.01$) and advanced ($\beta = -97.93$, $p < 0.001$) EMRs. This indicates that our findings are not a result of long resolution times during the early years of our sample when EMR use was relatively rare.

4.3.3. Alternative Measurement: Several measurement choices could be affecting the results. To investigate, we used alternative measures of claim resolution time and EMR use.

Research reported in *Health Affairs* on claim resolution time (Seabury et al. 2013) measures claim resolution time from the date the claim was filed (as we have done); alternatively, claim resolution time could be measured from the date the injury occurred. We reran our analysis using this alternative measure of resolution time (Table 10, Model M1) and continue to find that claim resolution time is significantly reduced with both basic ($\beta = -65.38$, $p < 0.01$) and advanced ($\beta = -105.59$, $p < 0.01$) EMRs. We do not observe the exact date when hospitals implemented component system that comprise EMRs; we only observe the year. Therefore, a claim occurring early in the year of

Table 9 EMRs and Malpractice Claim Resolution Time — Resolution Windows

Variables	Until 2005 Model W1	Until 2004 Model W2	5 years Model W3	6 years Model W4
Hospital fixed effects	yes	yes	yes	yes
Year indicators	yes	yes	yes	yes
Constant	−80.843	−238.176	−76.327	198.012
	(455.736)	(505.462)	(361.796)	(390.890)
Severity: Minor	132.934***	138.572***	134.313***	136.541***
	(18.308)	(19.850)	(14.702)	(15.972)
Severity: Significant	240.102***	241.001***	232.528***	231.864***
	(22.455)	(24.202)	(18.168)	(19.706)
Severity: Major	194.830***	203.672***	178.981***	182.197***
	(19.667)	(21.256)	(15.857)	(17.200)
Severity: Grave	213.173***	215.492***	223.308***	219.517***
	(27.936)	(30.058)	(22.713)	(24.629)
Severity: Death	283.702***	292.630***	251.471***	267.642***
	(18.213)	(19.723)	(14.681)	(15.929)
Court	270.810***	276.315***	237.838***	252.782***
	(11.872)	(12.455)	(10.011)	(10.710)
Employees (ln)	65.032	89.201	51.565	35.630
	(43.173)	(48.859)	(35.369)	(38.168)
Working Hours per Patient (std)	−61.414**	−57.390**	−41.168*	−49.445**
	(19.983)	(21.711)	(16.501)	(17.663)
Salary per FTE (std)	21.698	22.268	39.753*	25.584
	(20.827)	(23.241)	(16.611)	(17.923)
Net Revenue per Patient (std)	18.486	16.262	3.252	9.586
	(17.042)	(18.986)	(14.143)	(15.182)
Occupancy (std)	11.150	14.472	−17.420	10.178
	(21.986)	(24.569)	(17.332)	(18.683)
Geographic Index	6.094*	5.991	5.244*	4.013
	(3.055)	(3.367)	(2.367)	(2.564)
Case Mix Index	−151.899	−159.614	−108.212	−103.484
	(78.668)	(85.279)	(63.491)	(68.175)
Use of Accepted Practices (std)	−4.809	−2.614	−10.549	−10.840
	(13.234)	(15.033)	(10.558)	(11.374)
Basic EMR	−57.998*	−74.935**	−41.651*	−51.438**
	(24.785)	(29.040)	(18.252)	(19.842)
Advanced EMR	−117.243***	−124.999***	−82.167**	−97.931***
	(31.786)	(36.558)	(24.992)	(27.045)
Observations	12,756	11,716	12,882	13,265
R^2	0.158	0.133	0.166	0.171
F	107.0	84.09	109.9	117.6

OLS regression on claim resolution time (measured in number of days since claim filed); Sample restricted to pre-2005 (W1), pre-2004 (W2), within 5 years of claim filing (W3), and within 6 years of claim filing (W4). Standard errors (clustered by hospital) in parentheses; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; (std) indicates standardized variables.

implementation may not have had the EMR system in place at the time of the injury. To eliminate potential bias due to uncertainty about the specific implementation date, we ignored any claims during the year of EMR implementation (Table 10, Model M2). In this analysis, we find that claim resolution time is significantly reduced with both basic ($\beta = -63.07$, $p < 0.01$) and advanced ($\beta = -130.81$, $p < 0.001$) EMRs.

