5-11-2023

Evaluating user acceptance of an open-source mobile app for hospital price transparency rule

Saptarshi Purkayastha
Indiana University Purdue University Indianapolis, saptpurk@iupui.edu

Regina Merine
Indiana University Purdue University Indianapolis, remerine@iu.edu

Vyona Dsouza
Indiana University Purdue University Indianapolis, vdsouza@iu.edu

Pallavi Singh
Indiana University Purdue University Indianapolis, singpall@iu.edu

Judy Gichoya
Indiana University Purdue University Indianapolis, judywawira@emory.edu

EVALUATING USER ACCEPTANCE OF AN OPEN-SOURCE MOBILE APP FOR HOSPITAL PRICE TRANSPARENCY RULE

Research Paper

Saptarshi, Purkayastha, Indiana University Purdue University Indianapolis, USA, saptpurk@iupui.edu

Regina, Merine, Indiana University Purdue University Indianapolis, USA, remerine@iu.edu

Vyona, Dsouza, Indiana University Purdue University Indianapolis, USA, vyonadsz@iu.edu

Pallavi, Singh, Indiana University Purdue University Indianapolis, USA, singpall@iu.edu

Judy, Gichoya, Emory University Hospital, USA, judywawira@emory.edu

Abstract

In 2021, the US Center for Medicare and Medicaid Service mandated the Price Transparency Rule, requiring hospitals to publish a patient service price list called Charge Description Master. However, the mandated machine-readable formats made it difficult for patients to understand pricing and limited price transparency. To address this, we developed the LibreHealth Cost of Care Explorer App to provide patients with a user-friendly format of the CDM. We conducted a mixed-methods user study with 55 patients in two large US cities, one in a safety-net hospital and another in a for-profit hospital, and used PLS-SEM path modeling to analyze the app's acceptability using the Unified Theory of Acceptance and Use of Technology constructs. Behavioral Intention and Facilitating Conditions significantly impacted Usage Behavior. Effort Expectancy also had a positive impact. Further explanations for the observed model differences in the two hospital systems were obtained from think-aloud observations and semi-structured interviews.

Keywords: mHealth, price transparency, patient engagement, PLS-SEM

1 Introduction

Over the past decade, healthcare expenditure in the United States has increased at a tremendous annual average despite repeated efforts to reduce costs and improve patient outcomes. National healthcare expenditures grew at an average annual rate of 4.5 percent between 2010 and 2021. According to the Centers for Medicare and Medicaid Services data, total national health expenditures in the United States reached $4.3 trillion or $12,914 per person in 2021, representing a 2.7% increase from the previous year. Private health insurance expenditures increased by 5.8% to $1,211.4 billion, while Medicare spending increased by 8.4% to $900.8 billion, and Medicaid grew by 9.2% to $734.0 billion in 2021. (Hartman et al., 2022). According to international comparisons, diagnostic and therapeutic procedures cost 4-5 times higher in the United States than in Europe or other high-income nations (Bindman, 2020). Medical bills are a significant contributor to bankruptcy. A 2020 study showed that high medical bills were the cause of more than 60 percent of U.S. bankruptcies. That study also suggested that low income and lack of health insurance are the major contributors to medical debt and bankruptcy (Bielenberg et al., 2020). Many Americans must know about health insurance terms, concepts, and costs, including calculating out-of-pocket expenses. According to another 2020 study, the uninsured and underinsured had difficulty covering their medical expenses due to their inability to pay copays, deductibles, or coinsurance (Dean et al., 2020). The uninsured are less likely to have access to and use healthcare services than the insured due to their inability to pay full prices to the hospitals. It is apparent that these costs serve as a criterion for self-paying and cash-paying patients to choose where to receive care, and there is a high likelihood that some patients will delay or avoid receiving medical care due to the high costs (Attafuah et al., 2022). To eliminate these occurrences, the Centers for Medicaid and Medicare Services (CMS) notified the *Price Transparency Rule* that mandated hospitals to make their Charge Description Masters (CDM) publicly accessible in an importable, machine-readable format starting on January 1, 2021, in an effort to promote hospital price transparency. The main goal of this mandate is to lower out-of-pocket costs for uninsured and underinsured patients by enabling them to compare hospital pricing for requested services and make well-informed decisions about receiving medical treatment. In this approach, patients can prepare a budget for their care, request possible aid, and talk to their physician about the costs before beginning desired therapy (A. J. Lu et al., 2020). However, as of June 2022, CMS has issued 352 warning notices and 157 corrective action plan requests to hospitals for not following the rule correctly (Kona and Corlette, 2022), and only about 6% of hospitals have published the price list (Haque et al., 2022). Although the prices are sometimes displayed on hospital websites, several obstacles and restrictions prevent users from quickly accessing and comparing the information (Arvisais-Anhalt et al., 2021).

Challenges accessing Chargemasters

- **Informational Deficiency**: For instance, in most Dallas County hospital descriptions, a procedure was referred to as "circumcision." However, a huge price variation (252.00-7532.10 dollars) was found and reflected the distinction between an uncomplicated newborn circumcision performed in the newborn hospital and an older kid or adult with phimosis circumcised in the operating room while under anesthesia. These would indicate two distinct processes, but without more details, a consumer will not recognize the difference between the chargemasters (Arvisais-Anhalt et al., 2021).
- **Price Variations**: Hospital costs might differ for both simple and complicated treatments. However, CDM does not provide different prices for either. Researchers also conclude that costs may vary from those stated on websites and are occasionally negotiable. Likely, the listed rates will only apply to the initial cost of the services or procedures (Evashwick-Rogler et al., 2022).
- **Technical Complexity/ Lack of consumer friendliness**: The only way to access some charge description masters is by utilizing a "targeted query" in the search space. Web links are dispersed throughout the text, making it difficult to identify them, and most of the offered are inaccurate or broken (Linde and Egede, 2022).
- **Complex and varied terminology**: One of the frequent issues is that the treatments were listed in medical terminology and needed more basic terminology that a patient with minimal health literacy might comprehend. Researchers discovered that few hospitals had listed CPT codes for the services offered at their hospitals and that most standard hospital services were not searchable on the web because they were not listed with common terminology or had comprehensive acronyms (Reddy et al., 2022).
- **Others**: While estimating costs, most searches for price details needed a subscription to view the prices online (Kratka et al., 2018). The compliance report says that the price comparison tools are complex, and the critical sections of the website are hard to navigate. A few websites also asked for patients' personal information to reveal prices, neglecting the CMS protocol. For underinsured/partially insured patients, an ID is required, which is plan specific/insurer specific to find negotiated prices (Toci et al., 2022).

