


Outpatient portal use among pregnant individuals: Cross-sectional, temporal, and cluster analysis of use

Digital Health
Volume 8: 1–14
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076221109553
journals.sagepub.com/home/dhj


Evan Morgan¹, Patrick Schnell², Priti Singh¹ and Naleef Fareed¹ 

Abstract

Background: Outpatient portal technology can improve patient engagement. For pregnant individuals, the level of engagement could have important implications for maternal and infant outcomes.

Objective: This study: (1) cross-sectionally and temporally characterized the outpatient portal use among pregnant individuals seen at our academic medical center; and (2) identified clusters of the outpatient portal user groups based on the cross-sectional and temporal patterns of use.

Methods: We used outpatient portal server-side log files to execute a hierarchical clustering algorithm to group 7663 pregnant individuals based on proportions of outpatient portal function use. Post-hoc analyses were performed to further assess outpatient portal use on key encounter characteristics.

Results: The most frequently used functions were MyRecord (access personal health information), Visits (manage appointments), Messaging (send/receive messages), and Billing (view bills, insurance information). Median outpatient portal function use plateaued by the third trimester. Four distinct clusters were identified among all pregnant individuals: “Schedulers,” “Resulters,” “Intense Digital Engagers,” and “Average Users.” Post-hoc analyses revealed that the use of the Visits function increased and the use of the MyRecord function decreased over time among clusters.

Conclusions: Our identification of distinct cluster groups of outpatient portal users among pregnant individuals underscores the importance of avoiding the use of generalizations when describing how such patients might engage with patient-facing technologies such as an outpatient portal. These results can be used to improve user experience and training with outpatient portal functions and may educate maternal health providers on patient engagement with the outpatient portal.

Keywords

Patient portals, health information technology, prenatal care, log files, cluster analysis

Submission date: 16 March 2022; Acceptance date: 8 June 2022

Introduction

Background

Patient portals serve as the bridge between an electronic personal health record, which is designed to be managed by patients, and the electronic health record that is managed by the health care system.^{1, 2} In the outpatient setting, patients are able to communicate with health care providers using an outpatient portal and actively manage their health via

various portal functions. Actions include viewing laboratory

¹Department of Biomedical Informatics, The Ohio State University, USA

²Department of Biostatistics, The Ohio State University, USA

Corresponding author:

Naleef Fareed, Department of Biomedical Informatics College of Medicine, The Ohio State University, 460 Medical Center Drive, Columbus, OH 43210, USA.

Email: naleef.fareed@osumc.edu



results and visit summaries, scheduling appointments and prescription refills, and communicating with clinicians. Various studies have demonstrated that adult users of outpatient portals can experience positive clinical and quality outcomes, particularly in relation to chronic conditions such as diabetes, cancer, coronary artery disease, hyperlipidemia, and hypertension.^{3–9} Similar to individuals with chronic conditions, pregnant individuals are in frequent contact with health care providers and are encouraged to actively participate in their health care given the uncertainties and changes in health during this period. Analysis of log files generated from pregnant individual's use of an outpatient portal presents a unique opportunity to leverage large-scale behavioral data toward the study of an underrepresented population's interaction with a digital health tool.

While pregnant individuals do perceive the benefits of outpatient portal use,^{10–13} actual enrollment in and use of portals may vary based on factors such as pregnancy history, age, race, socioeconomic status, or internet access as well as experiences, preferences, and expectations related to the outpatient portal itself.^{14–16} Ukoha and colleagues examined whether enrollment in and use of secure messaging in an outpatient portal were associated with certain demographic and clinical covariates. They found that individuals with high-risk pregnancies were less likely to enroll in the outpatient portal, but did not differ in their use of secure messaging within the outpatient portal.¹⁶ Considering that an outpatient portal can be utilized in a more comprehensive manner than secure messaging alone, it is necessary to redefine use to account for the numerous ways this population can engage with an outpatient portal. Learning more about the behavioral interaction between pregnant individuals and the patient portal through the use of the generated log files could provide insight and guidance to help obstetrics and gynecology and maternal–fetal medicine programs improve the care delivered, patient satisfaction, and possibly clinical outcomes associated with pregnancy, as has been explored among the general population.^{17–19} Our study is the first to our knowledge that characterizes outpatient portal use and identifies user profiles in order to explore how this understudied patient population engages with their health care electronically. Furthermore, our study aims to demonstrate that outpatient portal use is heterogeneous and attempts to characterize outpatient portal use should avoid the generalization that all users interact with an outpatient portal equally. Categorizing outpatient portal use into a variety of profiles, or clustering, can benefit future works that aim to identify the clinical or quality outcomes associated with patterns of patient engagement with an electronic health care delivery technique.

Audit log files

Mining of outpatient portal audit log file data presents a unique pathway to explore how patients are managing their health via engagement with the outpatient portal.

Audit log files are an automatically produced, server-side documentation of user behavior within a particular system.²⁰ Originally designed for monitoring purposes, these files typically include four pieces of data: *who* accessed, *which* page, *when*, and the *action* they performed on that page.²¹ While the exact granularity of the log files depends on the vendor's software, the use of this information allows researchers to unobtrusively study a wide range of user activities within a system, such as an electronic health record or an outpatient portal. Log files are limited by their inability to precisely indicate why actions occur. However, the scale at which this activity information exists has been harnessed to summarize sequences of activity and further identify groups of common sequential actions.^{20–23}

Processing audit log file data presents a challenge; however, recent studies have developed, standardized, and shared such methods for processing patient portal audit log files and measuring patient portal use.^{24–26} These studies were performed across samples of an entire health system and did not examine how such metrics of use could be applied to patient portal use among unique patient populations. Jones and colleagues also presented a summary of data sources, data processing methods, and metrics of use along with results from a hierarchical cluster analysis among a population of users with chronic illnesses at a large academic medical center.²⁷ While this study offered a valuable view of clustered outpatient portal users, it did not address how use may change over time as users become more familiar with the outpatient portal. In addition, the motivations for and patterns of outpatient portal use among chronically ill patients are likely different from those of pregnant individuals given the durations and exposures to health care.

Objectives

The primary objective of this study was to cross-sectionally and temporally characterize the outpatient portal use among pregnant individuals seen by the obstetrics and gynecology and maternal–fetal medicine departments at our academic medical center. Our secondary objective was to then identify clusters of the outpatient portal user groups based on the cross-sectional and temporal patterns of use. Our overall hypotheses were that there will be differences in outpatient portal use among pregnant individuals and that there will be additional differences based on clinical risk.

Methods

Overview

This study builds on a larger assessment of patient portal technology, which has been described by Huerta et al., Di Tosto et al., and Fareed et al.^{24–26} Drawing on methods developed in these studies, we performed data processing

and analysis using coded-limited data requested from the Information Warehouse within our academic medical center in order to assess outpatient portal use. This study did not require approval from the Institutional Review Board of our academic medical center as it involved the use of a coded-limited data set reviewed and approved by the Honest Broker Committee of our academic medical center.