4.3.4. Sample Composition: Every hospital covariate is not available for all hospitals for which we have other data. As a result, these hospitals are excluded from our main analysis. We implemented an alternative model that uses only hospital fixed effects, which allowed us to include more hospitals in the analysis (Table 10, Model M3), increasing our sample size from 13,503 claims to 15,117 claims. We find that claim resolution time is significantly reduced with both basic ($\beta = -55.05$, $p < 0.01$) and advanced ($\beta = -93.74$, $p < 0.001$) EMRs.

Because our unit of analysis is the claim, our results implicitly grant more weight to hospitals with a larger number of claims. Our results could be influenced if hospitals with large claim volume all used (or did not use) EMRs. Therefore, we restricted our analysis to exclude claims originating in the top 10 hospitals by claim volume (Table 10, Model M4). Again, we find that claim resolution time is significantly reduced with both basic ($\beta = -68.77$, $p < 0.01$) and advanced ($\beta = -124.81$, $p < 0.001$) EMRs.

Based on these alternative analyses, we conclude that our focal findings are robust to a) model choice, b) potential truncation of claim resolution time, c) measurement choice, and d) sample composition. We believe that the focal model (ordinary least squares regressions for 1999–2006 with claim covariates, hospital covariates, a flexible time trend, and hospital fixed effects) presents the most parsimonious and complete analysis of claim resolution time.

4.4. Exploring the Effect of EMRs on Discovery

Our results show a significant and negative relationship between EMRs and malpractice claim resolution time. We believe this is because EMRs speed up discovery by permitting greater use of electronic discovery. However, it is possible that the negative relationship between EMRs and resolution time might be due to changes in other steps of the claim process. The relevant (i.e., post-filing) steps are discovery, mediation and negotiation, and trial (Berg 2016); the certificate of merit step is not relevant, because it must be filed with the initial complaint in

Table 10 EMRs and Malpractice Claim Resolution Time – Alternative Measurements and Sample Composition Checks

Variables	Based on Injury Date Model M1	Without EMR Adoption Year Model M2	Without Hospital Covariates Model M3	Without 10 Hospitals with the Largest Claim Volume Model M4
Hospital fixed effects	yes	yes	yes	yes
Year indicators	yes	yes	yes	yes
Constant	227.606	−58.954	786.365***	456.076
	(478.175)	(435.038)	(19.899)	(504.599)
Severity: Minor	202.398***	130.837***	125.590***	106.645***
	(19.609)	(17.993)	(16.517)	(20.337)
Severity: Significant	383.553***	243.949***	240.922***	210.405***
	(24.147)	(22.073)	(20.354)	(24.656)
Severity: Major	274.384***	186.601***	184.901***	142.702***
	(21.090)	(19.347)	(17.785)	(21.841)
Severity: Grave	365.264***	210.524***	210.583***	176.067***
	(30.198)	(27.605)	(25.084)	(30.639)
Severity: Death	334.694***	278.748***	271.861***	249.591***
	(19.533)	(17.922)	(16.435)	(20.238)
Court	359.270***	272.765***	258.532***	252.988***
	(13.010)	(11.836)	(10.981)	(13.088)
Employees (ln)	46.350	51.905		−26.835
	(46.675)	(42.271)		(54.374)
Working Hours per Patient (std)	−47.407*	−68.944***		−52.529*
	(21.499)	(19.649)		(21.733)
Salary per FTE (std)	−7.106	13.932		13.587
	(21.895)	(19.841)		(20.872)
Net Revenue per Patient (std)	−15.921	11.325		0.395
	(18.436)	(16.776)		(21.867)
Occupancy (std)	35.957	27.742		30.376
	(22.886)	(20.902)		(22.134)
Geographic Index	8.390**	4.804		3.023
	(3.139)	(2.842)		(3.008)
Case Mix Index	−153.295	−7.392		204.565
	(83.331)	(80.658)		(115.192)
Use of Accepted Practices (std)	−13.746	−9.615		−23.995
	(13.888)	(12.760)		(13.901)
Basic EMR	−65.376**	−63.068**	−55.048**	−68.769**
	(24.309)	(23.875)	(20.772)	(25.515)
Advanced EMR	−105.590**	−130.808***	−93.740***	−124.811***
	(33.129)	(31.467)	(26.888)	(36.411)
Observations	13,503	12,833	15,117	10,454
R^2	0.242	0.179	0.175	0.179
F	184.6	120.2	205.83	97.50