After conducting a thorough analysis of these obstacles to price transparency, we developed the Libre-Health Cost of Care Explorer app, an innovative, user-friendly, open-source mobile application that makes it easier for patients to compare the prices for their healthcare services. We created a web scraping backend that would go through a list of Medicare and Medicaid reimbursed hospitals' web pages to find price lists and CDMs every week and update the web service accessed by the mobile app.

Need for user-friendly platform
We built the app with a variety of features that can assist consumers in price discovery and compare service quality. As far as we know, this is the first mobile app with the functionality to show the CDM of all hospitals in the United States in a consumer-friendly way. On November 2, 2021, the CMS published the final rule, imposing fines of $500 to $5,500 per day on US hospitals that do not disclose their prices. By January 1, 2022, this would undoubtedly force hospitals to take action, but that has not been the case. It is unlikely that hospitals and health systems will try to develop comparable apps, given the intense competition prevalent in many urban locations. Since we won't see any immediate gains in patient outcomes, research groups will unlikely create apps like these. As a result, open-source groups, foundations, and the federal and state governments will likely have to take the lead in modifying and maintaining such apps.

2 Theoretical Framework and Hypothesis:

UTAUT Model: Our study utilized the Unified Theory of Acceptance and Use of Technology (UTAUT) constructs of user acceptability and behavior analysis. The UTAUT model predicts user acceptance of a specific technology by understanding the factors influencing usability. It aims to increase user acceptance through technology design. UTAUT has been extensively used in various domains, including mHealth and eHealth, to gain a uniform view of user acceptance. After reviewing earlier usability studies of mHealth apps using UTAUT, we selected the most widely employed constructs. The latent variables used in our conceptual model are Performance Expectancy (PE), Effort Expectancy (EE), Facilitating Conditions (FC), Behavioral Intension (BI), and Usage Behavior (UB). According to the hypothesis, PE and EE have a key role as direct predictors of user intention to adopt mHealth technologies, and FC and BI explain the actual use behavior of technology.

- **Performance expectancy (PE)**: PE is defined as "the degree to which an individual believes that using the system will help him or her to attain gains in job performance." A study conducted in 2012 concluded that PE is the strongest predictor of a user's behavioral intention (BI) to use and adopt the technology (Venkatesh, Thong, and Xu, 2012). One of the studies conducted in 2013, provided empirical evidence that the use of mobile health services increased with PE (Sun et al., 2013).
- **Effort Expectancy (EE)**: EE is defined as "the degree of ease associated with the use of the system." Studies from the past indicate that EE has a significant impact on users' intentions to embrace and accept mHealth applications. For example, a study conducted in 2019 revealed that EE positively impacts the use of the Physical Activity app among university students (Liu et al., 2019). Smartphone users' adoption of clinical decision support systems, mobile health, mobile health monitoring systems, and e-Health services is significantly influenced by EE.
- **Facilitating Conditions (FC)**: FC is defined as "the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system". A Research study revealed that FC has a favorable direct impact on behavioral intentions to use Health Information Systems (HIS)(Bawack and Kamdjoug, 2018). Another study confirmed that FC is a positive determinant of patients' acceptance and use of an Emergency Department (ED) wait-times website(Jewer, 2018).
- **Behavioral Intention (BI) and Usage Behavior (UB)**: Multiple research domains have extensively established the connection between Behavioral Intention (BI) and Usage Behavior (UB), proving that BI is a reliable predictor of UB. A study conducted in 2017 found that the relationship between BI and UB was significant in determining the actual use of the mobile phone-based Interactive Voice Response (mIVR) (Brinkel et al., 2017). There was empirical evidence that BI describes users' actual UB of technology in a study conducted in 2003 (Venkatesh, Morris, et al., 2003)

2.1 Methodology:

Study Design: This study followed a mixed-methods *Sequential Explanatory Design*. According to a research study, a sequential explanatory mixed methods design has two phases(Ivankova, Creswell, and Stick, 2006). The first phase includes gathering quantitative data to identify a phenomenon and then explaining and supporting the relationships using qualitative data (Subedi, 2016). In this study, the quantitative data consists of an electronic survey, while the explanatory qualitative data was collected through semi-structured interviews, screen recording observations, and think-aloud observations. Results of both phases were analyzed together, providing a broad overview through the quantitative data while allowing specific findings to be elaborated qualitatively.

Instrument Development: This study used two main instruments. A semi-structured interview guide and a 20-question survey with seventeen 3-point Likert scale questions were implemented in Qualtrics, a cloud-based survey tool. The survey collected Unified Theory of Acceptance and Use of Technology (UTAUT) constructs. The survey questionnaire was pretested for its validity and completion rate through a pilot study with the help of two research assistants from the Department of BioHealth Informatics at Indiana University (IU). The survey and the protocol were then approved by the IU Institutional Review Board (#10984) in the expedited categories (6 and 7), confirming participants' privacy and confidentiality. Semi-structured interview with simultaneous think-aloud app use, followed by the electronic survey, was designed to be completed within 15-30 minutes.

