

Article

Uncovering eHealth Engagement Patterns Through Latent Class Analysis and SHAP: A Data Mining Perspective on Telehealth Access

Ning Yang ^{1,*}  and Xin Yang ²

¹ Information Systems and Supply Chain Management, Quinlan School of Business, Loyola University Chicago, 16 E Pearson St, Chicago, IL 60611, USA

² Department of Communicative Disorders, The University of Alabama, 739 University Blvd, Tuscaloosa, AL 35487, USA

* Correspondence: nyang1@luc.edu

Abstract

Understanding how patients engage with digital health technologies is critical for improving the reach and equity of telehealth services. While prior research has largely focused on demographic predictors of telehealth use, this study applies a hybrid data mining approach to uncover behavioral engagement patterns and evaluate their predictive power. Using data from the 2022 U.S. Health Information Trends Survey (HINTS; N = 3525), we identified four distinct eHealth engagement typologies through Latent Class Analysis (LCA): (1) Highly Digital Engagers, (2) Moderate Digital Users, (3) Social Media and App Enthusiasts, and (4) Wearable and Health App Enthusiasts. We then modeled telehealth utilization as the outcome using multivariable logistic regression and eXtreme Gradient Boosting (XGBoost) with Shapley Additive Explanations (SHAP). Compared to Highly Digital Engagers, Moderate Digital Users had significantly lower odds of telehealth use (OR = 0.52), while the other two classes had higher odds. SHAP analyses confirmed that depression status and geographic region interacted with engagement profiles to shape telehealth access, with a notably negative effect of depression within Class 2. These findings demonstrate the value of integrating behavioral segmentation with interpretable machine learning to characterize heterogeneity in digital health engagement and its association with telehealth utilization. Our study offers a scalable, population-level analytic framework that can inform targeted telehealth planning and outreach strategies aligned with real-world patterns of digital engagement.

Keywords: eHealth engagement; telehealth; latent class analysis; latent class analysis; XGBoost; SHAP; healthcare data mining; digital inclusion



Academic Editor: Ognjen Arandjelović

Received: 26 January 2026

Revised: 12 February 2026

Accepted: 16 February 2026

Published: 19 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The growing integration of digital technologies in healthcare has enabled transformative innovations in how care is accessed, delivered, and experienced. Among these, telehealth, the remote provision of healthcare services via telecommunications technologies, has emerged as a scalable solution for expanding care delivery. Extensive empirical research has demonstrated the clinical and operational effectiveness of telehealth in both primary care and chronic disease management settings, improving outcomes, access, and continuity of care across diverse populations [1,2]. By reducing geographic and temporal barriers, telehealth has demonstrated considerable value in lowering costs, increasing

patient convenience, and improving healthcare accessibility [3–6]. Despite this potential, its uptake has been uneven, constrained by clinician resistance [7], policy barriers [8], infrastructure gaps [9], and persistent disparities in digital access [10]. The COVID-19 pandemic catalyzed rapid telehealth adoption, but post-pandemic sustainability demands a deeper understanding of underlying digital engagement behaviors [11].

Prior work conceptualizes telehealth and telemedicine as part of a broader digital health ecosystem that encompasses overlapping eHealth and mHealth technologies, rather than as isolated modes of care delivery [12,13]. These activities extend beyond individual tools to include the convergence of digital technologies (like AI, genomics, and big data) with health, healthcare, living, and society [14]. While digital health represents the overarching technological transformation of care, eHealth specifically describes the engagement behaviors of patients within this system. eHealth engagement, defined here as the consumer's active use of digital tools for information seeking, health monitoring, and self-management, has been linked to improved chronic disease management, better adherence to care plans, and enhanced health self-efficacy [15–17]. Although eHealth engagement and synchronous telehealth interactions are conceptually distinct, prior work suggests that they rely on shared enabling resources, including digital literacy, access to devices and broadband connectivity, and familiarity with digital health interfaces, which shape individuals' capacity to engage in virtual care [12,13,18]. For example, individuals who were identified with low health and digital literacy, disabilities, and other underserved characteristics were at heightened risk of telehealth exclusion [13]. Similarly, Ch et al. [18] show that telemedicine utilization is strongly shaped by socioeconomic status through disparities in access to high-speed internet, smart devices, and digital skills. Rush et al. [19] identify digital literacy as a primary barrier to telehealth adoption and situate telehealth within broader eHealth strategies that require familiarity with multiple digital tools. Although this body of work documents substantial disparities in access and adoption, it rarely examines whether and how engagement with eHealth tools translates into telehealth utilization differently for vulnerable populations. As Phuong et al. [20] argue, policies and regulatory contexts can create or constrain pathways to telehealth, suggesting that the relationship between digital engagement and telehealth use is context-dependent rather than universal.