OLS regression on claim resolution time measured in number of days since injury date (M1), without including year of EMR adoption (M2), without hospital covariates (M3), and without 10 hospitals with largest claim volume (M4); Standard errors (clustered by hospital) in parentheses; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; (std) indicates standardized variables.

the state of Florida (`http://www.ncsl.org/research/financial-services-and-commerce/medical-liability-malpractice-merit-affidavits-and-expert-witnesses.aspx`.) To assess which step(s) is affected by EMRs, we reran our focal regression for only those claims that did not go to trial, i.e., for which *Court* = 0 ($n = 11,354$). The coefficient for *Basic EMR* is $\beta = -62.69$ ($p < 0.01$), which represents a 9% reduction in *Resolution Time* for this sub-sample. The coefficient for *Advanced EMR* is $\beta = -105.87$ ($p < 0.001$), which represents a 15% reduction. These coefficients are similar to those from the focal model (see Table 7: -62.69 versus -65.43 for *Basic EMR* and -105.87 versus -109.92 for *Advanced EMR*), and because they are OLS coefficients, they are directly comparable. This similarity suggests that the speed benefits of EMRs occur during the discovery and mediation / negotiation steps as opposed to the trial step. Because discovery often spans the majority of the claim's timeline (Berg 2016, HG.org Legal Resources 2016), it seems likely that most of the speed benefits of EMRs are a result of faster discovery.

5. Conclusion

Electronic trace data about what individuals do and when they do it are increasingly generated, stored, and analyzed. A growing body of research investigates the use of trace data for targeting advertising, determining trustworthiness and creditworthiness, and improving matching within markets. We focus on a relatively under-explored implication of trace data: their effect on the resolution of lawsuits. Specifically, we study the effect of EMRs — which store trace data of what care was administered, when, and by whom — on the resolution time of malpractice claims. Theoretically, EMRs could either speed or slow claim resolution, depending on whether the data they contain clarify or obfuscate the relevant issues. We collected a unique dataset consisting of 13,503 resolved malpractice claims originating in 148 hospitals in Florida between 1999 and 2006. On average, each claim took a little more than two years to be resolved. We found that having basic EMR functionality in the hospital at the time of the injury was associated with a more than two-month reduction in resolution time. Having advanced EMR functionality was associated with a more than three-month reduction. We believe the reduction in resolution time is due to a streamlined discovery process made possible by the electronic paper trail generated by EMRs, including what care was administered (or not), when, and by whom. Faster resolution times are likely to generate economic savings, not to mention the emotional benefits they generate for providers and patients. Even if no payment is made to the patient, the cost of defending a malpractice claim is substantial. One study estimated the average cost to be $22,959 (standard deviation $41,687) (Seabury et al. 2012); another estimated it to be $47,783 (State of Connecticut Insurance Department 2014).