Study setting and sample

Our mid-western academic medical center implemented the outpatient portal in 2011 and sees about 7000 pregnant individuals annually. The outpatient portal is Epic's MyChart and contains the following functions that remained consistent over our study period: Messaging (send/receive messages), Visits (manage appointments), MyRecord (access personal health information such as test results, allergies, current health issues, and immunizations), Medical Tools (share records, participate in research, and connect devices), Billing (view bills and insurance information), Resources (frequently asked questions, terms and conditions, and patient education), Proxy (request proxy status and renew request), Preferences (change settings and notifications), and Custom (use of miscellaneous features). Consistent with the methods of Di Tosto et al., the use of the Resources function was not assessed as an individual's engagement with this function leads to external actions that are not logged in the audit log files. This study included a retrospective cohort of 14,658 individuals who were aged ≥ 18 years and received prenatal care from obstetrics and gynecology or maternal-fetal medicine providers at our academic medical center during the study period of 1/1/2016 to 8/1/2020. In the event an individual had multiple pregnancies during the study period, we only examined outpatient portal use during the individual's first pregnancy. We also excluded individuals who did not use the outpatient portal within 280 days of their estimated delivery date to ensure we only assessed outpatient portal use during an individual's pregnancy. Our final analyses only included 7664 individuals, who were considered active users of the outpatient portal during their pregnancy if they used the outpatient portal at least once within each trimester of pregnancy. There were 6233 individuals with a normal pregnancy and 1431 individuals with a high-risk pregnancy. An individual had a "high-risk" pregnancy if she required a visit with a maternal-fetal medicine provider for existing comorbidities (i.e. diabetes, high blood pressure, or any condition requiring high-risk care or fetal treatment).²⁸

Data sources

The audit log files from patients seen from 1/1/2016 to 8/1/2020 were the primary data source linked to electronic health record data. A categorical variable named "risk" was constructed to denote the level of risk associated with

the pregnancy, as mentioned previously. Trimesters were calculated for each pregnancy using the number of days prior to the estimated delivery date, according to the American College of Obstetricians and Gynecologists.²⁹

Data model

Our data model for outpatient portal use is presented in Table S1 of Appendix 1. We first categorized every patient's individual action within the outpatient portal, and then collapsed these action types into various outpatient portal functions. Sessions were defined as grouped sequences of actions that began with a patient's login and were completed by either a patient's manual logout or a period of inactivity between actions within the outpatient portal. The threshold for a period of inactivity was determined by calculating the 99th percentile of the time in seconds between actions (Module 2 in Appendix 2). A single session may involve the use of multiple actions or functions, and outpatient portal use can continue over numerous sessions throughout the course of a patient's pregnancy.

Using the timestamped actions and the estimated delivery date, we were able to examine outpatient portal use cross-sectionally at various timepoints and temporally as individuals progressed through pregnancy. Cross-sectional analysis of outpatient portal use was performed at the session and patient levels, while temporal analysis took place at the patient level. Clusters of patients were later identified based on the proportion of function use at the patient level. The movement of patients between clusters was assessed temporally as individuals progressed through pregnancy. Finally, all measurements of outpatient portal use were repeated to examine differences in outpatient portal use between individuals with a normal pregnancy and individuals with a high-risk pregnancy. The measurements of outpatient portal engagement, defined in recent studies,²⁴⁻²⁶ are listed in Table S2 of Appendix 1.

Data processing

A summary of the steps used for data processing is presented as a flowchart in Figure S1 of Appendix 1. Briefly, modules 1-7 were an application of the data processing methods outlined by Di Tosto et al., performed with the goal of translating the audit log files into data sets from which session-level outpatient portal engagement could be calculated. In module 8, we calculated the active status of each patient by using the first session start date that falls within the specified period of a patient's pregnancy (i.e. the first, second, or third trimester). Those with at least one session start date within each trimester period were considered active. The remaining modules were performed to measure outpatient portal engagement cross-sectionally and temporally among each pregnancy risk

group (i.e. all pregnancies, normal pregnancies, and high-risk pregnancies) and among clusters of users.

Data analysis

Our analyses involved: (1) the cross-sectional descriptive summary of function use at the session and patient levels; (2) the temporal summary of median function use at the patient level; (3) the cluster analyses according to the proportion of function use; (4) post hoc analyses to further characterize clusters of function use; (5) the temporal summary of movement between clusters, and (6) sensitivity cluster analyses examining outpatient portal use during the COVID-19 pandemic. Hierarchical agglomerative clustering algorithms were used to group individuals according to the proportion of outpatient portal function use. The proportion of outpatient portal function use calculation included the summated use of all nine outpatient portal functions as the denominator. Our hierarchical clustering analyses used Ward's method, which minimized within-cluster variance, as measured by the error sum of squares, across all proportions of use variables in order to determine similarity.³⁰ The final number of clusters was agreed upon following a review of a cluster dendrogram in conjunction with the use of cluster-analysis stopping rules.^{31, 32}

Additional post hoc analyses included the calculation of median values for the following metrics: number of sessions per pregnancy, pregnancy length, number of days to the first session of use, number of days to the last session of use, number of days to delivery, number of days to estimated delivery, number of maternal–fetal medicine visits, number of obstetrics and gynecology visits, and number of provider visits (Appendix 1, Table S2). Race and ethnicity for each cluster group were also assessed. A chi-square test was performed to test for a significant difference in proportions of active users. Kruskal–Wallis tests were performed to test for significant differences in median outpatient portal function use across trimesters, while Wilcoxon rank-sum tests were used to test for significant differences between trimesters and between pregnancy risk groups at each trimester. To test for significant differences in median proportions of outpatient portal function use, median post hoc metrics across an entire set of clusters,

and percent of clusters according to race, we performed Kruskal–Wallis tests within each pregnancy risk group at each cross-sectional time point (i.e. the entire pregnancy period and each trimester of pregnancy). Significant differences in mean proportion of outpatient portal use between individual clusters were assessed using Dunn's multiple comparison tests. Statistical analyses were performed using Stata 16.1, R programming language (v4.0.4), and Python programming language (v3.9). Our clustering method used the Stata cluster command with Ward's linkage and squared Euclidean distance.

Results

Outpatient portal active use status

Table 1 shows the classification of active and inactive outpatient portal users among pregnant individuals. Following data processing (Appendix 1, Figure S1), 7664 (52%) of the original 14,658 pregnant individuals were considered active users. A greater proportion of normal pregnancy individuals (54%) were active users compared to high-risk pregnancy individuals (45%) ($P < .001$).

Patient characteristics

Our study sample consisted primarily of non-Hispanic White (70%) individuals who were 30–34 years old (39%) and 25–29 years old (31%) at the time of their first visit (Table 2). Both pregnancy risk groups predominantly consisted of non-Hispanic White individuals between the ages of 30–34 at the time of the first visit. There were more non-Hispanic White individuals among the high-risk pregnancy group (75%) compared to those with a normal pregnancy (69%), and more individuals over the age of 35 years among high-risk pregnancy individuals (26%) compared to those with a normal pregnancy (16%).

Cross-sectional descriptive summary of function use

Table 3 shows the frequency of outpatient portal function use per session and per patient among active users, according to the pregnancy risk group.

Table 1. Portal active status per pregnancy, by pregnancy risk group.

	All pregnancies ($N = 14,658$) n (%)	Normal pregnancies ($N = 11,451$) n (%)	High-risk pregnancies ($N = 3207$) n (%)
Active status			
Active	7664 (52)	6233 (54)	1431 (45)
Inactive	6994 (48)	5218 (46)	1776 (55)

Table 2. Patient characteristics, by pregnancy risk group.

	All pregnancies (<i>N</i> = 7664) <i>n</i> (%)	Normal pregnancies (<i>N</i> = 6233) <i>n</i> (%)	High-risk pregnancies (<i>N</i> = 1431) <i>n</i> (%)
Age at first visit			
<20	110 (1)	95 (2)	15 (1)
20–24	809 (11)	675 (11)	134 (9)
25–29	2376 (31)	2000 (32)	376 (26)
30–34	2982 (39)	2450 (39)	532 (37)
35 +	1387 (18)	1013 (16)	374 (26)
Race-ethnicity			
Non-Hispanic White	5393 (70)	4326 (69)	1067 (75)
Non-Hispanic Black	1111 (15)	935 (15)	176 (12)
Hispanic	238 (3)	195 (3)	43 (3)
Race-ethnicity other	922 (12)	777 (13)	145 (10)

At the session level, the most frequently used functions were Visits (31%), MyRecord (24%), Messaging (10%) and Billing (10%). These are also the most frequently used functions for both pregnancy risk groups, with the use of these four functions being relatively similar between the two groups. The most frequently used functions at the patient level were the same as those at the session level, namely MyRecord (99%), Visits (97%), Messaging (92%) and Billing (89%). Such use was consistent across both pregnancy risk groups.