While prior studies have focused on individual predictors such as age, education, and digital literacy [21,22], such unidimensional profiling may fail to capture the behavioral complexity that underpins digital health engagement. Individuals may exhibit diverse combinations of digital behaviors, some highly active in social health forums but disengaged from formal patient portals, others using wearable devices but avoiding video consultations. These heterogeneous behavioral patterns require more sophisticated analytical approaches to uncover behavioral clusters that transcend simplistic demographic categorization. To address this complexity, Latent Class Analysis (LCA) offers a robust statistical framework for identifying unobserved (latent) subgroups based on distinct patterns of categorical responses [23].

In the broader informatics and population health domains, data-driven patient stratification has gained traction as a means of tailoring interventions to specific risk profiles. From chronic disease management to treatment optimization, unsupervised learning methods such as clustering and latent class analysis (LCA) have been used to uncover phenotypes or behavioral subgroups with clinical relevance [23,24]. Yet, such stratification approaches are rarely applied to digital engagement behaviors, despite their growing importance in health service delivery.

At the same time, explainable artificial intelligence (XAI) is emerging as a critical innovation in healthcare analytics, addressing the need for transparency and interpretability

in complex predictive models. Traditional machine learning techniques, though powerful, are often perceived as black boxes, limiting their adoption in domains requiring stakeholder trust and regulatory accountability [25–27]. SHAP (Shapley Additive Explanations), a widely adopted post hoc explanation method, has proven particularly useful in healthcare contexts by quantifying the relative contribution of features at both the global and local levels [26,28]. These tools allow for the generation of interpretable, stakeholder-relevant insights, enabling researchers, clinicians, and informatics professionals to make sense of model outputs in ways that support decision-making across different user roles.

This study responds to calls for transparent, behaviorally grounded, and population-scalable approaches to understanding telehealth utilization. It addresses two questions: (1) What behavioral clusters of eHealth engagement exist in the U.S. adult population? (2) How do these behavioral profiles, along with sociodemographic features, predict telehealth utilization? Using data from the 2022 U.S. Health Information National Trends Survey (HINTS), we develop a hybrid modeling pipeline that integrates unsupervised and supervised learning. First, we apply Latent Class Analysis to identify distinct typologies of eHealth engagement across seven digital behavior dimensions. We predict telehealth use based on individuals' digital engagement types, using both logistic regression and XGBoost models. To enhance interpretability, we apply SHAP analysis to quantify which behavior patterns and user characteristics most strongly influence telehealth adoption across subgroups.

By combining segmentation with explainable predictive modeling, this research contributes to the growing field of behavioral health informatics [29,30] and demonstrates how interpretable machine learning can enhance digital inclusion strategies. The approach offers a scalable framework for detecting vulnerable subgroups, informing targeted outreach, and aligning telehealth system design with real-world behavioral profiles, ultimately supporting more equitable and personalized digital health delivery.

2. Methods

Data for this study were drawn from the 2022 U.S. Health Information National Trends Survey (HINTS), Cycle 6. HINTS is a nationally representative survey administered by the U.S. National Cancer Institute (NCI) to collect population-level data on health communication, digital health behaviors, and technology use among non-institutionalized U.S. adults aged 18 and older. The survey employs a stratified random sampling design based on census tract characteristics, with oversampling of high-minority and underserved populations to enhance subgroup precision.

HINTS has been widely used in health informatics research to examine digital health disparities [31–33], health information-seeking behaviors [33–35], and eHealth adoption trends [36,37]. Its breadth of behavioral, sociodemographic, and technological indicators makes it especially valuable for behavioral segmentation and predictive modeling efforts. Building on this foundation, our study leverages HINTS to identify latent eHealth engagement typologies and examine their predictive value for telehealth utilization using both traditional and machine learning approaches.

2.1. Dependent Variables

Telehealth utilization was defined by the question: “In the past 12 months, did you receive care from a doctor or health professional using telehealth?”, Respondents who answered “Yes”, were classified as telehealth users, regardless of modality (video, phone call, or both), while those who answered “No”, were classified as non-users.