Our study makes three main contributions. First, it contributes to the understanding of how electronic trace data affects legal outcomes, specifically their resolution time. This is non-obvious a priori and warrants empirical examination, because there are plausible arguments for why trace data might either speed or slow lawsuit resolution. The effect of electronic trace data on lawsuits is likely to be increasingly important in the future, as more and more trace data is generated and stored. Second, our study has practical implications for the medical malpractice system. The lengthy time required to resolve malpractice claims creates substantial costs, both monetary and emotional, for providers and patients. There is a need to identify ways to make the system more efficient and timely. Whereas prior analysis has focused on the potential of tort reform to improve the system, we show that healthcare information technology (specifically EMRs) can improve the system by speeding claim resolution time, leading to potentially substantial welfare benefits. Third, our study adds to a growing understanding of the effects of EMRs on healthcare. EMRs offer many potential benefits, including better care and lower costs for individual patients as well as society (Devaraj and Kohli 2000, Koppel et al. 2005). Such benefits are at the root of the HITECH Act and the Meaningful Use standard (Blumenthal 2009). Our study offers empirical evidence of a previously unexamined consequence of EMR use. While not necessarily an intended or primary goal, the effect of EMRs on claim resolution time is important for everyone involved. Given evidence that healthcare providers may be reluctant to install EMRs due to liability concerns (Miller and Tucker 2014), our research indicates a silver lining.

This study has important limitations. First, our data do not indicate how extensively each hospital used the component systems that comprise EMRs. Thus, our results should be interpreted as the relationship between an average degree of basic / advanced EMR use and claim resolution time. Second, because of the need to match data across multiple sources, we do not have a random sample of Florida hospitals. We assessed the representativeness of our sample using year 2000 data from the Florida Agency for Healthcare Information. Our sample includes 78% of the Florida hospital population and is representative on multiple measures, including percentage of teaching, public, and non-profit hospitals, although the hospitals in our sample are slightly larger than average (317.8 beds versus 297.5 for the entire state). As noted in Section 4.3.4, we re-estimated our models without the time-variant hospital covariates but with the hospital fixed effects. This allowed us to include almost all of the Florida hospital population in our analysis, and we found similar results. Third, our analysis is specific to Florida, although the malpractice claim activity in Florida appears to be representative of many other states (Mello et al. 2003, Jena et al. 2011, Bishop et al. 2011). Fourth, because the research question necessitates studying a time period in which claims have had time to resolve (we used the 1999–2006 time period, having collected data through June 2015), we cannot be sure if our results hold for contemporary EMRs that may have more sophisticated functionality. As such, continued analysis of the relationship between EMRs and malpractice claim resolution is warranted. Fifth, although we employ multiple strategies to limit the possibility that unobserved attributes of claims and hospitals confound our results, we cannot rule out this possibility. For example, attributes such as the medical specialty associated with the claim may affect resolution time (Seabury et al. 2013), although we have no evidence that these attributes were distributed disproportionately across EMR-using and non EMR-using hospitals in our study. Sixth, although we believe that EMRs speed claim resolution via faster discovery, other mechanisms are possible. For example, EMRs could reduce claim resolution time by preventing errors that would otherwise lead to particularly long malpractice cases. Future research can further isolate the mechanism(s) through which EMRs speed claim resolution.

References

Abelson, R., J. Creswell, G. J. Palmer. 2012. Medicare Bills Rise as Records Turn Electronic. *The New York Times* 21 September.

Adomavicius, G., A. Tuzhilin. 2011. Context-Aware Recommender Systems. F. Ricci, L. Rokach, B. Shapira, P. B. Kantor, eds., *Recommender Systems Handbook*. Springer US, 217–253.

Agarwal, R., G. Gao, C. DesRoches, A. K. Jha. 2010. The Digital Transformation of Healthcare: Current Status and The Road Ahead. *Information Systems Research* **21**(4) 796–809.

Angst, C.M., R. Agarwal. 2009. Adoption of Electronic Health Records in the Presence of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion. *MIS Quarterly* **33**(2) 339–370.

Angst, C.M., R. Agarwal, V. Sambamurthy, K. Kelley. 2010. Social Contagion and Information Technology Diffusion: The Adoption of Electronic Medical Records in US Hospitals. *Management Science* **56**(8) 1219–1241.

Angst, C.M., S. Devaraj, C. Queenan, B. Greenwood. 2011. Performance Effects Related to the Sequence of Integration of Healthcare Technologies. *Production and Operations Management* **20**(3) 319–333.

