APPROVAL SHEET

Title of Dissertation: Assessing and Analyzing the Meaningful Use Compliance by Medicaid Providers

Name of Candidate: Swati Verma
Master of Science in Computer Science, 2019

Dissertation and Abstract Approved: ______________________________
Dr. Gunes Koru
Professor
Department of Information Systems

Date Approved: 07/11/2019

ABSTRACT

Title of Thesis: ASSESSING AND ANALYZING THE MEANINGFUL USE COMPLIANCE BY MEDICAID PROVIDERS

Swati Verma, Master of Science, 2019
Department of Computer Science and Electrical Engineering

Directed by: Dr. Gunes Koru
Professor
Department of Information Systems

Meaningful Use program in the U.S., encourages the use of Electronic Health Records(EHRs). A number of incentive programs have been set up to adopt EHRs and exhibit Meaningful Use. To obtain incentives, the providers are required to give an account of their medicaid encounters and meet the Meaningful Use standards. The electronic Medicaid Incentive Program Payment(eMIPP) system maintains this account. This system is used by the Maryland Department of Health(MDH), for registration and attestation of eligible providers. This study provides the factors to be considered while measuring EHR adoption levels. To assist in making informed decisions while providing incentives, this study aims at (1) determining factors affecting the meaningful use compliance for medicaid providers, (2) determining the level to which the determined factors in (1), affect the Meaningful Use compliance, and (3) providing a tool, that automates the process of computing descriptive statistics based on reported values.

ASSESSING AND ANALYZING THE MEANINGFUL USE COMPLIANCE BY MEDICAID PROVIDERS

By

Swati Verma

Thesis submitted to the Faculty of the Graduate School of the University of Maryland, Baltimore County, in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science

2019

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

With the growing need of changing the health information electronically and providing better quality care for patients, there is increased emphasis on use of Electronic Health Records(EHRs). Electronic Health Records is an initiative towards reforming the health sector, with the increase in use of internet and web technologies. EHRs assist medicaid providers better manage care for patients, by storing the patient data in a structured format. EHRs, in addition to a structured format, make patient records easily accessible and updatable at the point of care (HealthIT.gov, 2018).

To encourage the use of EHRs, the Centers for Medicare and Medicaid Services (CMS) introduced the Meaningful Use program. The goal of this program is to promote widespread adoption of EHR systems and improve the quality, safety and efficiency of patient care. This program provides incentive payments to eligible provides and eligible hospitals, who demonstrate meaningful use of certified EHR technology (CMS.gov, 2019). The providers are required to meet a number of measures set by the CMS, as a part of the program, to receive payments. The Electronic Medicaid Incentive Program Payment (eMIPP) system, is a system which facilitates the enrollment in the Maryland

Medicaid EHR Incentive Program. It is used for registration and attestation of eligible providers to obtain incentives in stages.

A number of previous works on EHR adoption have been carried out. Some of them suggest factors, for instance, technology interaction and social contagion among the health care providers as the factors influencing EHR adoption (Gan & Cao, 2014). The studies discuss how decisions made by other health care providers influence one's adoption of technology at the health care points. In addition, it is stated that the way in which technology interacts and impacts the performance of activities, of a health care association, is a factor influencing EHR adoption. Some other factors suggested by studies include the EHR system deployment affordability, regional variation and physician resistance (Inoue & Zhou, 2016).

This study determines the factors associated with the health care providers, that is provider characteristics, influencing EHR adoption. These provider characteristics stated during the study, influence meaningful use and can be taken in consideration while providing incentives to the medicaid providers. In addition to determination of factors, the study also provides the level to which these factors affect the Meaningful Use compliance. These measures about the compliance are provided using statistical values for instance, median and other numbers associated with the providers.

During the study, a data set with more than four thousand records is used. This data was collected by the eMIPP system. The calculation of statistical measures for a data set this large, is a tedious process and prone to human error. To avoid having errors while computing the measures, a reporting tool was developed during the study. The tool automates the complex task of computing the descriptive statistical values associated with the meaningful use measures. The tool generates reports as required by the user, that is, MDH in this case. In addition to generating reports for Meaningful Use, the tool also generates reports for another measurement domain known as Clinical Quality Measures.

With the feature of providing reports containing the descriptive statistical values for meaningful use and clinical quality measures, the tool also assists in fetching lists containing providers, in an order as required by the user. This is another type reports provided by the tool. The user, using the filters, can fetch the details about providers attested to a certain measure in the measurement domain. The providers fetched can be sorted in order of sort by filters, for instance, performance of a provider.

Apart from generating reports, the tool also provides data visualization feature in the form of pie-charts and histograms. This feature of graphical representation provides a better understanding and enables user to look into aspects that may have been overlooked in textual representation. Some other features of convenience provided by the tool are (1) different types of data filters for better insight into the data, (2) saving reports in pdf, excel and csv document formats in the user's personal directory, (3) features like copying text from tool, zoom in and zoom out of text in reports to make text more readable, (4) a detailed user manual, to direct the user.

This study is organized into seven sections including the introduction section. Section 2 provides background information for this study and elaborates on how health IT adoption can measured. Section 3 mentions the methodology adopted for data preparation and data analysis. It also states the formulas used for computing the values in reports produced by the reporting tool. Section 4 lists the results of the study including predictive data analysis and the software developed. Section 5 is the discussion section. It elaborates on similar kind of studies carried out in the past. It also discusses the finding of this study. Section 6 lists all the limitations associated with this study. Section 7 states the conclusion of the study.

Chapter 2

BACKGROUND

In this thesis, the background section develops from a general and more extensive dialog about adoption of technology among providers and how this compliance with technology is quantified.

2.1 Information Technology in Health Care

In the present scenario, much of what people use is created with the assistance of Information Technology(IT). The ongoing improvements in IT, have led to significant reforms in the health care sector. The main features of IT are digital storage, retrieval and exchange of information. The application of these features in health care has led to improvements in quality of care and how medical information is shared, stored and used. In addition, this has led to the notion of Health Information Technology(HIT). HIT is the comprehensive management of information among consumers, clinicians, government and insurers. A number of steps have been taken up by the government towards digitalizing health care. EHRs is one of such initiatives. EHRs help reduce medical errors, improve the patient safety and quality of care programs. The data relevant to the patient's care, demographics, problems, medications, past medical immunizations and medical history is included in EHRs. The providers use computer softwares, that is EHR systems, that maintain all aspects of patient care in addition to billing, scheduling. A number of health care providers adopt EHR systems to meet the Health Information Technology for Economic and Clinical Health Act (HITECH) and some to just modernize their operations. The HITECH Act's goal is to promote the application of EHR systems and facilitate electronic information sharing among the health care providers and patients (Piliouras et al., 2016). EHR systems automate access to information maintained by health care providers. While EHRs are used to store and transmit patient records, they can also be used to transform patient information in meaningful ways and make use of that information at the point of care. For instance, the patient data can be used for analysis and the treatments can be based on the outcomes. Other components of the health IT infrastructure, similar to EHRs, are Electronic Medical Record(EMR), Personal Health Record(PHR) and Health Information Exchange(HIE). As EHRs become vast as more data is collected, working with EHRs also have some challenges. Some of issues faced by providers are implementation challenges, maintenance of EHR systems, upgradation of EHR systems and data security (Ajami & Bagheri-Tadi, 2013).