2.2. Independent Variables

Independent variables of interest included seven composite scores related to eHealth engagement that were derived from 14 HINTS items. These scores included eHealth Information Access (EHIA), eHealth System Interaction (EHSI), mHealth Engagement (mHE), Social Media Health Decision and Discussion Engagement (SMDDE), Social Media Health Content Engagement (SMHCE), Health Information Sharing on Social Media (HISSM), and Wearable Device Tracking Health (WDTH) (see Table 1). Supplementary Table S1 provides a detailed breakdown of the HINTS questions included in each composite score. Composite scores were calculated by: (1) Responses to each question were converted to a standardized 5-point Likert scale as follows: “Almost every day”, = 5, “At least once a week”, or “Strongly agree”, = 4, “A few times a month”, or “Somewhat agree”, = 3, “Less than once a month”, or “Somewhat disagree”, = 2, and “Never”, or “Strongly disagree”, = 1; (2) averaging items within each composite; and (3) categorizing scores into “low”, “moderate”, or “high”, based on distribution percentiles (detailed in Supplementary Table S2). The categorization of composite scores into “low”, “moderate”, and “high”, was based on tertiles of the distribution for each composite score.

Table 1. Operationalization of digital health engagement constructs based on U.S. HINTS survey items.

Abbreviation	Name of Variables	Questionnaire Items and Variable ID in HINTS	Description
EHIA	eHealth Information Access	1. In the past 12 months have you used the Internet to look for health or medical information? (Electronic2_HealthInfo: B3a.) 2. In the past 12 months have you used the Internet to view medical test results? (Electronic2_TestResults B3c.)	Assesses an individual's ability to access online health-related resources.
EHSI	eHealth System Interaction	1. In the past 12 months have you used the Internet to look for health or medical information? (Electronic2_MadeAppts: B3d.) 2. In the past 12 months have you used the Internet to send a message to a health care provider or health care providers office?(Electronic2_MessageDoc: B3b)	Focuses on how individuals apply their online skills to engage in healthcare activities.
mHE	mHealth Engagement	1. In the past 12 months, have you used a health or wellness app on your tablet or smartphone? (UsedHealthWellnessApps2: B7) 2. Have you ever used an app like Apple Health Records or CommonHealth to combine your medical information from different patient portals or online medical records into one place? (UsedPortalOrganizerApp: E9)	Measures the usage of mobile health (mHealth) apps.
SMDDE	Social Media Health Decision and Discussion Engagement	1. How much do you agree or disagree—I use information from social media to make decisions about my health. (SocMed_MakeDecisions: B14a.) 2. How much do you agree or disagree—I use information from social media in discussions with my health care provider. (SocMed_DiscussHCP: B14b.)	Reflects both personal decision-making and professional consultations influenced by social media.
SMHCE	Social Media Health Content Engagement	B12e. In the last 12 months, how often did you watch a health-related video on a social media site (for example, YouTube)? (SocMed_WatchedVid: B12e.) B12d. In the last 12 months, how often did you interact with people who have similar health or medical issues on social media or online forums? (SocMed_Interacted: B12d.)	Reflects both passive content consumption and active engagement on social media.

Table 1. Cont.

Abbreviation	Name of Variables	Questionnaire Items and Variable ID in HINTS	Description
HISSM	Health Information Sharing on Social Media	1. In the last 12 months, how often did you share general health-related information on social media (for example, a news article)? (SocMed_SharedGen: B12c.) 2. In the last 12 months, how often did you share personal health information on social media? (SocMed_SharedPers: B12b.)	Focuses on the sharing behavior of individuals regarding health information on social media.
WDTH	Wearable Device Tracking Health	1. In the last 12 months, have you used an electronic wearable device to monitor or track health or activity? For example, a Fitbit, AppleWatch or Garmin Vivofit. (WearableDevTrackHealth: B8) 2. In the past month, how often did you use a wearable device to track your health? (FreqWearDevTrackHealth: B9.)	Measures the use of wearable devices to monitor or track health or activity.