Temporal summary of median function use

Figure 1 show the median number of sessions among (a) all active pregnant individuals and (b) according to pregnancy risk group. Among all individuals, median function use increased significantly ($P < .001$) from 14 sessions of function use in the first trimester to 22 sessions in the second and third trimesters. This trend of a plateauing median function use between the second and third trimesters continued when examining each pregnancy risk group. However, the median use among high-risk pregnancy individuals was significantly greater than use among normal pregnancy individuals ($P < .001$) at each trimester.

Figure 2 shows the median number of sessions with at least one use of the MyRecord, Visits, Messaging, or Billing function during each trimester, according to the pregnancy risk group. Among individuals with a normal pregnancy, median use of the MyRecord and Messaging functions significantly decreased between the second (7 sessions/trimester and 2 sessions/trimester, respectively) and third trimesters (6 sessions/trimester and 1 session/trimester, respectively; $P < .001$). Among high-risk pregnancy individuals, median use of the Messaging function did not significantly decrease between the second (3 sessions/trimester) and third trimesters (2 sessions/trimester, $P = .63$). Median use of the Visits function significantly increased each trimester among both pregnancy risk groups ($P < .001$), and high-risk pregnancy individuals had significantly more sessions of use during the second and third trimesters compared to normal pregnancy individuals ($P < .001$). There was no significant difference in median use of the Billing function between pregnancy risk groups, though use did significantly increase for each group between the first and second trimesters ($P < .001$).

Cluster analyses according to the proportion of use

The following sections present results from hierarchical cluster analyses during all trimesters of pregnancy, according to pregnancy risk groups. In addition, we present post hoc measurements used to further characterize clusters of function use. Cluster analyses and relevant post hoc measurements were also performed at each trimester time point, and the results are presented in Figures S2–S4 in Appendix 1. Unless otherwise specified, all differences in outpatient portal use and post hoc measurements across clusters and between clusters are statistically significant ($P < .05$ following Kruskal–Wallis tests and $P < .05$ following Dunn’s multiple comparison tests).

Figure 3 shows the results from the hierarchical cluster analyses during all trimesters of a pregnancy, according to pregnancy risk groups. A review of the Calinski/Harabasz pseudo-F indexes, Duda/Hart scores, and the dendrograms for cluster analyses using the study sample of each pregnancy risk group led to the selection of four clusters to categorize all pregnant individuals, three clusters to categorize normal pregnancy individuals and four clusters to categorize high-risk pregnancy individuals. While clusters are identified according to proportions of outpatient portal use, only the proportions for the top four outpatient portal functions are presented in each cluster (Figure 3; Appendix 1, Figures S2–S4) as the remaining functions were rarely used.

Among all pregnancies, outpatient portal users primarily spent their time using MyRecord (35%) and Visits (30%), while using Messaging (9%) and Billing (9%) with the remainder of their time. This outpatient portal use cluster is presented as the “Study Sample” column in the “All Pregnancy Episodes” section of Figure 3 and is used as the reference for the remaining clusters of outpatient

Table 3. Frequency of portal function use per session and pregnancy among active users, according to pregnancy risk group.

	Sessions of function use			Patients with function use		
	All pregnancies (N = 813,902), n (%)	Normal pregnancies (N = 622,471), n (%)	High-risk pregnancies (N = 191,431), n (%)	All pregnancies (N = 7664), n (%)	Normal pregnancies (N = 6233), n (%)	High-risk pregnancies (N = 1431), n (%)
Visits	256,167 (31)	199,069 (31)	57,098 (29)	7446 (97)	6035 (96)	1411 (98)
MyRecord	201,222 (24)	154,019 (24)	47,203 (24)	7659 (99)	6228 (99)	1431 (100)
Messaging	83,385 (10)	63,088 (10)	20,297 (10)	7087 (92)	5737 (92)	1350 (94)
Billing	83,124 (10)	66,929 (10)	16,195 (8)	6827 (89)	5565 (89)	1262 (88)
Proxy	10,719 (1)	8238 (1)	2481 (1)	1925 (25)	1555 (24)	370 (25)
Preferences	8089 (0)	6476 (1)	1613 (0)	3387 (44)	2764 (44)	623 (43)
Custom	5673 (0)	4507 (0)	1166 (0)	2500 (32)	1991 (31)	509 (35)
Medical tools	3874 (0)	3099 (0)	775 (0)	2254 (29)	1810 (29)	444 (31)

portal user groups. The “Average Users” cluster was 25% of the “Study Sample” and was most similar to the “Study Sample” with frequent use of MyRecord (35%) and Visit (30%). The “Schedulers” cluster was 30% of the “Study Sample” and was comprised of those who primarily used the Visits (55%) function to schedule, view details for, and check-in for an appointment. The “Resulters” cluster was 29% of the “Study Sample” and consisted of those who mostly used the MyRecord (52%) function to view test results as well as occasionally viewing personal health information like allergies, current health issues, and immunizations. Finally, the “Intense Digital Engagers” cluster was 16% of the “Study Sample.” They split most of their use between Messaging (44%) and MyRecord (32%) and consistently had the least sessions per pregnancy, which may indicate an aptitude with the digital health tool that enables intense, efficient engagement with their providers and personal health information.

Post hoc analyses from clusters among all pregnant individuals showed that the “Resulters” and “Intense Digital Engagers” have far fewer sessions of use per pregnancy (53 and 50, respectively) compared to the “Study Sample” (73), while the “Schedulers” have the most sessions per pregnancy (109). The “Schedulers” had the least days to the first session (60) and most days to the last session (96). Conversely, “Resulters” and “Average Users” started the latest (64) and “Resulters” and “Intense Digital Engagers” stopped the earliest (94) compared to other clusters. In addition, we saw that “Intense Digital Engagers” met with their obstetrics and gynecology providers the most (14

visits) throughout pregnancy compared to all other clusters (13). Finally, the “Study Sample” was predominantly represented by non-Hispanic White individuals (70%), followed by non-Hispanic Black individuals (15%), individuals of Other Race-Ethnicity (12%), and Hispanic individuals (3%). The “Intense Digital Engagers” contained the most non-Hispanic White individuals (77%) and least non-Hispanic Black individuals (10%), while the “Schedulers” had the fewest non-Hispanic White individuals (65) and the most non-Hispanic Black individuals (18). The “Resulters” cluster similarly contained many non-Hispanic White individuals (73%) and few non-Hispanic Black individuals (14%) compared to the “Schedulers.”

Individuals within the “Study Sample” of the normal pregnancy and high-risk pregnancy groups similarly used the MyRecord and Visits functions most often, while Messaging and Billing were used less. The common clusters between pregnancy risk groups were “Schedulers,” “Intense Digital Engagers,” and “Average Users.” The “Average Users” are the largest within both pregnancy risk groups, though more normal pregnancy individuals were “Average Users” (55%) compared to high-risk pregnancy individuals (36%). While “Schedulers” were equally present among both groups (25%), the “Resulters” were only present among high-risk pregnancy individuals (15%). Lastly, fewer normal pregnancy individuals were “Intense Digital Engagers” (20%) compared to high-risk pregnancy individuals (24%).