Armour, S. 2014. Borrowers Hit Social-Media Hurdles. *Wall Street Journal* 8 January.

August, T., M. F. Niculescu, H. Shin. 2014. Cloud Implications on Software Network Structure and Security Risks. *Information Systems Research* **25**(3) 489–510.

Balas, E.A. 2001. Information Systems Can Prevent Errors and Improve Quality. *Journal of the American Medical Informatics Association* **8**(4) 398–399.

Beisner, J.H. 2010. Discovering a Better Way: The Need for Effective Civil Litigation Reform. *Duke Law Journal* **60**(3) 547–596.

Berg, D. 2016. Timeline of a Medical Malpractice Lawsuit. URL `http://www.alllaw.com/articles/nolo/medical-malpractice/lawsuit-timeline.html`. Accessed 17 January 2016.

Bertrand, M., E. Duflo, S. Mullainathan. 2004. How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics* **119**(1) 249–275.

Bharadwaj, A., O.A. El Sawy, P.A. Pavlou, N. Venkatraman. 2013. Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly* **37**(2) 471–482.

Bhargava, H. K., A. N. Mishra. 2014. Electronic Medical Records and Physician Productivity: Evidence from Panel Data Analysis. *Management Science* **60**(10) 2543–2562.

Bishop, T.F., A.M. Ryan, L. Casalino. 2011. Paid Malpractice Claims for Adverse Events in Inpatient and Outpatient Settings. *Journal of the American Medical Association* **305**(23) 2427–2431.

Bleier, A., M. Eisenbeiss. 2015. Personalized Online Advertising Effectiveness: The Interplay of What, When, and Where. *Marketing Science* **34**(5) 669–688.

Blumenthal, D. 2009. Stimulating the Adoption of Health Information Technology. *New England Journal of Medicine* **360**(15) 1477–1479.

Burtch, G., A. Ghose, S. Wattal. 2015. The Hidden Cost of Accommodating Crowdfunder Privacy Preferences: A Randomized Field Experiment. *Management Science* **61**(5) 949–962.

Chan, S. 2008. When a MetroCard Is an Alibi. *The New York Times* 19 November.

Cohendet, P., F. Meyer-Krahmer. 2001. The Theoretical and Policy Implications of Knowledge Codification. *Research Policy* **30**(9) 1563–1591.

Cukier, K. 2010. Data, Data Everywhere. *The Economist* February 27.

Curfman, Gregory D., Stephen Morrissey, Jeffrey M. Drazen. 2013. High-Value Health Care — A Sustainable Proposition. *New England Journal of Medicine* **369** 163–1164.

Degnan, D. 2011. Accounting for the Costs of Electronic Discovery. *Minnesota Journal of Law, Science, & Technology* **12**(1) 151–190.

Devaraj, S., R. Kohli. 2000. Information Technology Payoff in the Health-Care Industry: A Longitudinal Study. *Journal of Management Information Systems* **16**(4) 41–67.

Dexter, P.R., S. Perkins, J.M. Overhage, K. Maharry, R.B. Kohler, C.J. McDonald. 2001. A Computerized Reminder System to Increase the Use of Preventive Care for Hospitalized Patients. *New England Journal of Medicine* **345**(13) 965–970.

DHHS Centers for Medicare and Medicaid Services. 2013. Hospital Compare. URL `http://www.cms.gov/HospitalQualityInits/11_HospitalCompare.asp`. August 8, 2013.

Dranove, D., S. Greenstein, A. Goldfarb, C. Forman. 2014. The Trillion Dollar Conundrum: Complementarities and Health Information Technology. *American Economic Journal: Policy* **6**(4) 239–270.

Gallagher, T.H., W. Levinson. 2005. Disclosing Harmful Medical Errors to Patients: A Time for Professional Action. *Archives of Internal Medicine* **165** 1819–1824.

Gart, M. 2008. Medical Malpractice: The First Stage. *Physician Executive* **34**(1) 77–79.