Data Collection The data collection was carried out at the Radiology department of Emory University Hospital Midtown (Hospital A), a large for-profit hospital in Atlanta, and at Sidney & Lois Eskenazi Hospital (Hospital B), a large safety-net hospital in Indianapolis. The study population comprised insured, uninsured, or underinsured populations who visited the hospital for radiology procedures or purchased medical supplies. Participants under the age of 18 and those unable to read and understand the English language were not allowed to participate in the study. The researchers explained the study procedures to each participant and received informed consent.

Participants: We recruited 27 subjects from Hospital A, 17 female, and 9 male, with 1 subject who did not want to be identified. There were 22 insured, 2 underinsured, and 3 uninsured patients. The average participant used the app and completed the survey in about 16-18 minutes, whereas the longest participant use duration was 42 minutes. We recruited 28 participants from Hospital B, 21 women, and 7 men. The majority of participants (n=16) had Medicaid or Medicare insurance. A few had coverage under the Healthy Indiana Plan (n=4). Three participants had UHG (UnitedHealth Group) insurance coverage. Two of them had unknown private insurance, and 3 of the participants had MD Healthwise insurance, managed health services insurance plan, and veterans' affairs insurance plan, respectively. The longest time participants took to use the app and complete the survey was 32 minutes, the shortest time taken was around 5 minutes, and the average time taken was around 12-15 minutes.

Data Analysis: We retrieved the survey results from Qualtrics, with preprocessing done using Python. We eliminated the metadata fields that were not necessary for the PLS-SEM modeling, such as start and end dates and duration. We linked all responses with the think-aloud recordings using the random ID generated by Qualtrics. Out of the 20 questions, we utilized questions from 4 to 20 for data analysis. These questions were then categorized according to the UTAUT model and included the latent variables and labels. Survey questions categorized by the construct are available as an appendix. Interview and think-aloud audio recordings were manually transcribed, and two researchers performed open coding using ATLAS.ti. Codes were exchanged after open coding the 3 same interviews to establish consensus. Finally, inter-coder reliability was calculated between the two researchers with a final 55 interviews with 83% match between the coding.
We utilized Partial Least Squares (PLS) Structured Equation Modeling (SEM) to investigate the relationship between constructs. PLS-SEM helps to model and estimate the cause-effect relationship between the

latent and observed variables. PLS-SEM helps measure the relationship between the latent variables and determines how efficiently the model explains the target constructs of interest. Using PLS-SEM enabled us to measure the model's predictive potential and to judge the quality of the model (Hair Jr et al., 2014). We evaluated the measurement model to test the reliability and validity of the research instrument using indicator reliability and internal consistency reliability metrics. Meanwhile, the validity test assessed both convergent and discriminant validity (Hair Jr et al., 2014). Once we established the reliability and validity of the outer models, we took a brute-force approach to evaluate the hypothesized relationships of the structural model. Every construct was connected to every other construct, and the model's path coefficients were compared. We eliminated weak relationships, and the final model is presented in the paper's results section. The model's quality is assessed based on its ability to predict endogenous constructs accurately. Path Coefficient Correlation and Coefficient of determination (R^2) are used to assess the structural model. Path coefficients represent the proposed connections between different constructs. These path coefficients are standardized on a scale that ranges from -1 to +1. Path coefficients closer to +1 indicate a strong positive relationship, while those closer to -1 indicate a strong negative relationship. Essentially, path coefficients show the strength and direction of the relationships between different constructs in the model. The degree of predictive precision a model exhibits is measured by the Coefficient of determination (R^2). Alternatively, it also indicates the collective impact of the exogenous variable on the endogenous variable(s). The scale of this impact ranges between 0 to 1, where 1 implies complete accuracy in prediction (Hair Jr et al., 2014). We used SmartPLS v4 to perform PLS-SEM modeling with settings listed in the appendix for all instances of Bootstrap and PLS.

3 Results

The following three sections report results. The first and second sections present the survey analysis (both structural and measurement models). The third section details the study hypothesis and findings.

3.1 Analysis of the Measurement Model

Construct	Indicators	Outer loadings	Composite Reliability	Average Variance Extracted (AVE)
Behavioral Intention	BI1	0.87	0.803	0.581
	BI2	0.591		
	BI3	0.799		
Effort Expectancy	EE1	0.899	0.792	0.659
	EE4	0.714		
Facilitating Condition	FC1	0.884	0.703	0.554
	FC2	0.571		
Performance Expectancy	PE1	0.875	0.746	0.511
	PE3	0.751		
	PE4	0.451		
Usage Behavior	UB1	0.783	0.835	0.717
	UB2	0.906		

Table 1. PLS Construct Reliability and Validity Report for Hospital A Model

Indicator Reliability: Indicators with an outer loading of 0.7 are considered a good indication of indicator reliability (Hair Jr et al., 2014). According to Hair et al.(2014), an indicator with an outer loading value between 0.4 and 0.7 should only be deleted, if its deletion increases the construct's composite reliability and average variance extracted (AVE). Any indicators with an outer loading value of less than 0.4 must be eliminated (Hair Jr et al., 2014). Both models had 17 indicators representing 5 latent constructs in the initial stages. After the 1st run of the PLS-SEM Model A, obtained from the survey data of Hospital A, indicators EE3, PE2, and FC3 were deleted as they had an outer loading of less than 0.4. EE2 and

UB3, with an outer loading of 0.617 and 0.591, respectively, were deleted as their deletion improved their construct's AVE and composite reliability. The outer loadings of all the remaining indicators for this model are given in Table 1. For the PLS-SEM Model B obtained from the survey data of Hospital B, indicator PE4 was deleted as it had an outer loading of less than 0.4. Indicators PE1 and FC2, with an outer loading of 0.625 and 0.423, respectively, were deleted to improve their construct's composite reliability and AVE. The outer loadings of all the remaining indicators for this model are given in Table 2.