Post hoc analyses demonstrated that high-risk pregnancy individuals have more sessions (97) and more visits with a

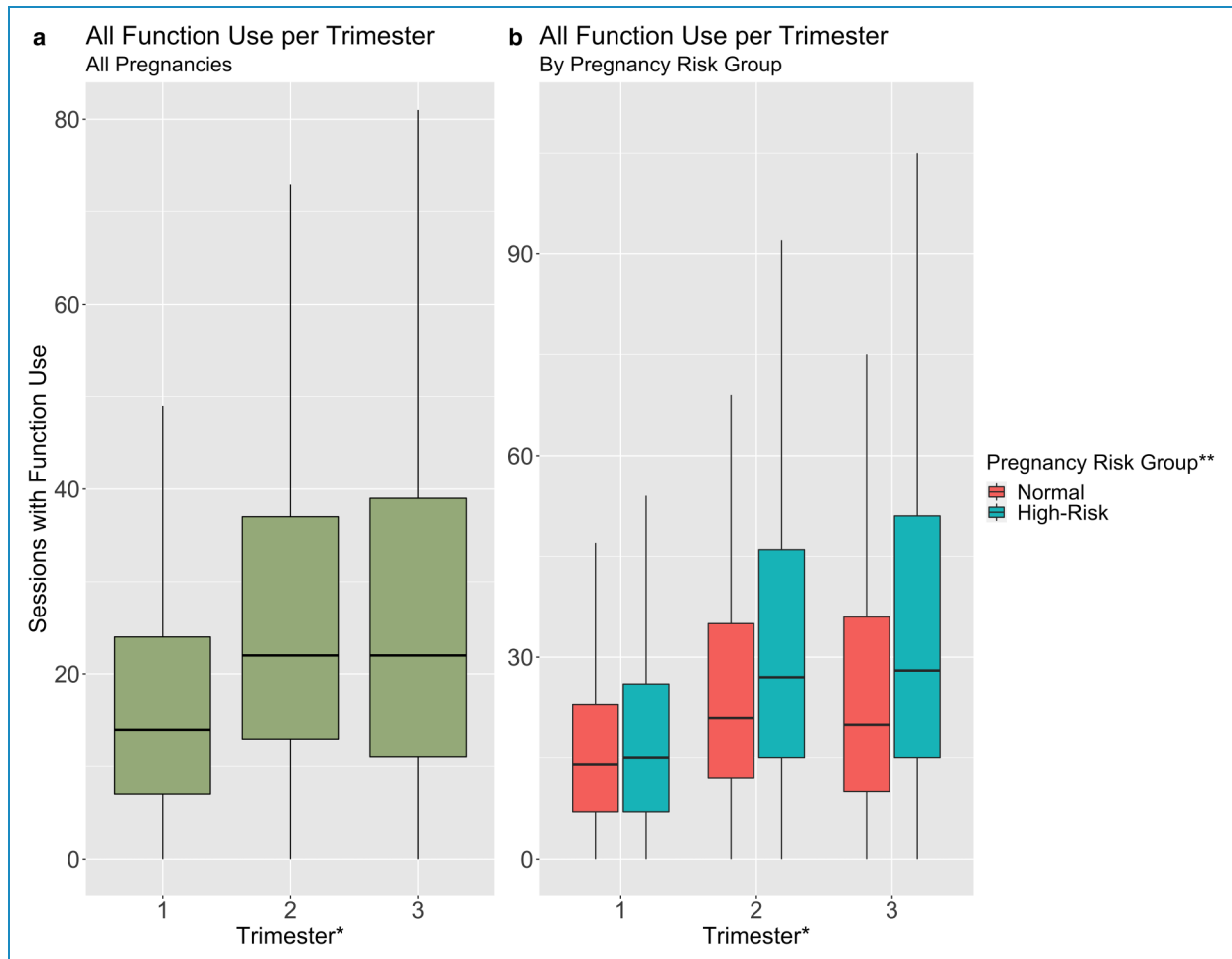


Figure 1. Median function use per trimester among (a) all women and (b) women with either a normal or high-risk pregnancy. *Median OPP function use with a significant difference at $P < .05$ level across trimesters; median OPP function use with significantly different values at $P < .05$ between trimesters 1-2 and 1-3 among (a) all women and (b) high-risk pregnancy group, and between all trimesters among (b) normal pregnancy group. **Median OPP function use with significantly different values at $P < .05$ between pregnancy risk groups at each trimester. Error bars indicate the interquartile range equal to the 75th and 25th percentiles of median function use at each trimester. Kruskal-Wallis used to test for overall significant differences in median OPP function use across all trimesters. Wilcoxon rank-sum tests were used to examine significant differences in median OPP function use between trimesters and pregnancy risk groups at each trimester.

provider (15) compared to normal pregnancy individuals (69 and 14, respectively), though this difference was not tested for significance. Among high-risk pregnancy individuals, we saw that “Schedulers” had the most sessions per pregnancy (134), while “Intense Digital Engagers” had the least (66). Comparatively, “Schedulers” among normal pregnancy individuals also had the most sessions per pregnancy (101) and “Intense Digital Engagers” had the least (46).

In addition, we saw that the “Schedulers” among normal pregnancy and high-risk pregnancy individuals had fewer days to the first session of use (60 and 63, respectively) and more days to the last session of use (96) compared to other clusters. In other words, individuals in these Visits-oriented clusters began using the outpatient portal slightly earlier and continued using it longer into their pregnancies compared to individuals in the “Study Sample.”

The “Intense Digital Engagers” of the pregnancy risk groups were the inverse of the “Schedulers” as they tend to start later and stop earlier compared to the “Study Sample” in their respective pregnancy group.

Finally, we saw that racial and ethnic differences among clusters persisted when comparing pregnancy risk groups. For instance, non-Hispanic White individuals were similarly common among “Intense Digital Engagers” and “Resulters” of high-risk pregnancy individuals (79% and 82%, respectively) compared to those “Intense Digital Engagers” of normal pregnancy individuals (78%). Non-Hispanic Black individuals were also similarly more common among the “Schedulers” of normal pregnancy individuals (20%) and high-risk pregnancy individuals (17%) compared to other clusters in the respective pregnancy groups.