Gelles, D. 2011. Inside Match.com. *The Financial Times* 11 July.

Government Accountability Office. 2003. *Medical Malpractice Insurance: Multiple Factors Have Contributed to Increased Premium Rates*. Government Affairs Office, Report No. GAO-03-702.

Gray, J.E., D.A. Goldmann. 2004. Medication Errors in the Neonatal Intensive Care Unit: Special Patients, Unique Issues. *Archives of Disease in Childhood: Fetal and Neonatal* **89**(6) F472–F473.

Hatton, C. 2015. China 'Social Credit': Beijing Sets Up Huge System. *BBC News* 29 October.

Healthcare Information and Management Systems Society. 2013. HIMSS Analytics Database.

Healthcare Information and Management Systems Society. 2015. EMR Adoption Model. URL `https://app.himssanalytics.org/stagesGraph.asp`. Accessed 17 January 2016.

Hersh, W. 2004. Health Care Information Technology: Progress and Barriers. *Journal of the American Medical Association* **292**(18) 2273–2274.

HG.org Legal Resources. 2016. Why Does a Lawsuit Take So Long? URL `http://www.hg.org/article.asp?id=31734`. Accessed 17 January 2016.

Hitsch, G. J., A. Hortaçsu, D. Ariely. 2010. Matching and Sorting in Online Dating. *The American Economic Review* **100**(1) 130–151.

Hobgood, C., J.H. Tamayo-Sarver, A. Elms, B. Weiner. 2005. Parental Preferences for Error Disclosure, Reporting, and Legal Action After Medical Error in the Care of Their Children. *Pediatrics* **116** 1276–1286.

Hoffman, S. 2010. Employing E-Health: The Impact of Electronic Health Records on the Workplace. *Kansas Journal of Law and Public Policy* **19** 409–528.

Hoffman, S., A. Podgurski. 2009. EHealth Hazards: Provider Liability and Electronic Health Record Systems. *Berkeley Technology Law Journal* **24**(4) 1523–1582.

Hsu, J.S-C., S-P. Shih, Y. W. Hung, P. B. Lowry. 2015. The Role of Extra-Role Behaviors and Social Controls in Information Security Policy Effectiveness. *Information Systems Research* **26**(2) 282–300.

Institute of Medicine. 2011. *Health IT and Patient Safety: Building Safer Systems for Better Care*. The National Academies Press.

Jena, A.B., S. Seabury, D. Lakdawalla, A. Chandra. 2011. Malpractice Risk According to Physician Specialty. *New England Journal of Medicine* **365**(7) 629–636.

Jha, A.K., C.M. DesRoches, E.G. Campbell, K. Donelan, S.R. Rao, T.G. Ferris, A. Shields, S. Rosenbaum, D. Blumenthal. 2009. Use of Electronic Health Records in US Hospitals. *New England Journal of Medicine* **360**(16) 1628–1638.

Kellermann, A.L., S.S. Jones. 2013. What It Will Take To Achieve The As-Yet-Unfulfilled Promises Of Health Information Technology. *Health Affairs* **32**(1) 63–68.

Kessler, D.P. 2011. Evaluating the Medical Malpractice System and Options for Reform. *Journal of Economic Perspectives* **25**(2) 93–110.

Kim, E.K.J. 2006. The New Electronic Discovery Rules: A Place for Employee Privacy? *The Yale Law Journal* **115**(6) 1481–1490.

Kleibergen, F., R. Paap. 2006. Generalized Reduced Rank Tests using the Singular Value Decomposition. *Journal of Econometrics* **133**(1) 97–126.

Koppel, R., J.P. Metlay, A. Cohen, B. Abaluck, A.R. Localio, S.E. Kimmel, B.L. Strom. 2005. Role of Computerized Physician Order Entry Systems in Facilitating Medication Errors. *Journal of the American Medical Association* **293**(10) 1197–1203.

Korin, J.B., M.S. Quattrone. 2007. Electronic Health Records Raise New Risks of Malpractice Liability. *New Jersey Law Journal* **19**.