2.2 Health Care Administration System in study

Health care Administration Systems are the systems designed to assist health care providers to collect, store and exchange patient Health care information more efficiently. They monitor and manage administrative tasks such as billing, registration and reimbursements, etc. The data collected by these systems, is collected by means of activities involving attestation for meaningful use of EHR data, patient details such as social security numbers and financial incentive details. The health care administration system mentioned in the study, manages a value based reimbursement program that encourages the use of EHRs and health information exchange. This health care administration system was adopted under HITECH act to promote the meaningful use of EHRs. The system is established to facilitate the enrollment in the EHR incentive program. The program managed by the system provides incentive payments to eligible providers and eligible hospitals as they adopt, manage and demonstrate meaningful use of certified EHR technology. The system undergoes evaluation and new features that use the collected data efficiently are added to the system by the developers.

2.3 Measure Health IT Adoption

In order to encourage the implementation of EHRs by providers, the department concerned with federal health care programs has set up incentive programs for eligible providers. The programs promote innovation and encourage reduction of burden on providers, to maintain patient records manually. The purpose of these programs is to set values for measures stated under the programs. The providers are required to meet the values set, in order to qualify for the incentives. The measures quantify the aspects of patient care, for instance, efficient use of health care resources, improving quality, safety and public health. The providers are required to demonstrate the meaningful use of EHR technology, in a manner that demonstrates electronic exchange of health information to improve the quality of care. The adoption of HIT by providers is measured using the numbers provided by the providers for the measures set by the programs.

Chapter 3

METHODS

The methods section here elaborates on the methodology followed for the data analysis and the preparation of the data for analysis to determine the relationship between Electronic Health Record (EHR) adoption levels and provider characteristics of medicaid providers.

3.1 Data Preparation

The data used in the study, was received from the Maryland Department of Health, collected by their health administration system, eMIPP. The data was in the form of text files: Sixty four text files for Clinical Quality Measures (CQM) and twelve text files for Meaningful Use Measures (MU). The data pre-processing was done in the R statistical environment. The goal of the data preparation procedure is to make data suitable for the analysis in R and to generate a database file which can be used by the reporting tool developed during the study. The initial steps of data processing focus on cleaning the data and removing unneeded lines. The data files received contain unneeded lines and characters for instance, the SQL query used to obtain the data dump given at the top of the file, blank lines, and the lines containing unneeded hyphen characters. These items are removed from the text files. Next, the text files are converted to comma separated files (CSV). The records in the files contain some duplicate records i.e., the rows that have the same National Provider Identifier (NPI) but blank values in the remaining columns. These duplicate records with blank values are removed using R script. The twelve CSV files for Meaningful Use are merged into a single CSV file on the basis of the NPI column in all files. This merge process avoids any duplication of columns that are common to all files in the resulting single CSV file. The same steps are followed for the sixty four file for CQM and a single CSV file is obtained. Finally, a database file is created and the single CSV files are imported as two separate tables in the database. The single CSV files created for the two domains MU and CQM are used for data analysis steps in R. The database created is used as a data source for the eMIPP reporting tool developed during the study.

Once the data is prepared, the descriptive statistics for MU and CQM measures are generated in R. The blank values are not considered during the process and only the numerical values are taken into consideration. The descriptive statistics generated, include mean, standard deviation, median, and percentiles. To obtain the mentioned descriptive statistics, the describe function in Hmisc package in R is used. The values are calculated in R first in order to compare the values calculated by the reporting tool. The comparison of values calculated using R, with the results later obtained from the tool indicate complete consistency at all times.

3.2 Formulas Used for Calculations

For Meaningful Use:

Attestation: A provider is said to have attested to a measure if, the numerator is a non-negative integer and denominator is a positive integer.

NUM: Numerator of a given measure

DENOM: Denominator of a given measure

Rate for a provider

$$rate = \frac{NUM}{DENOM} \tag{3.1}$$

blank values are ignored by the software when rate is calculated.

Number of unique providers attested to the measure (NUPA): count of all the providers who have attested to a given measure.

Number of unique providers who met the threshold (NPMT): count of all the providers whose calculated rate is greater than the threshold value for the given measure.

Number of Exclusions (NEXCL) = NUPA - NPMT

Percentage Exclusion

$$\%Exclusion = \frac{NEXCL}{NUPA} * 100 \tag{3.2}$$

Mean

$$mean = \frac{\sum rate}{NUPA} \tag{3.3}$$

blank values are ignored by the software when mean is calculated.

Standard deviation

$$sdev = \sqrt{\frac{\sum(rate - \bar{rate})^2}{NUPA - 1}} \tag{3.4}$$

For Clinical Quality Measures:

Attestation: A provider is said to have attested to a measure if, the numerator is a non-negative integer and denominator is a positive integer.

NUM: Numerator of a given measure

DENOM: Denominator of a given measure

Rate for a provider

$$rate = \frac{NUM}{DENOM} \tag{3.5}$$

Number of un-duplicated providers who selected (NUPS): count of all the providers who have attested to a given measure.

Number of providers who entered zero in the denominator: count of all the providers who entered zero in the denominator for a given measure.

Number of Exclusions (NEXCL): count of providers who entered data for an exclusion on the measure (when applicable) that was greater than zero.

Percentage Exclusion

$$\%Exclusion = \frac{NEXCL}{NUPS} * 100 \tag{3.6}$$

Number of Exceptions (NEXP): count of providers who entered data for an exception on the measure (when applicable) that was greater than zero.

Percentage Exception

$$\%Exception = \frac{NEXP}{NUPS} * 100 \tag{3.7}$$

Mean

$$mean = \frac{\sum rate}{NUPS} \tag{3.8}$$

blank values are ignored by the software when mean is calculated.

Standard deviation

$$sdev = \sqrt{\frac{\sum(rate - \bar{rate})^2}{NUPS - 1}} \tag{3.9}$$

3.3 Data Analysis

In the data analysis part, the relationship between the EHR adoption levels and the medicaid provider characteristics is determined. The data analysis is done in the R environment. The CSV files created from the text files, for the two domains: MU and CQM are used as the data set for the analysis. In addition, other files having provider addresses, median incomes, zipcodes and rurality status are used in the analysis. The rurality and median income are associated with the zip codes which in turn is associated with the National Provider Identifier (NPI). The files are merged on the basis of NPI for the analysis. The records in the files for the two measurement domains, are the eMIPP Medicaid provider attestations. Attestation is defined as a process where providers report their medicaid encounters to the Maryland state government via the eMIPP system, in order to obtain incentives. The providers have corresponding numerator and denominator values under each measure in the data set. The predictors used for the analysis are patient volume, median income, payment year, rurality, self paid providers, and self supported providers. In statistics, a predictor is an independent variable which is manipulated in a model to observe the effect on the outcome variable. As the data set in consideration does not contain much readily available predictors, the predictors patient volume, self paid providers, and self supported provider are created during the study for the purpose of the analysis. The (1) patient volume is determined by considering the values in the denominator column given for each measure, (2) self paid providers are determined by checking whether the National Provider Identifier (NPI) was the same as Payee NPI, (3) self supported providers are determined by checking if Payee NPI was the same as Organization NPI. The response variable in statistical analysis is the dependent variable or the outcome variable. Non-Compliance is taken as the response variable in the study. Non-Compliance for each measure is determined by subtracting the values in the numerator from the corresponding denominator values. The data analysis involves use of techniques including linear model estimation, cubic splines. The linear model estimation using Ordinary Least Squares from the Regression Modeling Strategies (rms) package in R is used for developing the statistical model. Cubic splines are used in order to relax the assumption of linearity in the model. The spline functions are piecewise polynomials used in curve fitting. Cubic splines are piecewise functions of polynomial degree 3. The plotting in the study is generated using ggplot from the rms package in R. In data analysis, Analysis of variance (ANOVA) is a way to find out if the model results are significant. Here, ANOVA is used in the study to compare two or more models. The model considering individual predictors is compared to complete model considering all predictors.