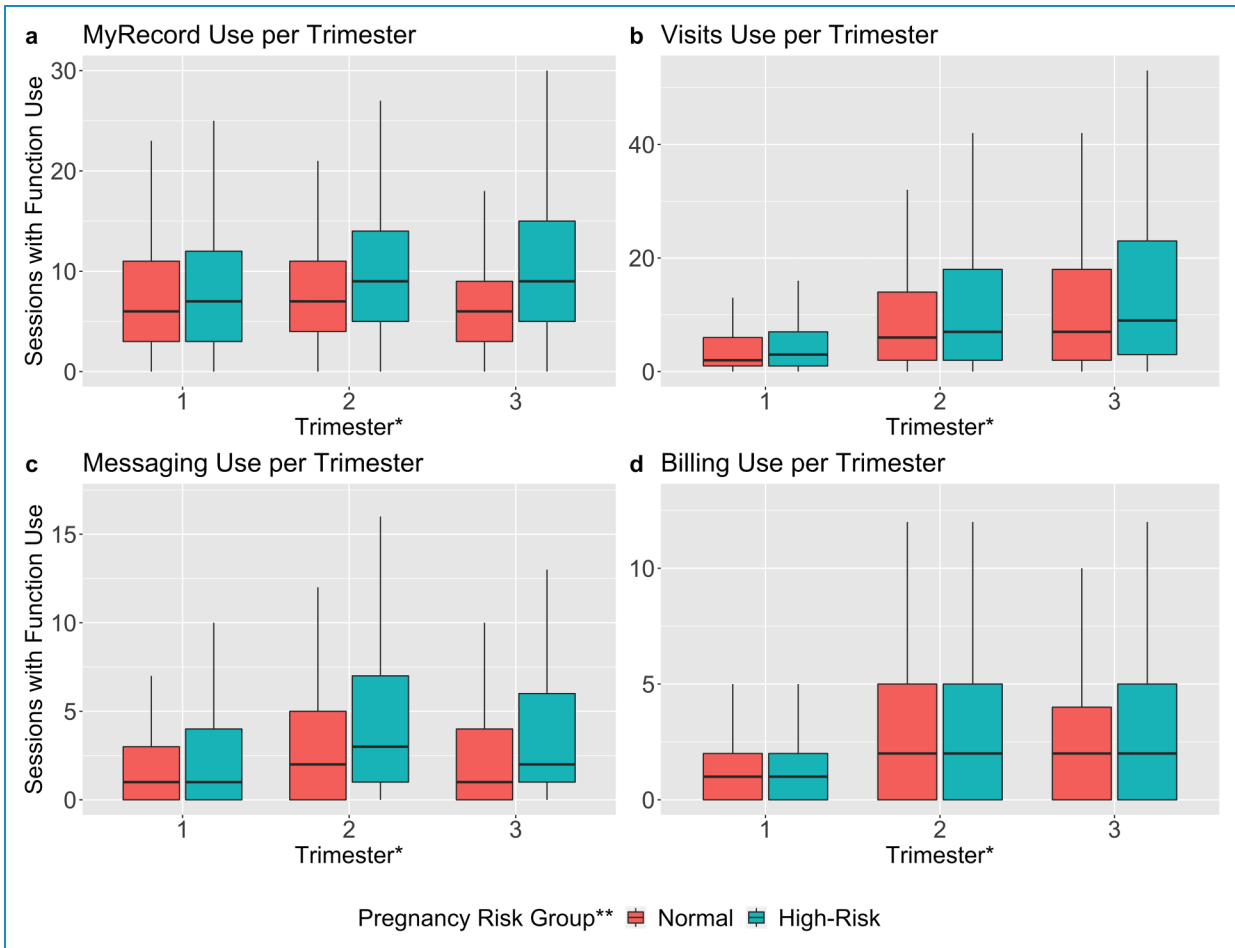


Figure 2. Median (a) MyRecord, (b) Visits, (c) Messaging, and (d) Billing use per trimester among all women. *Median OPP function use with a significant difference at $P < .05$ level across trimesters for all functions in (a)–(d); among normal pregnancy group, median OPP function use with significantly different values at $P < .05$ between all trimesters for all functions in (a)–(d); among high-risk pregnancy, median OPP function use with significantly different values at $P < .05$ between all trimesters for function in (b) and between trimesters 1–2 and 1–3 for functions in (a), (c), and (d). **Median OPP function use with significantly different values at $P < .05$ between pregnancy risk groups at all trimesters for functions in (a) and (c), and in second and third trimester for function in (b). Error bars indicate the interquartile range equal to the 75th and 25th percentiles of median function use at each trimester. Kruskal–Wallis used to test for overall significant differences in median OPP function use across all trimesters. Wilcoxon rank-sum tests were used to examine significant differences in median OPP function use between trimesters and pregnancy risk groups at each trimester.

When examining the characteristics of clusters at each trimester we observed two trends (Appendix 1, Figures S2–S4). First, the use of the Visits function increased and the use of the MyRecord function decreased over time. By the third trimester, the Visits function surpassed the MyRecord function as the most used function within the “Study Sample” of all pregnant individuals. Second, the median number of sessions per trimester plateaued after the second trimester among all pregnant users in a Visits-oriented cluster (e.g. “Schedulers”) and decreased each trimester among all pregnant users in “Average Users” and “Intense Digital Engagers.” Amidst these trends, we consistently observed that the “Schedulers” consisted of more non-Hispanic Black individuals compared to the Results- and Message-oriented clusters (“Resulters” and “Intense Digital Engagers”).

Temporal summary of movement between clusters

Figure 4 depicts the flow of all pregnant individuals between clusters throughout all trimesters of their pregnancy. Though the “Average Users” were typically one of the largest clusters in each trimester, the majority of “Average Users” subsequently moved to new clusters at the onset of a new trimester. The “Resulters” were similarly inclined to test new functions and move to new clusters over time. Comparatively, individuals in the “Intense Digital Engagers” and “Schedulers” typically remained focused on their preferred functions, with movement between clusters subsiding between second and third trimesters. “Non-users” are those who have logged a session in the outpatient portal during a trimester but did not engage

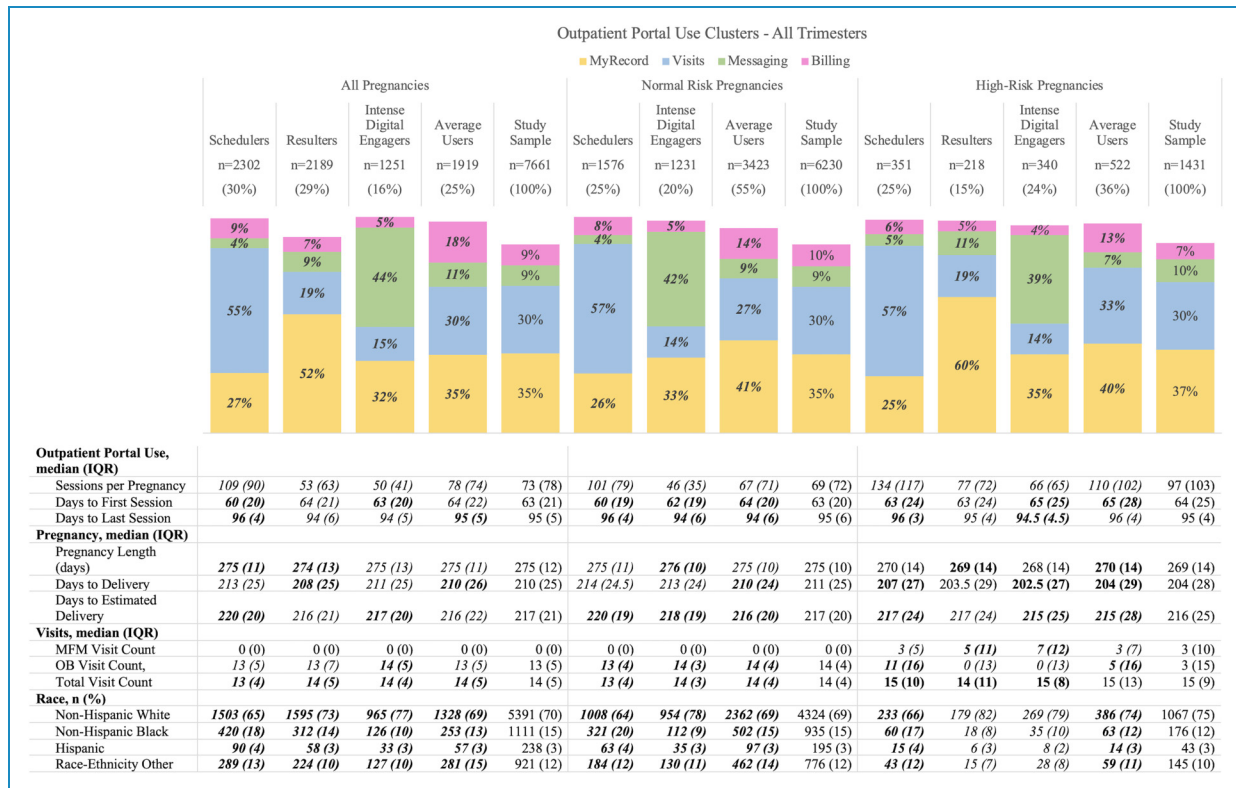


Figure 3. The median proportion of OPP use and summary statistics (median (IQR); n (%)) during all trimesters by OPP cluster types and study sample, according to pregnancy risk group. Italicized values indicate OPP function use or post hoc measurement with a significant difference at $P < .05$ level across clusters. Bold values indicate OPP function use or post hoc measurement significantly different values at $P < .05$ for a particular cluster relative to other clusters. Total proportions do not add to 1.00 as only the top four functions are shown. Kruskal-Wallis used to test for overall significant differences across the entire set of clusters among each pregnancy risk group; Dunn's multiple comparison tests were used to examine significant differences in means across individual clusters of each pregnancy risk group.

with any of the functions listed. Few “Non-users” persist over two trimesters and there were no pregnant individuals who were “Non-users” across all trimesters. Movement between clusters was also examined between pregnancy risk groups, and results are seen in Figures S5–S6 in Appendix 1. We continued to see similar movement among the “Average Users,” “Intense Digital Engagers,” and “Schedulers” in both pregnancy risk groups.