Kuperman, G.J., A. Bobb, T.H. Payne, A.J. Avery, T.K. Gandhi, G. Burns, D.C. Classen, D.W. Bates. 2007. Medication-related Clinical Decision Support in Computerized Provider Order Entry Systems: A Review. *Journal of the American Medical Informatics Association* **14**(1) 29–40.

Lambrecht, A., C. Tucker. 2013. When Does Retargeting Work? Information Specificity in Online Advertising. *Journal of Marketing Research* **50**(5) 561–576.

Leuven, E., B. Sianesi. 2003. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. URL `http://ideas.repec.org/c/boc/bocode/s432001.html`.

Lunt, C., N. Galbreath, J. Winner. 2010. Ranking Search Results Based on the Frequency of Clicks on the Search Results by Members of a Social Network who are within a Predetermined Degree of Separation. US Patent 7,788,260.

Luoma, V. M. 2006. Computer Forensics and Electronic Discovery: The New Management Challenge. *Computers & Security* **25**(2) 91–96.

Mangalmurti, S.S., L. Murtagh, M.M. Mello. 2010. Medical Malpractice Liability in the Age of Electronic Health Records. *New England Journal of Medicine* **363**(21) 2060–2067.

McNickle, M. 2011. 10 Things You Hate about Your EMR. URL `http://www.healthcareitnews.com/news/10-things-you-hate-about-your-emr`. 6 October; Accessed 17 January 2016.

Mearian, L. 2015. Lawyers Smell Blood in Electronic Medical Records. *ComputerWorld* April 13.

Mello, M. M., D. M. Studdert, T. A. Brennan. 2003. The New Medical Malpractice Crisis. *New England Journal of Medicine* **348** 2281–2284.

Mello, M.M., A. Chandra, A.A. Gawande, D.M. Studdert. 2010. National Costs of the Medical Liability System. *Health Affairs* **29** 1569–1577.

Mendelson, H., R. R. Pillai. 1998. Clockspeed and Informational Response: Evidence from the Information Technology Industry. *Information Systems Research* **9**(4) 415–433.

Menon, N.M, R. Kohli. 2013. Blunting Damocles' Sword: A Longitudinal Model of Healthcare IT Impact on Malpractice Insurance Premium and Quality of Patient Care. *Information Systems Research* **24**(4) 918–932.

Menon, N.M., B. Lee, L. Eldenburg. 2000. Productivity of Information Systems in the Healthcare Industry. *Information Systems Research* **11**(1) 83–92.

Miller, A.R., C.E. Tucker. 2009. Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records. *Management Science* **55**(7) 1077–1093.

Miller, A.R., C.E. Tucker. 2014. Electronic Discovery and the Adoption of Information Technology. *Journal of Law, Economics, and Organization* **30**(2) 217–243.

Miller, J., J. Glusko. 2003. Standing Up to the Scrutiny of Medical Malpractice. *Nursing Management* **34**(10) 20–22.

Mitra, S., S. Ransbotham. 2015. Information Disclosure and the Diffusion of Information Security Attacks. *Information Systems Research* **26**(3) 565–584.

Morris, C. 2013. The Sexiest Job of the 21st Century: Data Analyst. *CNBC* 5 June.

Pace, N.M., L. Zakaras. 2012. *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery*. RAND Corporation.

Page, L. 2001. Method for Node Ranking in a Linked Database. US Patent 6,285,999.

Peckham, C. 2015. Medscape Malpractice Report 2015: Why Most Doctors Get Sued. URL `http://www.medscape.com/features/slideshow/public/malpractice-report-2015`. 9 December; Accessed 17 January 2016.

Phillips, R., A. S. Şimşek, G. van Ryzin. 2015. The Effectiveness of Field Price Discretion: Empirical Evidence from Auto Lending. *Management Science* **61**(8) 1741–1759.

Pierce, L., D. C. Snow, A. McAfee. 2015. Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity. *Management Science* **61**(10) 2299–2319.