Anova in R is a hypothesis testing technique for analyzing the amount of variance that is contributed to a sample due to various factors. It is a collection of estimation procedures, which can be used to analyze the differences among the group means in a data set. In addition, it can be used to test different groups to see if there is a similarity between them. Using this feature in this study, models for different measures have been compared, using the p-values and the partial sum of square values produced by anova in R.

The partial sum of square values represent the variation or deviation from the mean. It is the sum of the squares due to the source. It can be taken into account when there is a need to assess significance of each independent variable. The partial sum of squares for a specific variable measures the increase in the regression sum of squares by adding the variable to the model. The partial sum of squares and the p-values for the independent variables in the study have been listed in the table below, for the various models in the study.

Table 3.1: Partial Sum of Squares and p-values from Anova models 1-4

	M1	M2	M3	M4
RURAL				
Partial SS	3847.909	137070.381	385.91243	1.094206×10^4
P-value	0.652	0.0053	0.3789	0.4093
Median				
Partial SS	200859.989	703283.883	10346.21953	2.318494×10^5
P-value	0.0051	<0.0001	<0.0001	0.0008
Nonlinear				
Partial SS	140925.343	61080.779	8742.94834	8.849313×10^4
P-value	0.0065	0.0622	<0.0001	0.0191
PYMNT_YEAR				
Partial SS	119877.071	121984.892	989.26533	1.648631×10^4
P-value	0.012	0.0085	0.1592	0.3111
VOLUME				
Partial SS	405193.824	1284365.489	79391.07562	6.630888×10^6
P-value	<0.0001	<0.0001	<0.0001	<0.0001
Nonlinear				
Partial SS	112846.623	441848.884	10568.50052	1.692771×10^5
P-value	0.0148	<0.0001	<0.0001	0.0012
SELF_SUPPORTED				
Partial SS	16552.592	1719.107	49.92833	7.787040×10^4
P-value	0.3497	0.7541	0.7515	0.0279
SELF_PAID				
Partial SS	48822.49	19037.778	204.61377	7.326377×10^1
P-value	0.1084	0.2974	0.5216	0.9462
TOTAL NONLINEAR				
Partial SS	315150.18	542664.237	18259.79528	2.475215×10^5
P-value	0.0003	<0.0001	<0.0001	0.0005
REGRESSION				
Partial SS	769139.664	2416940.627	109773.07767	6.975839×10^6
P-value	<0.0001	<0.0001	<0.0001	<0.0001
ERROR				
Partial SS	15823287.442	13501646.577	205905.94833	1.345188×10^7

Table 3.2: Partial Sum of Squares and p-values from Anova models 5-8

	M5	M6	M7	M8
RURAL				
Partial SS	98759.179	2.325030×10^{2}	29824.728	107193.815
P-value	0.0335	0.6138	0.062	0.0002
Median				
Partial SS	264178.502	9.747959×10^{3}	64809.593	182314.299
P-value	0.0024	0.0049	0.0228	<0.0001
Nonlinear				
Partial SS	37964.451	4.181561×10^{3}	37472.251	110543.357
P-value	0.1871	0.0325	0.0365	0.0002
PYMNT_YEAR				
Partial SS	13359.784	6.690112×10^{-1}	74626.736	38316.789
P-value	0.4338	0.9784	0.0032	0.0274
VOLUME				
Partial SS	5435718.927	1.527786×10^{5}	2521234.208	45451888.362
P-value	<0.0001	<0.0001	<0.0001	<0.0001
Nonlinear				
Partial SS	368852.218	1.385845×10^{4}	1246.998	74870.561
P-value	<0.0001	0.0001	0.7025	0.0021
SELF_SUPPORTED				
Partial SS	3011.814	2.206477×10^{2}	2520.822	4896.148
P-value	0.7101	0.623	0.5872	0.4298
SELF_PAID				
Partial SS	16803.819	1.472910×10^{3}	8125.011	6649.922
P-value	0.38	0.2042	0.3297	0.3576
TOTAL NONLINEAR				
Partial SS	396945.406	2.248295×10^{4}	38334.162	197438.046
P-value	0.0001	<0.0001	0.1066	<0.0001
REGRESSION				
Partial SS	5806924.761	1.959723×10^{5}	3008902.97	52418263.551
P-value	<0.0001	<0.0001	<0.0001	<0.0001
ERROR				
Partial SS	20283201.131	8.182052×10^{5}	10187346.519	7197213.439

Chapter 4

RESULTS

The following section elaborates on the results of the study: the predictive data analysis and the reporting tool developed.

4.1 Predictive Data Analysis

The predictive data analysis carried out during the study, is done using R. In the analysis, the model performs multivariate analysis, that is, the effect of each predictor variable on the response variable is observed, while other predictors are kept constant. From the plots generated from the analysis, it is observed that, (1) the models where the patient volume predictor is significant, the response variable, non-compliance, increases at a slower rate than patient volume, (2) the models where the payment year predictor is significant, non-compliance decreases linearly over the payment year. The result in (1) depicts that a larger patient volume leads to a higher EHR compliance and according to result in (2), providers using EHR technology for a longer time period are more compliant. In addition, from the ANOVA results, it is observed that the p-value for volume predictor is significant (0.0021) for higher degree. This indicates a non-linear relationship with the

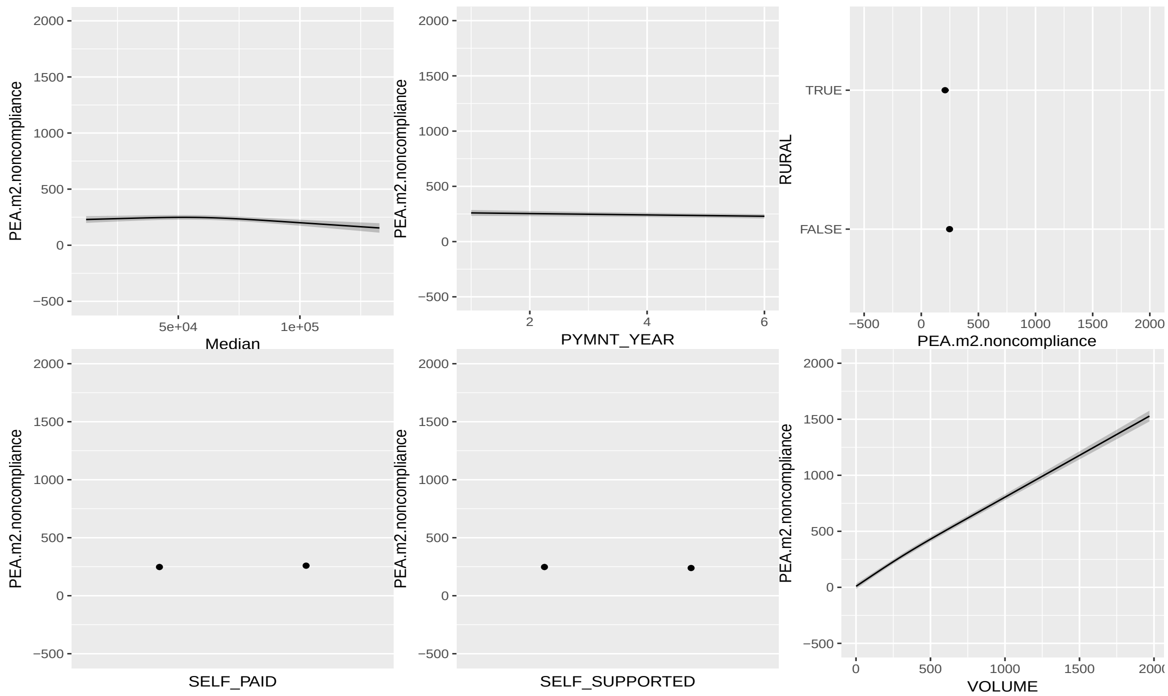

Figure 4.1: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Patient Electronic Access (PEA) measure 2

response variable. The ANOVA results and plots generated for all the models are shown in the Appendix A of the report.