Internal consistency Reliability: The construct measurement's internal consistency reliability is assessed using composite Reliability. Since the composite reliability of all indicators in both the models is greater than the recommended 0.70 (Hair Jr et al., 2014), the construct measurement's internal consistency reliability is confirmed. Table 1 and Table 2 contain the composite reliability of all the constructs for Model A and Model B, respectively.

Construct	Indicators	Outer loadings	Composite Reliability	Average Variance Extracted (AVE)
Behavioral Intention	BI1 BI2 BI3	0.755 0.672 0.761	0.774	0.534
Effort Expectancy	EE1 EE2 EE3 EE4	0.771 0.758 0.645 0.81	0.835	0.561
Facilitating Condition	FC1 FC3	0.94 0.596	0.756	0.619
Performance Expectancy	PE2 PE3	0.856 0.917	0.88	0.786
Usage Behavior	UB2 UB3	0.829 0.884	0.847	0.734

Table 2. PLS Construct Reliability and Validity Report for Hospital B Model

Convergent Validity: An average variance extracted (AVE) value greater than 0.5 confirms convergent validity (Hair Jr et al., 2014). Since the AVE values of all the constructs in both models are above 0.5, convergent validity is established.

Multicollinearity assessment: Before assessing the structural model, both models were tested for collinearity. There was no collinearity, as the Variance Inflation Factor scores of the constructs in both models were less than 1.5.

	Behavioral Intention	Effort Expectancy	Facilitating Condition	Performance Expectancy	Usage Behavior
Behavioral Intention	0.762				
Effort Expectancy	0.499	0.812			
Facilitating Condition	0.691	0.319	0.744		
Performance Expectancy	0.674	0.649	0.381	0.715	
Usage Behavior	0.811	0.507	0.508	0.541	0.847

Table 3. Fornell-Larcker Criterion for Hospital A Model

Discriminant Validity: Cross loadings and the Fornell-Larcker criterion are used to establish the discriminant validity of the constructs (Hair Jr et al., 2014). The discriminant validity is confirmed since the cross-loading value of each indicator in its own construct is higher than the cross-loading value for other constructs in both models (Hair Jr et al., 2014). Table 3 and Table 4 represent the Fornell-Larcker criterion for Hospital A Model and Hospital B Model, respectively.

	Behavioral Intention	Effort Expectancy	Facilitating Conditions	Performance Expectancy	Usage Behavior
Behavioral Intention	0.73				
Effort Expectancy	0.67	0.749			
Facilitating Conditions	0.197	0.172	0.787		
Performance Expectancy	0.101	0.295	0.468	0.887	
Usage Behavior	0.551	0.497	0.617	0.427	0.857

Table 4. Fornell-Larcker Criterion for Hospital B Model

3.2 Analysis of the structural model

The preliminary models were built based on the hypothesized relationships of the UTAUT framework; PE → BI, PE→UB, FC→BI, FC→UB, EE→BI, EE→UB, BI→UB. Path loadings were sorted to find the lowest loading, and then the smallest path was removed before rerunning the bootstrapping. Once the weakest paths were trimmed, we checked for alternate paths outside the standard UTAUT model for relevant loadings; PE→EE, PE→FC, FC→EE, FC→PE, EE→PE, EE→FC.

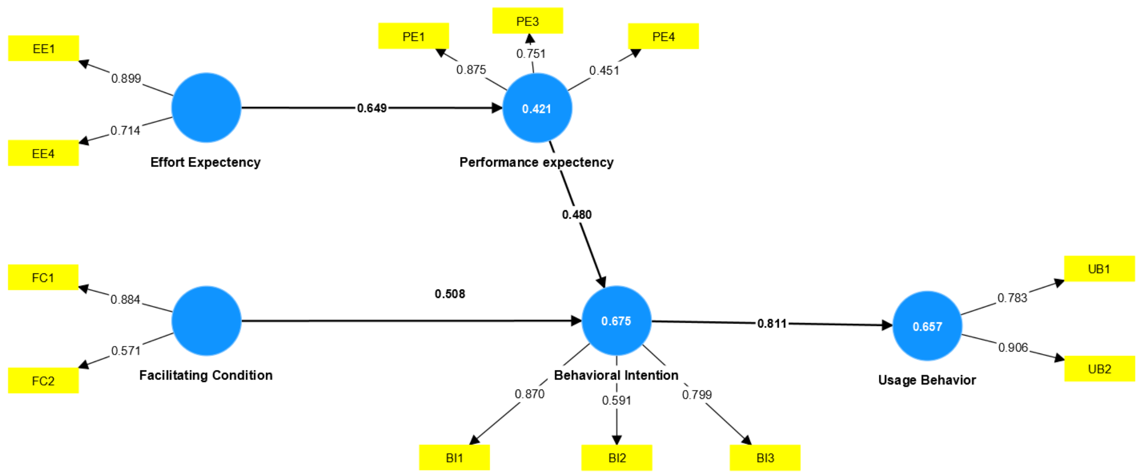

Figure 1. PLS-SEM Hospital A Model obtained from SmartPLS v4

The preliminary PLS-SEM model of the survey data from Hospital A indicated that the strongest relationship was between BI→UB. Additionally, the models showed clear relationships between PE→BI, and FC→BI. Weak relationships were seen between EE→BI, EE→UB, PE→UB, and FC→UB. These paths were trimmed from the model. From the alternate paths, we kept only EE→PE as it had the strongest connection. The final structural model that was obtained is represented in Figure 1. Table 5 contains the path coefficient report for Hospital Model A.