Sensitivity analysis: Outpatient portal use during the COVID-19 pandemic

A final cluster analysis was performed among a subsample of individuals who only used the outpatient portal between 1/1/2020 and 8/1/2020 in order to determine if clusters of outpatient portal use varied during the COVID-19 pandemic compared to our original sample. Appendix 3 contains tables with the proportions of outpatient portal use among this subsample of individuals, listed by pregnancy risk group and trimesters. Table S1 in Appendix 3 demonstrates that the Visits function was used more than two times more than the MyRecord function (57% vs 25%,

respectively) among the entire population of pregnant individuals. This disparity in the use of these functions is very different from the relatively even use that was seen in the original sample. In addition, the Proxy function was occasionally used among clusters and resulted in a new “Proxy Users” cluster. The use of the Messaging function was rarer among these individuals, resulting in the loss of the “Intense Digital Engagers” entirely. Throughout all trimesters during the COVID-19 pandemic, the “Schedulers” and “Average Users” remained the largest, while the “Proxy Users” and Results-oriented clusters remained the smallest. These trends continue, though not as clearly, when examining clusters among individuals with a normal pregnancy and individuals with a high-risk pregnancy.

Discussion

Principal results

We performed cross-sectional and temporal analyses that characterized outpatient portal use among pregnant individuals and identified profiles of outpatient portal user

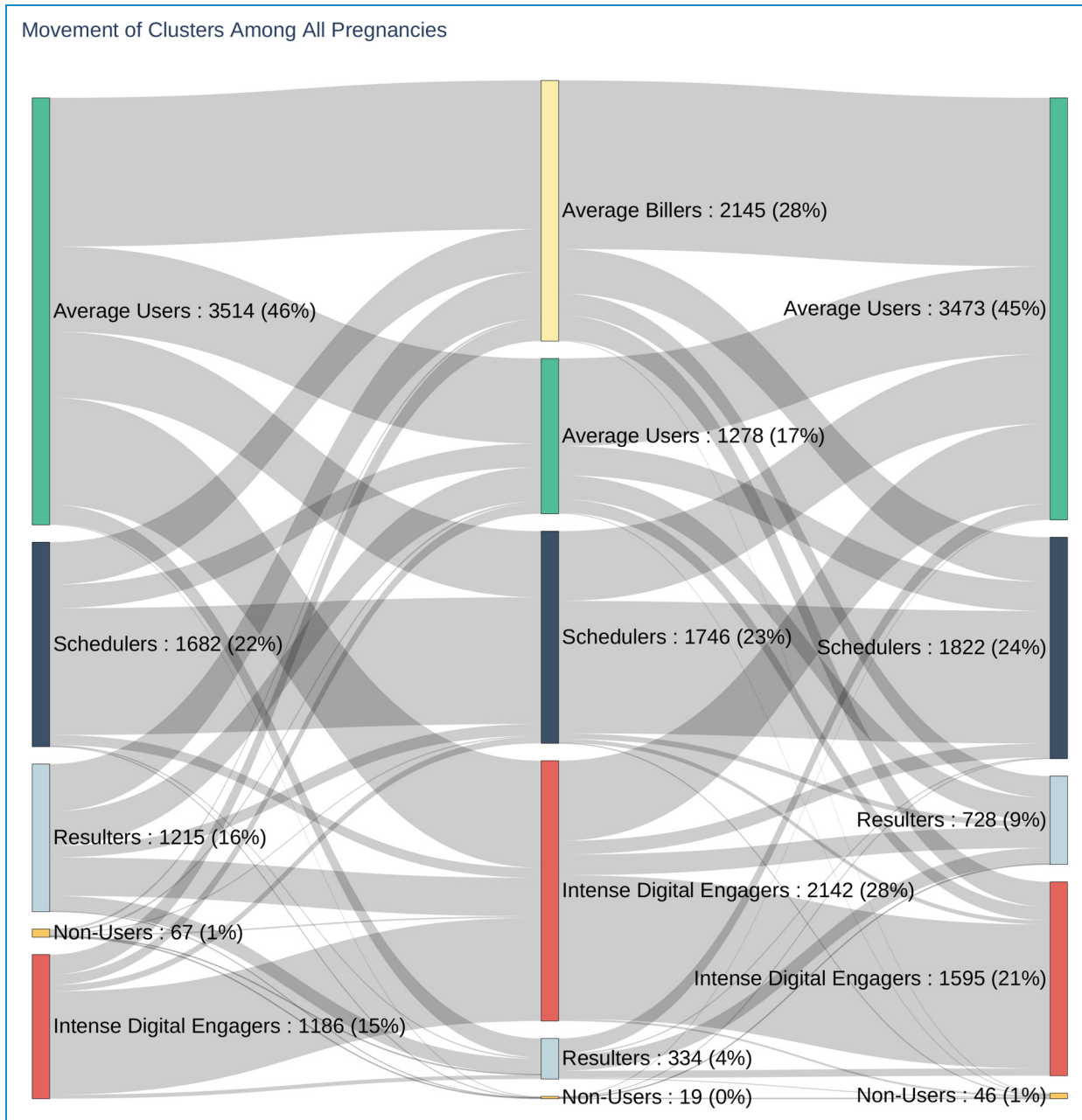


Figure 4. Movement of clusters among all pregnancies, during the first, second, and third trimester. Cluster names are listed with the number of women and percent of the total sample in each cluster.

groups based on the characteristics of outpatient portal use. We explored differences in the characteristics and clusters of outpatient portal use between individuals with normal and high pregnancy-related risks in order to discern how different patient populations engage with the outpatient portal.

Regarding the characterization of outpatient portal use, we discovered that pregnant individuals in our study predominantly used only four outpatient portal functions: MyRecord, Visits, Messaging, and Billing. These

individuals spent 55% of their sessions either using MyRecord to check PHI or managing appointments using Visits, while 20% of sessions were spent either contacting providers using Messaging or managing payments for care using Billing (Table 3). In addition, the median number of sessions during which an individual used an outpatient portal function plateaued between the second and third trimesters among all pregnant individuals, suggesting that they became more comfortable engaging with the outpatient portal midway through a pregnancy. In comparison

to Di Tosto et al., the patients in our study sample primarily used the same four outpatient portal functions, though the use of the Messaging function was far less common in our study.²⁵ It is yet to be determined why these samples differ so greatly in their usage of the Messaging function, and it could be associated with departmental implementation and cultural artifacts that may contribute to how the outpatient portal is used by patients and providers.^{10, 14, 15}

With respect to the second objective of our study, we were able to identify four unique clusters of outpatient portal users (“Schedulers,” “Resulters,” “Intense Digital Engagers,” and “Average Users”), though these were not entirely consistent across trimesters. These clusters can be used to guide and drive future hypothesis generation related to portal use and associated outcomes. Health care teams may also hope to identify such users within their patient population in order to provide more comprehensive, balanced outpatient portal training to address any gaps in learning among users.