4.2 Use Scenarios

To obtain a better understanding of the features required to automate the reporting task, all possible scenarios were identified (Dennis, 2012). The user scenarios identified are described below.

User Scenario 1: The user wants to generate the reports for measurement domain Meaningful Use or Clinical Quality Measures without application of any data filters. The domain can be selected by the user, for which the report is required.

Table 4.1: Analysis of Variance for noncompliance for Meaningful Use objective Patient Electronic Access (PEA) measure 2

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	107193.815	107193.815	13.66	0.0002
Median	2	182314.299	91157.149	11.61	<0.0001
Nonlinear	1	110543.357	110543.357	14.08	0.0002
PYMNT_YEAR	1	38316.789	38316.789	4.88	0.0274
VOLUME	2	45451888.362	22725944.181	2895.52	<0.0001
Nonlinear	1	74870.561	74870.561	9.54	0.0021
SELF_SUPPORTED	1	4896.148	4896.148	0.62	0.4298
SELF_PAID	1	6649.922	6649.922	0.85	0.3576
TOTAL NONLINEAR	2	197438.046	98719.023	12.58	<0.0001
REGRESSION	8	52418263.551	6552282.944	834.83	<0.0001
ERROR	917	7197213.439	7848.652		

User Scenario 2: The user wants to generate Meaningful Use report after application of filter specific to Meaningful Use. The user is first required to select the domain MU and then select whether the filter is required to be applied to the data. Once filter active, the use can select the filter criteria for the report.

User Scenario 3: The user wants reports with list of providers sorted in order of their rate or performance. The user will first select the measurement domain for which the report is required. According to the domain selected, the measures for the domain must be auto-filled. Then the user is required to select the values such as the sorting criteria, for order of results, the group from the results, the threshold by and the number of results required.

User Scenario 4: The user requires a visual representation of data that is plots: Histograms and pie-charts. The user is required to switch to the plots tab first. Once in the plots tab, the user should select the type of plot required and measure for which the plot is required.

User Scenario 5: The user wants to copy a particular section of the report generated. The user can click on the report and select the region to be copied and pasted to the local machine.

User Scenario 6: The user wants to save the report generated on the local machine. The user can do so in the formats required based on the operating system, the tool is being used.

User Scenario 7: The user wants to have a better view of the text in the report. The user can do so by using the Zoom in and Zoom out feature in the tool.

User Scenario 8: The user wants to scroll down a lengthy report. This can be done using the scroll bar at the side for vertical scrolling or scroll bar at the bottom for horizontal scrolling.

Considering the user scenarios identified, different ideas on how to develop the appropriate user inter- face for the eMIPP reporting tool was discussed.

4.3 Software

The study involves development of a software that automates the task of computing descriptive statistical values for the data reported to the eMIPP system, by the providers. The software, that is, the eMIPP reporting tool is a cross platform software, which works on both Linux and Windows operating systems. The reporting tool uses a TCL/Tk based installer program for the installation of the tool onto the user's system. The installer program provides user with a feature to select the directory, where the tool should be installed. Once the tool is installed successfully in the directory, it is ready to use. The reporting tool generates reports that provide descriptive statistics for the measures that fall under the two measurement domains: Meaningful Use and Clinical Quality Measures.

The software is developed using the TCL scripting language. The graphical user interface is developed using the TCL extension, Tk. The backend used for the tool is a database in SQLite3. SQLite3 is the relational database management system, embedded into the end program. The database is prepared during the data preparation steps, using the text files received from Maryland Department of Health (MDH). The database contains the Medicaid provider attestation data. The TCL/Tk program uses the TCL package sqlite3 for making use of the SQLite queries to compute the descriptive statistics for the reports.

The three main functions of the eMIPP Reporting Tool are providing reports with descriptive statistics for the measures in the two domains, providing plots: pie-charts and histograms that plot Meaningful Use Compliance Rate with the provider count and providing reports where providers are ranked on the basis of their performance and rates. Provider performance and rate are computed using rate and the average normalized rank formulas mentioned in the section 3.2. The tool consists of two tabs, the Report tab and the Plot tab. The reports tab displays the descriptive statistics and the provider reports for the domain as selected by the user and also the reports with providers ranked in order of their performance. The plots tab is used for the graphical representation of the measures in a domain. The tool consists of a multi-select checkbox bar, which allows the user to select year or combination of years for which the report is required. In the reports tab, the reports are generated using report, struct and matrix modules of TCL/Tk. The report module makes use of a template to generate the report outline. The struct and matrix modules assist in generating matrix of calculated values. In case, the user selects Meaningful Use from the domain dropdown and checks the checkbox for MU report filter, a threshold dropdown gets active. This filter allows the user to see descriptive statistics results for providers who are above or below and equal to the threshold value for a given measure. The threshold values for each measure under the two domains is found from the documentation for the measures, provided by the CMS.

In addition to the descriptive statistics reports, as mentioned above, the tool provides user with the reports having providers ranked by their performance score and rates. In the reports tab of the tool, once the user selects the measurement domain, the user can select the ranking criteria from the sort by dropdown. In addition to the sort by dropdown, the tool provides the user with the measure, group, order, threshold by and number dropdowns. The measure dropdown contains the measures that fall under the measurement domain selected initially in the tool. This dropdown is auto filled on the basis of the domain selected. The group dropdown provides user with the option to select the top group of resulting list of providers or the bottom ones. The order dropdown allows user to select the order, that is, whether the providers should be displayed in increasing order of rate/performance or in decreasing order of rate/performance. The threshold by dropdown contains count and performance values, which state that whether the number of providers to be displayed should by count or percentile number. For instance, 10 providers in number or 10 percentile providers out of the providers in the result. The number dropdown takes the value of count or percentile providers to be displayed. The dropdown contains some numbers such as 5, 10, 25 etc. in addition to the custom option. In case, the user selects the custom option, the user is provided with a text box, which takes the custom number required by the user.

The tool enables the user to save the reports generated, in three file formats for a windows system, namely, Excel Spreadsheets (.xlsx), Portable Document Format (.Pdf) and Comma separated files (.CSV) and in the latter two formats on Linux based systems. Packages exist in TCL/Tk, such as pdf4tcl, csv, CAWT and twapi, that enable the tool to perform this functionality. Some of the features have been added to the tool in order to enhance the usability of the tool. The (1) ability to select and copy a desired portion of report, (2) the zoom in and zoom out of the report, (3) a help menu consisting a detailed user guide, are some features of convenience provided by the tool. The data visualization functionality is added to the tool using the plotcharts module in TCL/Tk.

With assistance of different modules of language TCL/Tk, eMIPP Reporting Tool provides the user with automated reporting and data visualization components in a single tool.

Chapter 5

DISCUSSION

The HITECH act of 2009 has significantly increased the adoption and use of health IT by health care providers and hospitals. The incentivization of the use of EHRs by the HITECH Act of 2009 and the Meaningful Use have spurred the adoption of EHRs. As a result, the health sector is making considerable progress in promoting interoperability and digital health information exchange. The adoption of EHR by health care providers enable multiple providers, regardless of location, to simultaneously access patient records from any electronic device. They allow more efficient collaboration on multiple facets of care. The benefits of EHRs have been categorized on the basis clinical, organizational and societal outcomes, by a number of researchers (Menachemi & Collum, 2011). The clinical outcomes such as the data can be quickly shared, results can be better managed with negligible or no errors. Out of the the organizational outcomes, some are financial and operational performance and satisfaction among patients and clinicians. In addition to organizational and clinical outcomes, the benefits of EHRs include improved patient care, increased patient participation, improved diagnostics and patient outcomes.