	Path Coefficient	Sample mean	Standard deviation	T -statistics	P-values
Behavioral Intention -> Usage Behavior	0.811	0.794	0.147	5.514	0.001
Effort Expectancy -> Performance Expectancy	0.649	0.673	0.183	3.54	0.001
Facilitating Condition -> Behavioral Intention	0.508	0.542	0.174	2.921	0.004
Performance expectancy -> Behavioral Intention	0.48	0.438	0.176	2.729	0.006

Table 5. Path coefficient, mean, standard deviation, t-values, p-values for Hospital A Model

The preliminary PLS-SEM model of the survey data from Hospital B, indicated that the strongest relationship was between EE→BI. Additionally, the models showed clear relationships between FC→UB, and BI→UB. A potential weak relationship was seen between PE→UB. Relationships between EE→UB, PE→BI, and FC→BI were trimmed from the model as they had very low path loadings and very weak relationships. The model did not include alternate paths other than UTAUT as they did not exhibit a strong connection. Figure 2 represents the final structural model for the Hospital B survey data. Table 6 contains the path coefficient report for Hospital Model B.

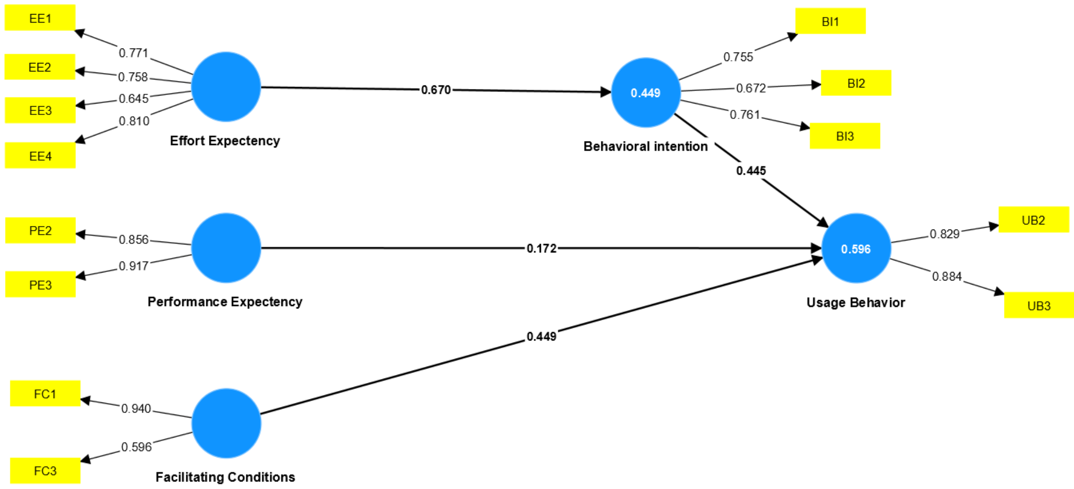

Figure 2. PLS-SEM Hospital B Model obtained from SmartPLS v4

	Original sample	Sample mean	Standard deviation	T-statistics	P-values
Behavioral intention - Usage Behavior	0.445	0.459	0.14	3.185	0.001
Effort Expectancy - Behavioral Intention	0.67	0.71	0.172	3.885	0.001
Facilitating Conditions - Usage Behavior	0.449	0.41	0.183	2.454	0.014
Performance Expectancy - Usage Behavior	0.172	0.201	0.193	0.891	0.373

Table 6. Path coefficients - mean, standard deviation, t-statistic, p-values of Hospital Model B

Coefficient of determination: The model's predictive accuracy is measured by the coefficient of determination (R^2) (Hair Jr et al., 2014). Based on the R^2 values of Performance expectancy (0.421), Behavioral Intention (0.675), and Usage Behavior (0.657), Hospital Model A had moderate predictive accuracy (Hair Jr et al., 2014). Hospital Model B also had moderate predictive accuracy based on the R^2 values for Behavioral Intention (0.449) and Usage Behavior (0.556) (Hair Jr et al., 2014).

3.3 Hypothesis from the survey data obtained from Hospital A and Hospital B

H1: Behavioral intention affects by Usage Behavior: For Hospital A Model, the Hypothesized path of Behavioral Intentions and Usage Behavior had a path coefficient of 0.811 and was statistically significant (Table 5). The Behavioral intention construct had high reliability and validity (Table 1). The hypothesis was also supported by Hospital B Model. Behavioral intention and Usage Behavior had a path coefficient of 0.445 and were statistically significant (Table 6). The behavioral intention construct of the Hospital B Model also had high reliability and validity (Table 2). Therefore, H1 was supported by models of both hospitals. The behavioral intention had positively and significantly affected Usage behavior.

H2: Effort Expectancy affects Performance Expectancy: For Hospital A Model, the hypothesized path of Effort Expectancy and Performance Expectancy had a path coefficient of 0.649 and was statistically significant (Table 5). The Effort Expectancy construct also had high reliability and validity (Table 1). H2 was supported by Hospital Model A, and Effort Expectancy had a positive relationship with Performance Expectancy. For the Hospital B model, there was no relationship between Effort Expectancy and Performance Expectancy

H3: Facilitating Conditions is affected Behavioral Intention: The hypothesized path of Facilitating conditions and Behavioral Intention had a statistically significant path coefficient of 0.508 in Hospital A Model(Table 5). Facilitating conditions construct had high reliability and validity (Table 1). H3 was supported by Hospital A Model, and Facilitating Condition positively affected Behavioral Intention. No relationship was observed between the two constructs in the case of the Hospital B Model.

H4: Performance Expectancy affects Behavioral intention: In the case of the Hospital A Model, the hypothesized path of Performance Expectancy and Behavioral intention had a statistically significant path coefficient of 0.480 (Table 5). The Performance Expectancy construct also had high validity and reliability (Table 1). H4 was supported by Hospital A Model. There was a positive relationship between Performance expectancy and Behavioral intention. In the case of the Hospital B model, there was no relationship between Effort Expectancy and Performance Expectancy.