The categorization reflects a deliberate analytic tradeoff between measurement precision and interpretability. While discretization necessarily reduces within-category variance, this approach aligns with the study’s population-level and policy-oriented objectives, where actionable stratification often relies on interpretable groupings rather than fine-grained psychometric distinctions [24]. In the context of national survey data such as HINTS, tertile-based categorization has been widely used to support segmentation, comparability across heterogeneous indicators, and downstream modeling stability, particularly when integrating unsupervised clustering with predictive analytics [37]. This method was chosen to ensure that each engagement dimension was consistently evaluated across similar relative levels, regardless of differences in scaling or score distributions.

To adjust for confounding variables, we also include covariates in the models, such as race (Non-Hispanic White, Non-Hispanic Black, Asian and Other, Hispanic), education (High School Graduate or lower, Some College, College Graduate or higher), household income (Less than \$20,000, \$20,000 to \$35,000, \$35,000 to \$75,000, \$75,000 or More), gender at birth (male, female), age group (18–34, 35–49, 50–64, 65–74, 75+), census region (Midwest, Northeast, South, West), depression status (Yes, No), lack of transportation for medical care (Yes, No), chronic conditions such as diabetes, hypertension and heart conditions, frequency to visit doctors in the past 12 months, and Health insurance status (yes, no).

2.3. Statistical Analysis

Descriptive statistics, including frequencies and weighted percentages, were used to summarize the sample characteristics for each category within variables. These statistics were also stratified by telehealth utilization to examine group differences. Associations between categorical variables and telehealth use were tested using chi-squared tests. To categorize individuals based on their digital health engagement patterns and identify hidden subgroups with similar behaviors, we conducted Latent Class Analysis (LCA) on the seven eHealth engagement variables using the “poLCA”, package in R software version 4.4.1 [38]. The selection of four latent classes was informed by several model fit indices and the interpretability of the resulting subgroups (Supplementary Table S2, Figure 1). The four-class model provided the optimal balance between model fit and parsimony, demonstrating the lowest Bayesian Information Criterion (BIC = 35,094) compared to 3-class and 5-class alternatives. While the Akaike information criterion (AIC) continued to decrease with additional classes, the BIC plateaued at the four-class solution, supporting its selection.

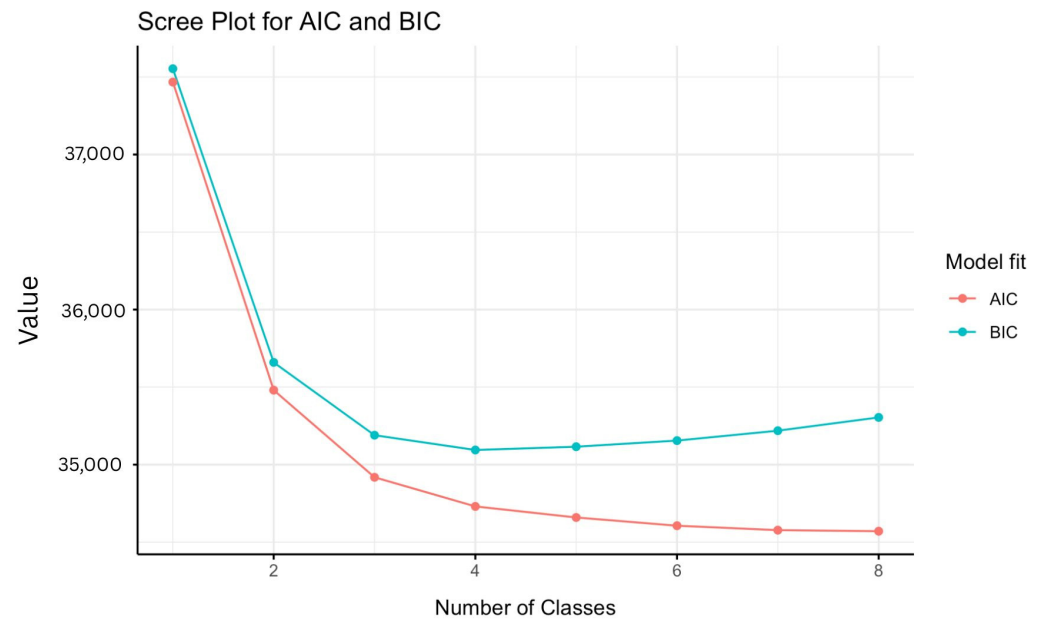


Figure 1. Scree plot for evaluating model fit in latent class analysis to determine the optimal number of classes.