Temporal change among the proportions of function use, which served as the foundation for clustering, was not uncommon, as every cluster experienced at least a 15% change in its sample of patients during the second and third trimesters (Appendix 1, Tables S3–S4). Individuals in our largest clusters were engaging with a variety of different outpatient portal functions over time. For instance, 45% of “Average Users” in the second trimester came from a different first-trimester cluster, and 79% of third-trimester “Average Users” came from a different cluster. These users may have been unfamiliar with the outpatient portal at baseline, and could benefit from reinforcing outpatient portal training early in their period of use in order to ensure broad engagement with the tool. Individuals who were “Intense Digital Engagers” during the second and third trimesters were among the most consistent outpatient portal users in terms of proportions of use, as these clusters frequently saw the least percent change within their sample of patients. The “Schedulers” also saw low turnover between the second and third trimesters. These results suggest that few users began their outpatient portal use with a strong baseline preference, and might greatly benefit from outpatient portal training and exposure. It is not clear if the preferences developed between the second and third trimesters are based on prior experience with the outpatient portal, nor what clinical implications may be associated with this type of outpatient portal use.

There were notable consistencies among certain clusters across the trimesters with regards to sessions of outpatient portal use per pregnancy and race-ethnicity of pregnant individuals. Across all trimesters, the Results- and Messaging-oriented clusters had a low number of sessions per pregnancy, while the “Schedulers” had high session counts per pregnancy (Figure 3; Appendix 1, Figures S2–S4). These clusters with more sessions per pregnancy had a higher percentage of non-Hispanic Black individuals,

and a lower percentage of non-Hispanic White individuals compared to the clusters that had fewer sessions of outpatient portal use per pregnancy. Our results indicate an association between a pregnant individual’s race-ethnicity and her use of the outpatient portal, such that a pregnant individual using the outpatient portal to schedule and monitor in-person appointments with her provider is more likely to use the outpatient portal more often and to be a non-Hispanic Black individual compared to the typical user. In contrast, a pregnant individual who uses the outpatient portal to check lab results and digitally message with her provider is more likely to use the outpatient portal less often and to be a non-Hispanic White individual compared to the typical user. These results support previous research that demonstrated African American/Black users look at their results in a patient portal less often compared to White users.³³ Future work could determine whether these racial disparities in outpatient portal use are driven by the health literacy of individuals in these clusters, which would add to research examining a digital divide among pregnant users of patient portals.¹⁵

Comparison with prior work

In comparison to previous studies utilizing a clustering algorithm to identify groups of patient portal users, our study identified only four clusters among the entire sample while Jones et al. profiled eight electronic patient portal users and Fareed et al. identified five clusters of inpatient portal users.^{26, 27} All studies were able to identify clusters of users who primarily used functions related to monitoring schedules and results, though the remaining clusters were quite different. For instance, since Fareed et al. defined clusters based on inpatient portal use, the remaining clusters were defined by the use of functions that were not available in the outpatient portal (e.g. “Getting Started” and “Dining on Demand” functions). These inpatient users likely engaged with this portal in a different manner given their proximity to clinical staff and hospital resources. While the clusters identified by Jones et al. were based on the use of an electronic patient portal that allowed users to perform similar functions (e.g. review labs, secure messaging, and manage appointments), they identified a greater variety of clusters that were typically smaller in size. Clusters unique to their study included those defined by their frequency and duration of use (i.e. “eDabblers” and “Infrequent, intense users”) and those who used unique settings offered within the portal (i.e. “Proxy moms” and “Record updaters”).

Unexpectedly, the frequency of portal function use was quite similar among normal pregnancy and high-risk pregnancy individuals (Table 3). However, this supports previous research that found no difference in the likelihood of engaging with a secure messaging portal when comparing individuals with and without a high-risk pregnancy,

which was defined using various clinical characteristics.¹⁶ In regards to our temporal analysis of outpatient portal use, high-risk pregnancy individuals demonstrated a significantly greater median outpatient portal function use per trimester when compared to normal pregnancy individuals (Figures 1 and 2). In contrast to the findings from Ukoha et al., these results, and those from a recent study³⁴ suggest that high-risk pregnancy individuals may engage with various outpatient portal functions more than their counterparts with a normal pregnancy over time. Therefore, it may be useful to consider the use of multiple outpatient portal functions when examining differences between pregnancy risk groups rather than looking at the use of a messaging-oriented function alone.

The primary difference between the clusters of pregnancy risk groups is that there were more clusters of outpatient portal users among high-risk pregnancy individuals than normal pregnancy individuals. In comparison to normal pregnancy individuals, high-risk pregnancy individuals may be more inquisitive in their use of the outpatient portal in that they engage with and explore the outpatient portal through diverse interactions with the outpatient portal functions, resulting in a wider variety of combined outpatient portal function uses. These individuals also use the portal more frequently according to median use per trimester. We have yet to determine why high-risk pregnancy individuals may use an outpatient portal differentially, though previous research indicates that the complexity of their pregnancy may intensify their inclination to engage with their health using the outpatient portal and diminish their perception of barriers to outpatient portal use.³⁵

Our results add to the foundation of research that suggests outpatient portal use is highly heterogeneous.^{26, 27} We were able to identify unique clusters of outpatient portal users among an understudied population, based on differences in proportions of outpatient portal function use, median outpatient portal use per trimester and patient characteristics such as race and ethnicity. Future studies might benefit from our work if they intend to explore whether such patterns of outpatient portal use could influence the patient quality and clinical outcomes. For instance, it could be helpful to know if obstetrics and gynecology or maternal–fetal medicine teams are retaining individuals who primarily use the Visits function more than individuals in other clusters. In addition, providers might be interested in whether medication adherence is greater among high-risk pregnancy individuals who frequently track their lab results using the MyRecord function compared to those who check these infrequently. Such knowledge could influence health care teams and centers to provide further training on outpatient portal use for both providers and patients. It would also be beneficial for future work to examine whether there are qualitative themes of outpatient portal use, rather than quantitative alone, that might help expand our understanding of outpatient portals as a beneficial tool

in the effort to improve quality and possibly clinical outcomes.

Limitations

Our study is not without limitations. First, the generalizability of our study is somewhat limited given that it was performed using data from users of the single outpatient portal at a single academic medical center. Nevertheless, our work adds to the larger picture of outpatient portal use among pregnant individuals. Next, a portion of our study timeframe overlapped with the COVID-19 pandemic, which forced our academic medical center to rapidly transition the majority of their care to the outpatient setting with the understanding that the outpatient portal would be used more frequently to engage with patients. We performed a final cluster analysis among individuals who only used the outpatient portal between 1/1/2020 and 8/1/2020 in order to determine how clusters of outpatient portal use during the COVID-19 pandemic differed from those in our original sample. Second, we only utilized one clustering algorithm (i.e. hierarchical clustering) to identify groups of outpatient portal users. While *k*-means clustering is a commonly employed algorithm, we were not able to use this method given our lack of *a priori* information regarding the types of outpatient portal users within this patient population. Given the clusters we have now identified, future research could explore the optimal clustering algorithm for use on this data set. Finally, the number of individuals with a normal pregnancy may be inflated as there could be individuals who were referred to but did not visit with maternal–fetal medicine for a suspected high-risk. Future studies may be able to identify additional individuals with a high-risk pregnancy if they can successfully track referrals and subsequent visits.

Conclusion

In conclusion, characterizing outpatient portal use and identifying profiles of outpatient portal user groups among pregnant individuals demonstrates how specific patient populations might engage with their health. We were able to identify distinct cluster groups of outpatient portal users among pregnancy groups, which underscores the importance of avoiding the use of generalizations when describing how patients in this population of pregnant individuals might engage with patient-facing technologies such as an outpatient portal. These results can be used to improve user experience and training with outpatient portal functions and may educate obstetrics and gynecology and maternal–fetal medicine providers on patient engagement with the outpatient portal.