H5: Effort Expectancy affects Behavioral Intention: In Hospital B Model, the hypothesized path of Effort Expectancy and Behavioral Intention had a statistically significant path coefficient of 0.670 (Table 6). The Effort Expectancy construct had high reliability and validity (Table 2). H5 is supported by Hospital B Model, and Effort Expectancy has a significantly positive relationship with Behavioral Intention. No relationship was observed between the two constructs in the case of the Hospital A Model.

H6: Facilitating Conditions affects Usage Behavior: The hypothesized path of Facilitating conditions and Usage Behavior had a statistically significant path coefficient of 0.449 for Hospital B Model (Table 6). Facilitating conditions construct had high reliability and validity (Table 2). H6 was supported by Hospital B Model, and Facilitating Condition had a positive relationship with Usage Behavior. In the case of the Hospital A model, there was no relationship between Facilitating conditions and Usage Behavior.

H7:Performance Expectancy affects Usage Behavior: For Hospital B Model, the hypothesized path of Performance Expectancy and Usage Behavior had a path coefficient of 0.172 and was statistically insignificant (Table 6). The null hypothesis is rejected as there is no positive relationship between Performance expectancy and Behavioral intention. No relationship was observed between the two constructs in the case of the Hospital A Model.

3.4 Explanations through qualitative analysis

The qualitative analysis had a high inter-coder agreement, likely for two reasons. Firstly, the linked think-aloud and semi-structured interviews received semantically similar discussions regarding the app across the two sites and patient groups, even though the interviewers differed at both study sites. Second, the two coders had PharmD educational training and participated in the pilots of the instruments from the beginning.

Six codes, two each from three themes, were identified as different between the two sites. Two themes were related to pricing transparency (1.) procedure lookup by price and (2.) filtering procedures by price, and one theme was related to (3.) quality of service. Participants at Hospital B focused on the pricing parts of the app (right side of Figure 3) and themes related to price filtering and price search, as highlighted by the following quotes:

> *S6: The most important feature of the app would be to be able to find procedures such as [dialysis] or by diagnosis like [breast cancer] for mammography and compare or search the price at nearby hospitals.*

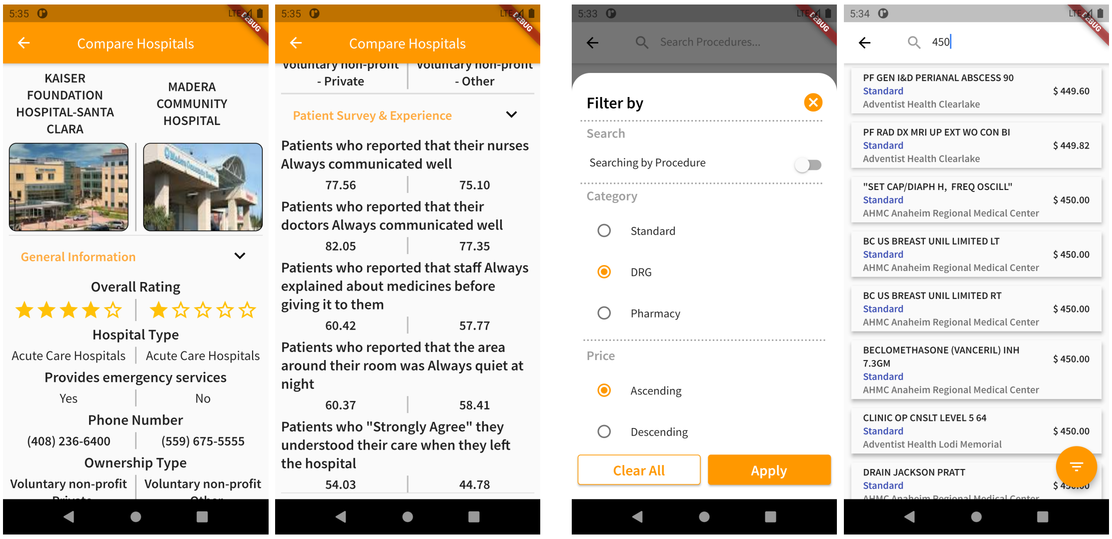

Figure 3. Left: Hospital A - Quality of Service. Right: Hospital B - Price filtering

> *S6: Arranging the search results for services by Ascending order of pricing was the best feature of the app*

Or another quote from an uninsured patient in Hospital B highlights that extra effort to locate and download multiple hospital CDMs and compare them was acceptable:

> *S9: Instead of searching by procedure, [if I] can I get the entire pricelist and then search by a dollar value and not the procedure?*
>
> *S9: Even if I have to click and download a few nearby hospital chargemasters, I would like to use that to compare the price between hospitals for [mammography]*

Thus, as we see in the Hospital B model, Effort Expectancy affecting Performance Expectancy did not impact the users' behavioral intention to use the app until other factors, such as performance expectation or Behavioral Intention, were high. More so, Facilitating Conditions played a limited role in Hospital B participants compared to Hospital A participants. Facilitating Conditions was highly important for predicting Usage Behavior in the Hospital A model. Facilitating Conditions included questions such as: FC1 = "The app allows comparing between hospitals effectively based on its general information and patient experiences", or
FC2 = "I want to get information on the evidence-based treatment options/medical procedures obtained from a registered medical practitioner for the medical condition given in the CDM of the app."
A few quotes presented in Theme 3 (Quality of Service - participants focused on the Left side of Figure 3 of the app) are as follows:

> *S31: The distance search and then facility comparison by the rating is an important feature. I wonder where this patient satisfaction information is coming from into your app?*
>
> *S39: The overall rating, particularly when compared to the national average for [this] facility, was wonderful. It is an important feature that will make me use this app*
>
> *S44: While price comparison is important, I want to see if my insurance company will cover the service at one hospital or the other hospital. ([V1: this is out of the app's scope]) S44: but that's what matters to me instead of the hospital's pricelist*

These three differentiated themes between the sites help explain why the PLS-SEM models demonstrate strong or weak path coefficients between the UTAUT constructs and help explain the differences between patients who come to safety-net or for-profit hospital systems.