Despite the positive impact of EHR adoption, there are certain factors that influence the level of adoption of EHRs. Some of the studies elaborate on the hurdles faced during EHR implementation that affect the EHR adoption levels. The usability challenges that make providers unable to process patient information, legal challenges, variance of computer literacy among health care providers, privacy of patient records, cost of EHR systems, rejection of EHRs by patients etc, are some of the mentioned hurdles (Farala Agno & L Guo, 2013). In addition, studies indicate that main reasons for low levels of EHR adoption are due to: legal issues after implementation, faced because of problems such as poor implementation, social barriers caused by influence from other providers, return on investment for providers after investing on implementation and lack of financial support(Palabindala V, 2016).

After going through a number of studies conducted on EHR adoption and use, the factors influencing the adoption levels can be categorized into social, legal, ethical and technological factors. This study elaborates on factors associated with providers. The study takes into account the characteristics of health care providers such as the rurality, whether the provider is self paid and self supported, the payment year and the patient volume. These characteristics are computed using the data entered by the providers themselves into a health care administration system, in order to obtain incentives for their display of meaningful use of EHRs. Using these factors, conclusions are made about the compliance of providers with EHRs. As also mentioned in the results section, the models where the patient volume predictor is significant, non-compliance increases at a smaller rate than volume. This result indicates that the providers working in larger medical setups may have better Meaningful use compliance than providers in smaller medical setups. The analysis also indicates that the number of providers adopting EHRs, increase over the payment year, that is, non- compliance decreases linearly over the payment year where the payment year is significant. This indicates that providers who have adopted EHR technology for a long time may have higher Meaningful Use compliance rate. Therefore, it may not be fair to consider larger and smaller medical set ups together, in the process of computing meaningful use compliance, as a larger medical set up, due to the presence of large patient volume can be a false indicator of being more compliant than the smaller medical set up. The results of the study can assist the decision makers, when it comes to quantifying the EHR adoption levels by providers while providing incentives.

The data collected by the eMIPP health care administration system contains records that are large in number and difficult to handle manually, when it comes to making conclusions. This study involved development of a cross platform tool, eMIPP reporting tool. The tool focuses on automating the process of computing the descriptive statistical values for the measures falling under different measurement domains and providing providers list ranked in order of their performance and rates. The reports provided by the tool assists the health care administration in making decisions when providing incentives to health care providers. The research contributes by development of the tool, as before the tool was developed, the tasks now done using the tool were conducted manually.

Chapter 6

LIMITATIONS

This research performs predictive data analysis on health care data. The research suffers from various threats to validity. Since the research deals with statistical computations, threats associated with Statistical Conclusion Validity or Conclusion Validity can affect the conclusions about relationship among variables. Conclusion validity is solely concerned with whether there exist a relationship or not (Stephanie, 2015d). It does not deal with the specifics about the type of the relationship. Threats to Statistical Conclusion Validity, fall under the categories of (1) fishing, leading to incorrect conclusions demonstrating relationship, when there is no relationship, (2) low statistical power, i.e., incorrect conclusion stating that there is no relationship between your variables, (3) restriction of range that can lead to incorrect estimates, (4) unreliable measures, that can result in under or over estimation in size of relationship among variables.

Another type of validity is Internal Validity (Stephanie, 2015c). This type of validity measures if the research is done right. Internal Validity is concerned with what is responsible for the change in the response or the dependent variable. As it deals with the affects on the response variable, therefore, threats to internal validity are of concern for this research. In addition to the Internal validity, there is External Validity. External validity is associated with if the research can be applied to the real world (Stephanie, 2015a). In case, the research can be applied to real life situations, then external validity is high. The threats to external validity can undermine the application of the research to real world problems. The threats are explanations to how the results can be wrong when applied to other research problems. Threats under this category can be due to presence of external factors. It can be the case, where the results of the study are correct but there exist hidden variables or factors in addition to the independent variables, influencing the results. Population validity and Ecological validity are types of External validity. These answer the questions the questions about how the research on a sample can be applied to the population as a whole and how results are applicable to different setting respectively.

Another type of validity associated with results from data analysis is Replicability. This type of validity exists when the results from a study can be replicated, if research is repeated in the exact same manner (Stephanie, 2015b). The threats to replicability can be due to a slight difference in experimental procedures for the research. The difference in procedure can be due to lack of experience or lack of interest in replication.

Although the results of this study can assist in measuring Meaningful Use compliance and provide incentives to health care providers accordingly, threats to validity of the conclusions exist.

Chapter 7

CONCLUSION

The study performed, makes conclusions compliance of medicaid providers with meaningful use of Electronic Health Records. The analysis conducted during the study produces results that state, that the providers working in large medical set ups or having a longer period of time of using EHR technology, have a better meaningful compliance. The factors such as median income, patient volume, self-paid, self-supported providers, rurality are considered to make the conclusions. From the results of the study, it can be said that it will be unfair to consider large and small size medical set-ups together, when providing incentives. This is because a large medical set up, due to the presence of larger patient volume can be a false indicator of being more compliant than a smaller. In addition to the conclusions made regarding EHR compliance, the eMIPP reporting tool developed during the study, provides its users, reports for the required measurement domains. The reports containing both descriptive statistical values and provider rankings in order of their rate or performance score can be generated using the tool. The tool is capable of handling a large number of data and generate reports in an optimized time period. The users can generate reports using the automated solution, the eMIPP reporting tool and devote more time on analyzing the reports rather than computing the values manually. The development of reporting tool and the analysis conducted during the study, can assist the decision makers in making decisions while providing incentives to the medicaid providers, for their demonstration of meaningful use of EHR technology.

Appendix A

Model 1: Meaningful Use Measure Computerized Order Entry measure 1

Significant predictors: median (indicates median income), payment year, volume

Non significant predictors: rural, self-paid, self-supported

Response variable: non-compliance

Table 1: Analysis of Variance for noncompliance for Meaningful Use Measure Computerized Order Entry measure 1

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	3847.909	3847.909	0.20	0.6520
Median	2	200859.989	100429.994	5.31	0.0051
Nonlinear	1	140925.343	140925.343	7.45	0.0065
PYMNT_YEAR	1	119877.071	119877.071	6.34	0.0120
VOLUME	2	405193.824	202596.912	10.72	<0.0001
Nonlinear	1	112846.623	112846.623	5.97	0.0148
SELF_SUPPORTED	1	16552.592	16552.592	0.88	0.3497
SELF_PAID	1	48822.490	48822.490	2.58	0.1084
TOTAL NONLINEAR	2	315150.180	157575.090	8.34	0.0003
REGRESSION	8	769139.664	96142.458	5.09	<0.0001
ERROR	837	15823287.442	18904.764		

The predictors median and volume indicate a non-linear relationship with the response variable non-compliance.