Ethical approval: This study did not require approval from the Institutional Review Board of The Ohio State University

Medical Center as it involved the use of a coded-limited data set reviewed and approved by the Honest Broker Committee (REQ674151) of The Ohio State University Medical Center.

Guarantor: NF

Contributorship: EM, PS, and NF researched literature and conceived the study. EM, PS, PS, and NF contributed to the conceptualization and drafting of the analysis. EM, PS, and NF were involved in gaining data set approval. EM, PS, and NF performed data collection and data analysis. EM wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgments: We would like to thank Shonda Vink of The Ohio State University Medical Center for her assistance in this research.

Author contributions:

Declaration of Conflicting Interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Informed consent: Not applicable, because this article does not contain any studies with human or animal subjects.

ORCID iD: Naleef Fareed  <https://orcid.org/0000-0001-8764-1625>

Trial registration: Not applicable, because this article does not contain any clinical trials.

Supplemental material: Supplemental material for this article is available online.

References

1. Irizarry T, DeVito Dabbs A and Curran CR. Patient portals and patient engagement: A state of the science review. *J Med Internet Res* 2015; 17: e148.
2. What are the differences between electronic medical records, electronic health records, and personal health records? *FAQ*, <https://www.healthit.gov/faq/what-are-differences-between-electronic-medical-records-electronic-health-records-and-personal-2021>.
3. Kuo A and Dang S. Secure messaging in electronic health records and its impact on diabetes clinical outcomes: A systematic review. *Telemed E-Health* 2016; 22: 769–777.
4. Amante DJ, Hogan TP, Pagoto SL, et al. A systematic review of electronic portal usage among patients with diabetes. *Diabetes Technol Ther* 2014; 16: 784–793.
5. Tang PC, Overhage JM, Chan AS, et al. Online disease management of diabetes: Engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. *J Am Med Inform Assoc JAMIA* 2013; 20: 526–534.
6. Osborn CY, Mayberry LS, Mulvaney SA, et al. Patient web portals to improve diabetes outcomes: A systematic review. *Curr Diab Rep* 2010; 10: 422–435.
7. Krist AH, Woolf SH, Rothemich SF, et al. Interactive preventive health record to enhance delivery of recommended care: A randomized trial. *Ann Fam Med* 2012; 10: 312–319.
8. Chrischilles EA, Hourcade JP, Doucette W, et al. Personal health records: A randomized trial of effects on elder medication safety. *J Am Med Inform Assoc JAMIA* 2014; 21: 679–686.
9. Devkota B, Salas J, Sayavong S, et al. Use of an online patient portal and glucose control in primary care patients with diabetes. *Popul Health Manag* 2016; 19: 125–131.
10. Shaw E, Howard M, Chan D, et al. Access to web-based personalized antenatal health records for pregnant women: A randomized controlled trial. *J Obstet Gynaecol Can* 2008; 30: 38–43.
11. Forster M, Dennison K, Callen J, et al. Maternity patients' access to their electronic medical records: Use and perspectives of a patient portal. *Health Inf Manag J* 2015; 44: 4–11.
12. Wieland D, Gibeau A, Dewey C, et al. Patient portal readiness among postpartum patients in a safety net setting. *Appl Clin Inform* 2017; 08: 698–709.
13. Peacock S, Reddy A, Leveille SG, et al. Patient portals and personal health information online: Perception, access, and use by US adults. *J Am Med Inform Assoc* 2017; 24: e173–e177.
14. Ukoha EP and Yee LM. Use of electronic patient portals in pregnancy: An overview. *J Midwifery Womens Health* 2018; 63: 335–339.
15. Kim J, Mathews H, Cortright LM, et al. Factors affecting patient portal use among low-income pregnant women: Mixed-methods pilot study. *JMIR Form Res* 2018; 2: e6.
16. Ukoha EP, Feinglass J and Yee LM. Disparities in electronic patient portal use in prenatal care: Retrospective cohort study. *J Med Internet Res* 2019; 21: e14445.
17. Ammenwerth E, Schnell-Inderst P and Hoerbst A. The impact of electronic patient portals on patient care: A systematic review of controlled trials. *J Med Internet Res* 2012; 14: e162.
18. Fareed N, MacEwan S, Vink S, et al. Relationships between patient portal account status and patient satisfaction scores. *Am J Manag Care*. 2022; 28: 25–31.
19. Goldzweig CL, Orshansky G, Paige NM, et al. Electronic patient portals: Evidence on health outcomes, satisfaction, efficiency, and attitudes: A systematic review. *Ann Intern Med* 2013; 159: 677–687.
20. Ten Klooster I, Noordzij ML and Kelders SM. Exploring how professionals within agile health care informatics perceive visualizations of log file analyses: Observational study followed by a focus group interview. *JMIR Hum Factors* 2020; 7: e14424.
21. Rule A, Chiang MF and Hribar MR Using electronic health record audit logs to study clinical activity: A systematic review of aims, measures, and methods. *J Am Med Inform Assoc* 2020; 27: 480–490.
22. Durojaiye AB, Levin S, Toerper M, et al. Evaluation of multidisciplinary collaboration in pediatric trauma care using EHR data. *J Am Med Inform Assoc JAMIA* 2019; 26: 506–515.

23. Sinsky CA, Rule A, Cohen G, et al. Metrics for assessing physician activity using electronic health record log data. *J Am Med Inform Assoc JAMIA* 2020; 27: 639–643.
 24. Huerta T, Fareed N, Hefner JL, et al. Patient engagement as measured by inpatient portal use: Methodology for log file analysis. *J Med Internet Res* 2019; 21: e10957.
 25. Di Tosto G, McAlearney AS, Fareed N, et al. Metrics for outpatient portal use based on log file analysis: Algorithm development. *J Med Internet Res* 2020; 22: e16849.
 26. Fareed N, Walker D, Sieck CJ, et al. Inpatient portal clusters: Identifying user groups based on portal features. *J Am Med Inform Assoc* 2019; 26: 28–36.
 27. Jones JB, Weiner JP, Shah NR, et al. The wired patient: Patterns of electronic patient portal use among patients with cardiac disease or diabetes. *J Med Internet Res* 2015; 17: e42.
 28. The Ohio State University Wexner Medical Center. High Risk Pregnancy, <https://wexnermedical.osu.edu/obstetrics-gynecology/high-risk-pregnancy>. (accessed 11 April 2020).
 29. How is the length of my pregnancy measured? *How Your Fetus Grows During Pregnancy*, <https://www.acog.org/womens-health/faqs/how-your-fetus-grows-during-pregnancy>, 2021.
 30. Hair JF. *Multivariate data analysis*. Eighth edition. Andover, Hampshire: Cengage, 2019.
 31. Calinski T and Harabasz J. A dendrite method for cluster analysis. *Commun Stat - Theory Methods* 1974; 3: 1–27.
 32. Duda RO, Hart PE and Stork DG. *Pattern classification*. 2nd ed. New York: Wiley, 2001.
 33. Foster B and Krasowski MD. The use of an electronic health record patient portal to access diagnostic test results by emergency patients at an academic medical center: Retrospective study. *J Med Internet Res* 2019; 21: e13791.
 34. Singh P, Jonnalagadda P, Morgan E, et al. Outpatient portal use in prenatal care: Differential use by race, risk, and area social determinants of health factors. *J Am Med Inf Assoc*. 2022; 29: 364–371.
 35. Powell KR. Patient-Perceived facilitators of and barriers to electronic portal use: A systematic review. *Comput Inform Nurs CIN* 2017; 35: 565–573.
-