4 Discussion

The findings of this study highlight the factors that affect a patient's Usage Behavior in utilizing the newly developed open-source mobile application that provides patient-friendly access to charge masters. The findings indicate that Behavioral Intention significantly improved Usage Behavior. These results aligned with earlier research performed by Efiloğlu Kurt and Tingöy, 2017, which demonstrated higher intention to use due to its user-friendly interface and perceived ease of use. While user-friendly UI was necessary for the intention to use the app, it had a similar impact on participants at the safety-net and for-profit hospitals. However, Effort Expectancy is mediated through Behavioral Intention for Usage Behavior in the Hospital B model, whereas Effort Expectancy is mediated through Performance Expectancy in the Hospital B model.

Our theoretical model has multiple practical uses. Firstly, the models can be used as a lens to explain what we observed about the limited use of CDM by patients since 2021. The goal of the *Price Transparency Rule* was to be able to use market forces for price discovery and allow patients to make informed choices of where to receive services based on pricing. However, many fully insured patients whose services are covered by payors have a limited concerns regarding access to CDMs. They will unlikely try to discover prices at the health system where they regularly visit or compare their hospital to other hospitals based on pricing.

Secondly, our model can also be used to predict the health systems and patient population that will benefit the most from this app and access to CDMs. As Value-Based Programs (Hirsch et al., 2017) become the primary way to reimburse providers, it might be important for the CMS to reflect on the impact of the *Price Transparency Rule*. Weak sanctions on the hospitals that are not following the rule is diluting the goals of the rule. If market forces have to work to reduce prices, competitive pricing can only work if hospitals know that patients will make a choice on the pricing (Miller, Stearns, and Berwick, 2022).

5 Limitations

Certain limitations to this study need to be considered when interpreting the findings. Firstly, the level of education was not collected during the survey, which limits our ability to investigate these variables' potential influence on the app's acceptance. We collected age in the survey but didn't get an adequate response rate at the safety-net hospital. Future studies should aim to collect additional demographic data to investigate the potential influence of these factors on user behavior. Secondly, the data collected was from a small subset of the population, limiting our findings' generalizability. Our conclusions about the app's effectiveness for the safety-net or uninsured patients may not be valid for all kinds of treatments. Finally, the study relied on self-reported data subject to biases. Participants may have overestimated or underestimated their actual usage of the app, which could affect the validity of our conclusions. Future research should aim to replicate our findings in larger, more diverse populations, investigate the potential influence of additional variables, and explore the app's effectiveness for other types of treatments and patient populations.

6 Conclusion

Our study demonstrates a clear difference in patient opinion based on the hospital system they visit, resulting in a different acceptance model for apps that highlight pricing transparency. Hospital A, with a for-profit focus, catered to adequately insured patients who did not have out-of-pocket expenses. These patients cared less for pricing comparisons, and transparency mattered less to them than the quality of service. Instead, patients who visited the safety-net Hospital B were likely uninsured or underinsured. Price transparency truly mattered to them, and other Facilitating Conditions in the app affected Usage Behavior. Thus, the CMS *Price Transparency Rule* and our open-source app are more impactful for safety-net or uninsured patients.

7 Appendix

Q#	Question	Latent variable	Label
4	The app was easy to use.	Behavioral intention to use the system	BI1
5	It was easy to navigate within the app.	Behavioral intention to use the system	BI2
6	The app has a user-friendly interface.	Behavioral intention to use the system	BI3
7	This app has the functions and capabilities to manage the cost of care effectively.	Performance expectancy	PE1
8	I am satisfied with the ease of completing the tasks within the app.	Effort expectancy	EE1
9	I have no difficulty entering, searching, or comparing hospitals, procedures, and prices within the app.	Effort expectancy	EE2
10	I could quickly find all the nearby hospitals by entering the location and customizing the radius.	Effort expectancy	EE3
11	The app provides enough medical procedure list categories and their prices for each hospital and allows a quick comparison of medical procedure costs.	Performance expectancy	PE2
12	I am satisfied with the amount of time it took to complete the overall tasks of downloading each hospital's CDM and comparing medical procedures' prices.	Usage behavior	UB1
13	I could quickly sort the prices in ascending or descending order from the filter options.	Usage behavior	UB2
14	I can use the app without written instructions/training.	Usage behavior	UB3
15	The information in the app is well-organized, so I could easily find the information I needed.	Effort expectancy	EE4
16	The app allows comparing between hospitals effectively based on its general information and patient experiences.	Facilitating conditions	FC1
17	I do not notice any inconsistencies as I move between the home screen, download CDM, main menu, and settings.	Performance expectancy	PE3
18	Information in the app is poorly presented or is subject to misinterpretation.	Performance expectancy	PE4
19	I want to get information on the evidence-based treatment options/medical procedures obtained from a registered medical practitioner for the medical condition given in the CDM of the app.	Facilitating conditions	FC2
20	I want to get information on the treatment options/medical procedures from Wikipedia rather than from a registered medical practitioner for the given medical condition in the CDM of the app.	Facilitating conditions	FC3

Table 7. Appendix A: Survey questions categorized by the construct

References

Arvisais-Anhalt, S., S. McDonald, J. Y. Park, K. Kapinos, C. U. Lehmann, and M. Basit (2021). "Survey of hospital chargemaster transparency." *Applied clinical informatics* 12 (02), 391–398.

Attafuah, P. Y. A., I. H. Everink, C. Lohrmann, A. A. Abuosi, and J. M. Schols (2022). "Improving health and social care services for slum-dwelling older adults: Perspectives of health professionals." *Frontiers in Public Health*, 3649.