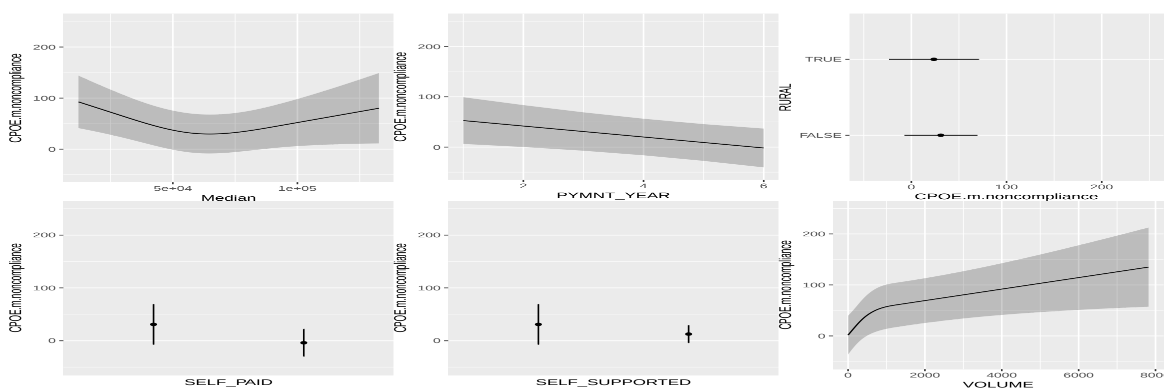

Figure 1: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Computerized Provider Order Entry measure 1

Model 2: Meaningful Use Measure Computerized Order Entry measure 2

Significant Predictors: rural, payment year, volume

Non significant predictors: median, self-paid, self-supported

Response variable: non-compliance

Table 2: Analysis of Variance for noncompliance for Meaningful Use Measure Computerized Order Entry measure 2

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	137070.381	137070.381	7.83	0.0053
Median	2	703283.883	351641.941	20.08	<0.0001
Nonlinear	1	61080.779	61080.779	3.49	0.0622
PYMNT_YEAR	1	121984.892	121984.892	6.97	0.0085
VOLUME	2	1284365.489	642182.744	36.67	<0.0001
Nonlinear	1	441848.884	441848.884	25.23	<0.0001
SELF_SUPPORTED	1	1719.107	1719.107	0.10	0.7541
SELF_PAID	1	19037.778	19037.778	1.09	0.2974
TOTAL NONLINEAR	2	542664.237	271332.119	15.49	<0.0001
REGRESSION	8	2416940.627	302117.578	17.25	<0.0001
ERROR	771	13501646.577	17511.863		

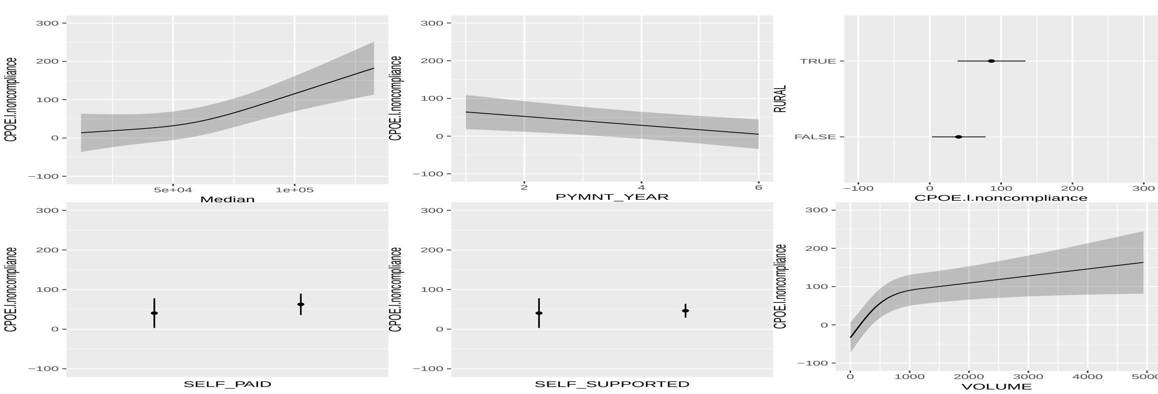

Figure 2: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Computerized Provider Order Entry measure 2

The predictor volume indicates a non-linear relationship with the response variable non-compliance.

Model 3: Meaningful Use Measure Computerized Order Entry measure 3

Significant predictors: median, volume

Non significant predictors: rural, payment year, self-paid, self-supported

Response variable: non-compliance

Table 3: Analysis of Variance for noncompliance for Meaningful Use Measure Computerized Order Entry measure 3

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	385.91243	385.91243	0.78	0.3789
Median	2	10346.21953	5173.10976	10.40	<0.0001
Nonlinear	1	8742.94834	8742.94834	17.58	<0.0001
PYMNT_YEAR	1	989.26533	989.26533	1.99	0.1592
VOLUME	2	79391.07562	39695.53781	79.81	<0.0001
Nonlinear	1	10568.50052	10568.50052	21.25	<0.0001
SELF_SUPPORTED	1	49.92833	49.92833	0.10	0.7515
SELF_PAID	1	204.61377	204.61377	0.41	0.5216
TOTAL NONLINEAR	2	18259.79528	9129.89764	18.36	<0.0001
REGRESSION	8	109773.07767	13721.63471	27.59	<0.0001
ERROR	414	205905.94833	497.35736		

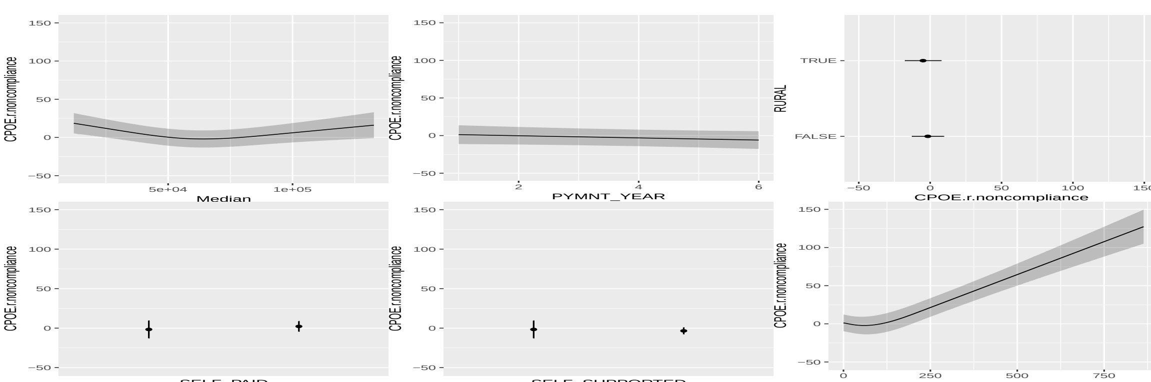

Figure 3: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Computerized Provider Order Entry measure 3

The predictors median and volume indicate a non-linear relationship with the response variable non-compliance.

Model 4: Meaningful Use Measure Electronic Prescribing (eRx)

Significant predictors: median, volume, self-supported

Non significant predictors: rural, payment year, self-paid

Response variable: non-compliance

Table 4: Analysis of Variance for noncompliance for Meaningful Use Measure Electronic Prescribing

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	1.094206×10^4	10942.06296	0.68	0.4093
Median	2	2.318494×10^5	115924.72482	7.22	0.0008
Nonlinear	1	8.849313×10^4	88493.13451	5.51	0.0191
PYMNT_YEAR	1	1.648631×10^4	16486.31078	1.03	0.3111
VOLUME	2	6.630888×10^6	3315444.21046	206.54	<0.0001
Nonlinear	1	1.692771×10^5	169277.13483	10.55	0.0012
SELF_SUPPORTED	1	7.787040×10^4	77870.39777	4.85	0.0279
SELF_PAID	1	7.326377×10^1	73.26377	0.00	0.9462
TOTAL NONLINEAR	2	2.475215×10^5	123760.74881	7.71	0.0005
REGRESSION	8	6.975839×10^6	871979.93568	54.32	<0.0001
ERROR	838	1.345188×10^7	16052.36439		