Quinn, M. A., A. M. Kats, K. Kleinman, D. W. Bates, S. R. Simon. 2012. The Relationship Between Electronic Health Records and Malpractice Claims. *Archives of Internal Medicine* **172**(15) 1187–1189.

Ransbotham, S., D. Kiron, P.K. Prentice. 2015. Minding the Analytics Gap. *MIT Sloan Management Review* **56**(3) 63–68.

Rosenbaum, Paul R. 2002. *Observational Studies*. Springer.

Rosenbaum, P.R. 2005. Sensitivity Analysis in Observational Studies. B.S. Everitt, D.C. Howell, eds., *Encyclopedia of Statistics in Behavioral Science*, vol. 4. John Wiley & Sons, Ltd, 1809–1814.

Sage, W.M. 2004. Reputation, Malpractice Liability, and Medical Error. V. A. Sharpe, ed., *Accountability: Patient Safety and Policy Reform*, chap. 10. Georgetown University Press, Washington (DC), 159–183.

Seabury, S., A. Chandra, D. Lakdawalla, A.B Jena. 2012. Defense Costs of Medical Malpractice Claims. *New England Journal of Medicine* **366** 1354–1356.

Seabury, S.A., A. Chandra, D.N. Lakdawalla, A.B Jena. 2013. On Average, Physicians Spend Nearly 11 Percent of Their 40-Year Careers with an Open, Unresolved Malpractice Claim. *Health Affairs* **32**(1) 111–119.

Sen, M. 2014. How Judicial Qualification Ratings May Disadvantage Minority and Female Candidates. *Journal of Law and Courts* **2**(1) 33–65.

Sidorov, J. 2006. It Ain't Necessarily So: The Electronic Health Record And The Unlikely Prospect Of Reducing Health Care Costs. *Health Affairs* **25**(4) 1079–1085.

Soumerai, S., R. Koppel. 2012. A Major Glitch for Digitized Health-Care Records. *The Wall Street Journal* 17 September.

State of Connecticut Insurance Department. 2014. Connecticut Medical Malpractice Annual Report. URL `http://www.ct.gov/cid/lib/cid/2014_Connecticut_Medical_Malpractice_Report.pdf`. Accessed 17 January 2016.

Stock, J.H., M. Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. J.H. Stock, D.W.K. Andrews, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*.

Studdert, D.M., M.M. Mello, T.A. Brennan. 2004. Medical Malpractice. *New England Journal of Medicine* **350**(3) 283–292.

Studdert, D.M., M.M. Mello, A.A. Gawande, T.K. Gandhi, A. Kachalia, C. Yoon, A.L. Puopolo, T.A. Brennan. 2006. Claims, Errors, and Compensation Payments in Medical Malpractice Litigation. *New England Journal of Medicine* **354**(19) 2024–2033.

Swartz, N. 2007. Firms Unprepared for E-Discovery. *Information Management Journal* **41**(6) 6.

Thorpe, K.E. 2004. The Medical Malpractice "Crisis": Recent Trends and the Impact of State Tort Reforms. *Health Affairs* **23** 20–30.

Victoroff, M.S., B. M. Drury, E. J. Campagna, E. H. Morrato. 2012. Impact of Electronic Health Records on Malpractice Claims in a Sample of Physician Offices in Colorado: A Retrospective Cohort Study. *Journal of General Internal Medicine* **28**(5) 637–644.

Virapongse, A., D.W. Bates, P. Shi, C.A. Jenter, L.A. Volk, K. Kleinman, L. Sato, S.R. Simo. 2008. Electronic Health Records and Malpractice Claims in Office Practice. *Archives of Internal Medicine* **168**(21) 2362–2367.

Volonino, L. 2003. Electronic Evidence and Computer Forensics. *Communications of the Association for Information Systems* **12** 457–468.

Wei, Y., P. Yildirim, C. Van den Bulte, C. Dellarocas. 2016. Credit Scoring with Social Network Data. *Marketing Science* Forthcoming.

Zuboff, S. 1988. *In the Age of the Smart Machine: The Future of Work and Power*. Basic Books.