Bawack, R. E. and J. R. K. Kamdjoug (2018). “Adequacy of UTAUT in clinician adoption of health information systems in developing countries: The case of Cameroon.” *International journal of medical informatics* 109, 15–22.

Bielenberg, J. E., M. Futrell, B. Stover, and A. Hagopian (2020). “Presence of any medical debt associated with two additional years of homelessness in a Seattle sample.” *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 57, 0046958020923535.

Bindman, A. B. (2020). “Rising prices and health care "empires".” *Jama* 323 (9), 815–816.

Brinkel, J., J. May, R. Krumkamp, M. Lamshöft, B. Kreuels, E. Owusu-Dabo, A. Mohammed, A. Bonacic Marinovic, P. Dako-Gyeke, A. Krämer, et al. (2017). “Mobile phone-based interactive voice response as a tool for improving access to healthcare in remote areas in Ghana–an evaluation of user experiences.” *Tropical Medicine & International Health* 22 (5), 622–630.

Dean, C. A., J. Wiltshire, E. Liu, M. A. Amamoo, E. Garcia Colato, and K. Elder (2020). “Confidence in understanding health insurance and challenges paying medical bills among men in the United States.” *American Journal of Men’s Health* 14 (4), 1557988320943359.

Efiloğlu Kurt, Ö. and Ö. Tingöy (2017). “The acceptance and use of a virtual learning environment in higher education: an empirical study in Turkey, and the UK.” *International Journal of Educational Technology in Higher Education* 14 (1), 1–15.

Evashwick-Rogler, T. W., S. W. Dooley, C. D. Murawski, M. S. Fourman, M. V. Hogan, and N. A. Smyth (2022). “The mandated publication of individual hospital charge description masters does not permit the estimation of complex procedure charges.” *JAAOS-Journal of the American Academy of Orthopaedic Surgeons* 30 (1), e118–e123.

Hair Jr, J. F., M. Sarstedt, L. Hopkins, and V. G. Kuppelwieser (2014). “Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research.” *European business review*.

Haque, W., M. Ahmadzada, S. Janumpally, E. Haque, H. Allahrakha, S. Desai, and D. Hsiehchen (2022). “Adherence to a Federal Hospital Price Transparency Rule and Associated Financial and Marketplace Factors.” *JAMA* 327 (21), 2143–2145.

Hartman, M., A. B. Martin, B. Washington, A. Catlin, N. H. E. A. Team, et al. (2022). “National Health Care Spending In 2020: Growth Driven By Federal Spending In Response To The COVID-19 Pandemic: National Health Expenditures study examines US health care spending in 2020.” *Health Affairs* 41 (1), 13–25.

Hirsch, J. A., A. B. Rosenkrantz, S. A. Ansari, L. Manchikanti, and G. N. Nicola (2017). “MACRA 2.0: are you ready for MIPS?” *Journal of neurointerventional surgery* 9 (7), 714–716.

Ivankova, N. V., J. W. Creswell, and S. L. Stick (2006). “Using mixed-methods sequential explanatory design: From theory to practice.” *Field methods* 18 (1), 3–20.

Jewer, J. (2018). “Patients’ intention to use online postings of ED wait times: A modified UTAUT model.” *International journal of medical informatics* 112, 34–39.

Kona, M. and S. Corlette (2022). “Hospital And Insurer Price Transparency Rules Now In Effect But Compliance Is Still Far Away.” *Health Affairs Forefront*.

Kratka, A., C. A. Wong, R. Herrmann, K. Hong, A. Karediya, I. Yang, and P. A. Ubel (2018). “Finding health care prices online—how difficult is it to be an informed health-care consumer?” *JAMA Internal Medicine* 178 (3), 423–424.

Linde, S. and L. E. Egede (2022). “Do Chargemaster Prices Matter?: An Examination of Acute Care Hospital Profitability.” *Medical Care* 60 (8), 623–630.

Liu, D., R. Maimaitijiang, J. Gu, S. Zhong, M. Zhou, Z. Wu, A. Luo, C. Lu, Y. Hao, et al. (2019). “Using the unified theory of acceptance and use of technology (UTAUT) to investigate the intention to use physical activity apps: cross-sectional survey.” *JMIR mHealth and uHealth* 7 (9), e13127.

Lu, A. J., E. M. Chen, E. Vutam, J. Brandt, and P. Sadda (2020). “Price transparency implementation: accessibility of hospital chargemasters and variation in hospital pricing after CMS mandate.” In: *Healthcare*. Vol. 8. 3. Elsevier, p. 100443.

Miller, A. S., S. A. Stearns, and D. M. Berwick (2022). "Hospital noncompliance with US price transparency regulations." *The Lancet Regional Health–Americas* 12.

Reddy, S., G. Daly, S. Baban, A. Kadesh, A. E. Block, and C. L. Grimes (2022). "Accessibility and usability of hospital chargemasters in New York State." *Journal of General Internal Medicine* 37 (8), 2130–2131.

Subedi, D. (2016). "Explanatory sequential mixed method design as the third research community of knowledge claim." *American Journal of Educational Research* 4 (7), 570–577.

Sun, Y., N. Wang, X. Guo, and Z. Peng (2013). "Understanding the acceptance of mobile health services: a comparison and integration of alternative models." *Journal of electronic commerce research* 14 (2), 183.

Toci, G. R., W. A. Hassan, M. J. Lambrechts, B. Wellens, and A. R. Vaccaro (2022). "Price Transparency: The Good, the Bad, and the Ugly." *Contemporary Spine Surgery* 23 (6), 1–5.

Venkatesh, V., M. G. Morris, G. B. Davis, and F. D. Davis (2003). "User acceptance of information technology: Toward a unified view." *MIS quarterly*, 425–478.

Venkatesh, V., J. Y. Thong, and X. Xu (2012). "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology." *MIS quarterly*, 157–178.