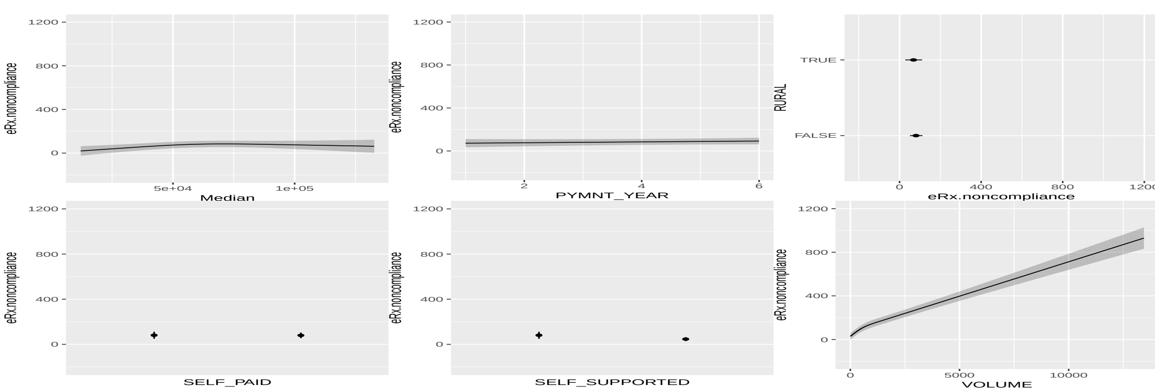

Figure 4: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Electronic Prescribing

The predictors median and volume indicate a non-linear relationship with the response variable non-compliance.

Model 5: Meaningful Use Measure Patient Specific Education

Significant predictors: rural, volume

Non significant predictors: median, payment year, self-paid, self-supported

Response variable: non-compliance

Table 5: Analysis of Variance for noncompliance for Meaningful Use Measure Patient Specific Education

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	98759.179	98759.179	4.53	0.0335
Median	2	264178.502	132089.251	6.06	0.0024
Nonlinear	1	37964.451	37964.451	1.74	0.1871
PYMNT_YEAR	1	13359.784	13359.784	0.61	0.4338
VOLUME	2	5435718.927	2717859.464	124.75	<0.0001
Nonlinear	1	368852.218	368852.218	16.93	<0.0001
SELF_SUPPORTED	1	3011.814	3011.814	0.14	0.7101
SELF_PAID	1	16803.819	16803.819	0.77	0.3800
TOTAL NONLINEAR	2	396945.406	198472.703	9.11	0.0001
REGRESSION	8	5806924.761	725865.595	33.32	<0.0001
ERROR	931	20283201.131	21786.467		

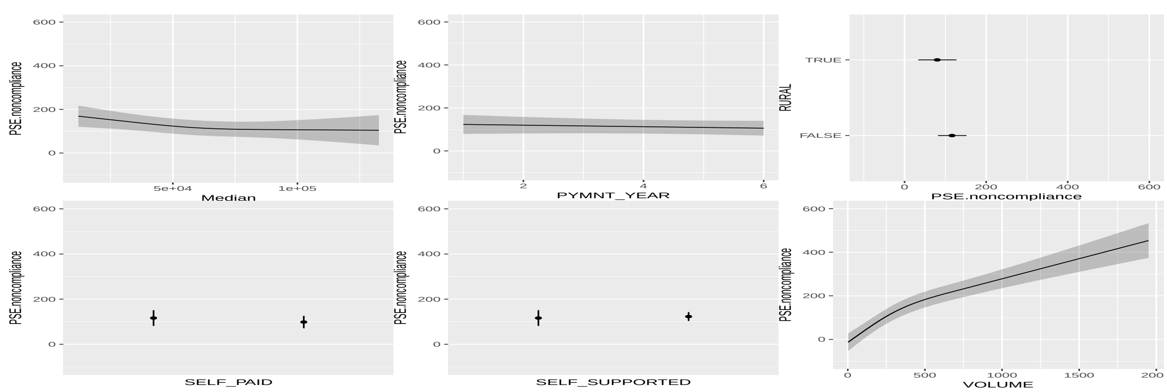

Figure 5: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Patient Specific Education

The predictor volume indicates a non-linear relationship with the response variable non-compliance.

Model 6: Meaningful Use Measure Medication Reconciliation

Significant predictors: median, volume

Non significant predictors: rural, payment year, self-paid, self-supported

Response variable: non-compliance

Table 6: Analysis of Variance for noncompliance for Meaningful Use Measure Medication Reconciliation

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	2.325030×10^{2}	232.5030098	0.25	0.6138
Median	2	9.747959×10^{3}	4873.9795094	5.34	0.0049
Nonlinear	1	4.181561×10^{3}	4181.5609727	4.58	0.0325
PYMNT_YEAR	1	6.690112×10^{-1}	0.6690112	0.00	0.9784
VOLUME	2	1.527786×10^{5}	76389.3120117	83.75	<0.0001
Nonlinear	1	1.385845×10^{4}	13858.4455289	15.19	0.0001
SELF_SUPPORTED	1	2.206477×10^{2}	220.6476837	0.24	0.6230
SELF_PAID	1	1.472910×10^{3}	1472.9097344	1.61	0.2042
TOTAL NONLINEAR	2	2.248295×10^{4}	11241.4725711	12.32	<0.0001
REGRESSION	8	1.959723×10^{5}	24496.5357425	26.86	<0.0001
ERROR	897	8.182052×10^{5}	912.1574059		

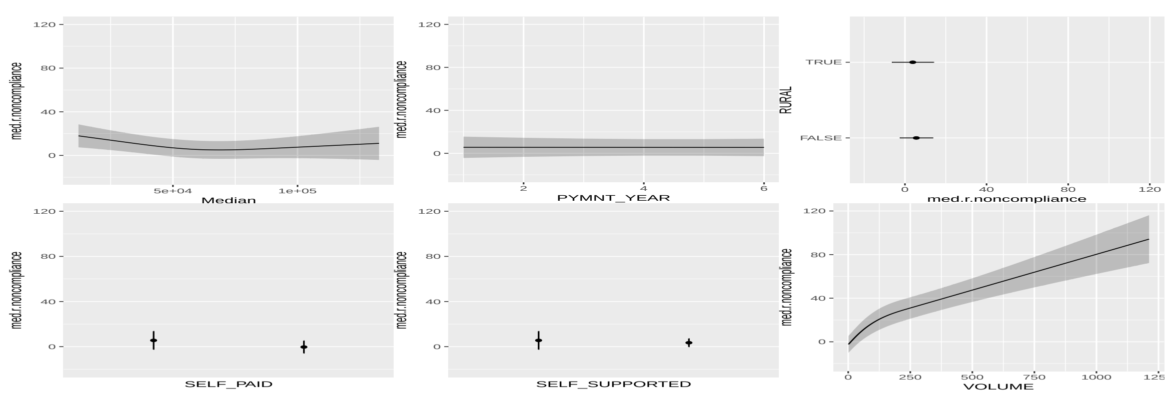

Figure 6: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Medication Reconciliation

The predictor volume indicates a non-linear relationship with the response variable non-compliance.

Model 7: Meaningful Use Measure Patient Electronic Access measure 1

Significant predictors: median, payment year

Non significant predictors: rural, volume, self-paid, self-supported

Response variable: non-compliance

Table 7: Analysis of Variance for noncompliance for Meaningful Use Measure Patient Electronic Access measure 1

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	29824.728	29824.728	3.49	0.0620
Median	2	64809.593	32404.796	3.79	0.0228
Nonlinear	1	37472.251	37472.251	4.38	0.0365
PYMNT_YEAR	1	74626.736	74626.736	8.73	0.0032
VOLUME	2	2521234.208	1260617.104	147.50	<0.0001
Nonlinear	1	1246.998	1246.998	0.15	0.7025
SELF_SUPPORTED	1	2520.822	2520.822	0.29	0.5872
SELF_PAID	1	8125.011	8125.011	0.95	0.3297
TOTAL NONLINEAR	2	38334.162	19167.081	2.24	0.1066
REGRESSION	8	3008902.970	376112.871	44.01	<0.0001
ERROR	1192	10187346.519	8546.432		

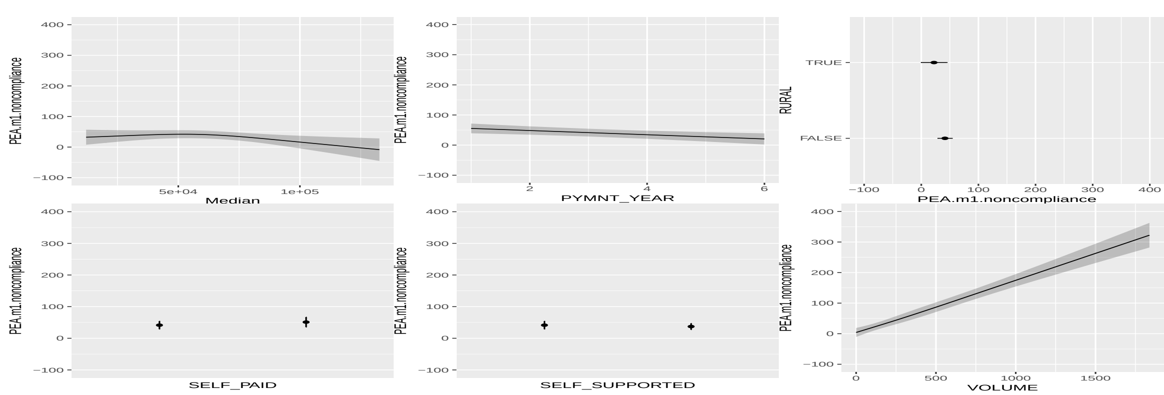

Figure 7: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Patient Electronic Access measure 1

The predictor median indicates a non-linear relationship with the response variable non-compliance.

Model 8: Meaningful Use Measure Patient Electronic Access measure 2

Significant predictors: median, payment year, rural, volume

Non significant predictors: self-paid, self-supported

Response variable: non-compliance

Table 8: Analysis of Variance for noncompliance for Meaningful Use Measure Patient Electronic Access measure 2

	d.f.	Partial SS	MS	*F*	*P*
RURAL	1	107193.815	107193.815	13.66	0.0002
Median	2	182314.299	91157.149	11.61	<0.0001
Nonlinear	1	110543.357	110543.357	14.08	0.0002
PYMNT_YEAR	1	38316.789	38316.789	4.88	0.0274
VOLUME	2	45451888.362	22725944.181	2895.52	<0.0001
Nonlinear	1	74870.561	74870.561	9.54	0.0021
SELF_SUPPORTED	1	4896.148	4896.148	0.62	0.4298
SELF_PAID	1	6649.922	6649.922	0.85	0.3576
TOTAL NONLINEAR	2	197438.046	98719.023	12.58	<0.0001
REGRESSION	8	52418263.551	6552282.944	834.83	<0.0001
ERROR	917	7197213.439	7848.652		

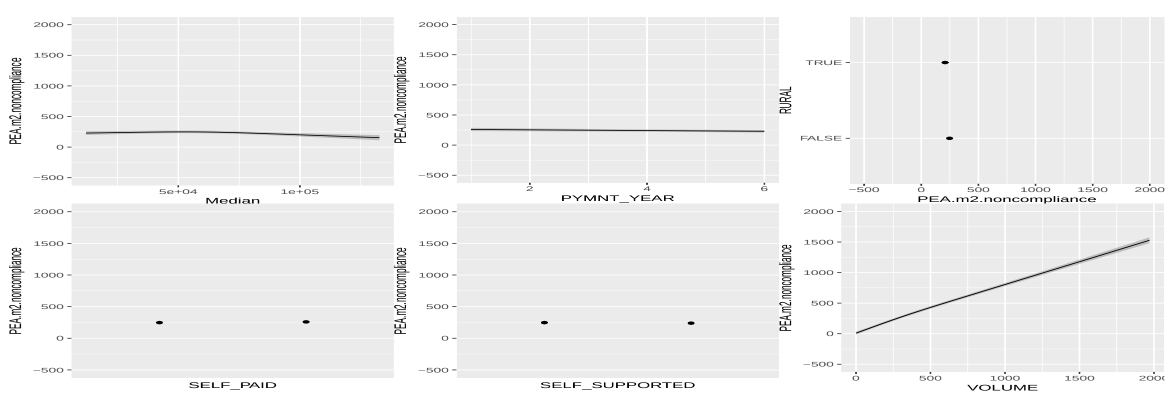

Figure 8: Relationship between each predictor and the response, noncompliance in the fitted model for Meaningful Use objective Patient Electronic Access measure 2

The predictor median and volume indicate a non-linear relationship with the response variable non-compliance.

References

Ajami, S., & Bagheri-Tadi, T. (2013, 05). Barriers for adopting electronic health records (ehrs) by physicians. *Acta informatica medica : AIM : journal of the Society for Medical Informatics of Bosnia Herzegovina : Ä■asopis DruÅątva za medicinsku informatiku BiH*, *21*, 129-34. doi: 10.5455/aim.2013.21.129-134

CMS.gov. (2019). *Promoting interoperability.* https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html?redirect=/EHRincentiveprograms. ((Accessed on 02/08/2019))

Dennis, A. (2012). *Systems analysis and design* (5th ed.). Wiley Publishing.

Farala Agno, C., & L Guo, K. (2013, 07). Electronic health systems: Challenges faced by hospital-based providers. *The health care manager*, *32*, 246-52. doi: 10.1097/HCM.0b013e31829d76a4

Gan, Q., & Cao, Q. (2014, Jan). Adoption of electronic health record system: Multiple theoretical perspectives. In *2014 47th hawaii international conference on system sciences* (p. 2716-2724). doi: 10.1109/HICSS.2014.352

HealthIT.gov. (2018). *Advantages of electronic health records.* https://www.healthit.gov/faq/what-are-advantages-electronic-health-records. ((Accessed on 02/08/2019))

Inoue, A., & Zhou, D. (2016, May). A modern, lean deployment of ehr systems: Two lessons from meaningful use in the united states. In *2016 5th international conference on informatics, electronics and vision (iciev)* (p. 1179-1184). doi: 10.1109/ICIEV.2016.7760186

Menachemi, N., & Collum, T. (2011, 05). Benefits and drawbacks of electronic health record systems. *Risk management and healthcare policy*, *4*, 47-55. doi: 10.2147/RMHP.S12985

Palabindala V, J. N., Pamarthy A. (2016, 10). Adoption of electronic health records and barriers. *Journal of Community Hospital Internal Medicine Perspectives*. doi: 10.3402/jchimp.v6.32643

Piliouras, T., Yu, P. L., Salloum, L., Chavez, J., Suss, R. J., & and. (2016, April). Impacts of meaningful use and what's next: The next generation accountable care organization model. In *2016 ieee long island systems, applications and technology conference (lisat)* (p. 1-6). doi: 10.1109/LISAT.2016.7494110

Stephanie. (2015a). *External validity definition examples.* `https://www.statisticshowto.datasciencecentral.com/external-validity/`. ((Accessed on 03/08/2019))

Stephanie. (2015b). *External validity definition examples.* `https://www.statisticshowto.datasciencecentral.com/external-validity/`. ((Accessed on 03/08/2019))

Stephanie. (2015c). *Internal validity: Definition and examples.* `https://www.statisticshowto.datasciencecentral.com/internal-validity/`. ((Accessed on 03/08/2019))

Stephanie. (2015d). *Statistical conclusion validity.* https://www.statisticshowto.datasciencecentral.com/statistical-conclusion-validity/. ((Accessed on 03/08/2019))