Furthermore, the model demonstrated strong classification quality. The entropy was 0.72, and the Average Posterior Probabilities (AvePP) for assignment to each class ranged from 0.80 to 0.86, exceeding the recommended 0.70 threshold. This confirms that the identified groups—Highly Digital Engagers, Moderate Digital Users, Social Media Enthusiasts, and Wearable Enthusiasts—are statistically distinct and well-separated, aligning with theoretical expectations of varying digital health behaviors.

The resulting classes were not driven by uniform engagement intensity across all dimensions, but instead reflected distinct combinations of behaviors, such as high system interaction with low social media use, suggesting that the identified classes capture meaningful behavioral configurations. The selection of the four-class model was supported by multiple fit indices (Supplementary Table S2). The four-class solution demonstrated the lowest Bayesian Information Criterion (BIC = 35,094), indicating optimal fit relative to model complexity. Furthermore, the classes showed strong distinctiveness, with an entropy of 0.72. The Average Posterior Probabilities (AvePP) for assignment to each class ranged from 0.80 to 0.86, exceeding the recommended 0.70 threshold and confirming that individuals were reliably assigned to well-separated groups.

Multiple logistic regression models were employed to assess associations between telehealth utilization and key predictors, including race, education, household income, urbanicity, census region, transportation barriers to medical care, depression status, and eHealth engagement class membership. Model performance was validated using 5-fold cross-validation, yielding an average Area Under the Curve (AUC) of 0.697. To complement and validate findings from logistic regression, an eXtreme Gradient Boosting (XGBoost) model—a high-performing machine learning algorithm—was also developed using the same set of predictors, achieving a 5-fold cross-validated AUC of 0.852. To interpret the results from the XGBoost model, Shapley Additive Explanations (SHAP) values were calculated, providing insight into variable importance and contribution to the model's predictions [39]. SHAP values represent the contribution of each feature to the model's prediction. Positive SHAP values indicate an increase in the predicted probability of telehealth use, while negative values indicate a decrease. The overall analytical framework, illustrating the data processing pipeline from feature engineering and latent class identifi-

cation to predictive modeling and interpretation, is presented in Supplementary Figure S2. To ensure rigorous validation, predictive performance (AUC) was evaluated using 5-fold cross-validation. To maximize statistical power for the variable importance and SHAP interaction analyses, the final model was retrained on the full dataset using the hyperparameters validated during the cross-validation step. This approach ensures that our reported performance metrics are unbiased (preventing overfitting) while our interpretation of factors utilizes the complete information available in the dataset. A p -value of <0.05 was considered statistically significant. All analyses were conducted using R statistical software [40].

3. Results

Table 2 summarizes the characteristics of 3525 participants, with inclusion and exclusion details in Supplementary Figure S1. Most were aged 18–65 years (87.1%), Non-Hispanic White (62.7%), had some college education or higher (77.7%), a household income $\geq \$75,000$ (29%), and resided in the South (39%). Table 3 shows that eHealth engagement was predominantly low for SMDDE, SMHCE, HISSM, and WDTN ($\geq 60\%$), while over half had high engagement in EHSI and EHIA. For mHE, moderate engagement was slightly less common than low or high levels. Telehealth use varied significantly by census region, depression status, and all eHealth engagement variables except SMDDE ($p < 0.05$ for all significant differences in Table 3).

Table 2. Patient-level sociodemographic characteristics of study population by telehealth use ($N = 3525$).

	All		TeleHealth = Yes		TeleHealth = No		p
	n	% (SE)	n	% (SE)	n	% (SE)	
Age							
18–34	713	30.7 (1.3)	301	26.2 (2.2)	412	34 (1.6)	0.024 *
35–49	912	29.3 (1.3)	438	32.4 (2.1)	474	27 (1.7)	
50–64	1045	27.1 (1.1)	474	27.5 (2)	571	26.8 (1.4)	
65–74	611	8.9 (0.5)	280	9.3 (0.8)	331	8.6 (0.7)	
75+	244	4 (0.3)	113	4.6 (0.6)	131	3.5 (0.4)	
Gender							
Female	2137	50.6 (0.7)	1043	57.9 (1.7)	1094	45.1 (1.4)	<0.001 ***
Male	1388	49.4 (0.7)	563	42.1 (1.7)	825	54.9 (1.4)	
Race							
NHW	2095	62.7 (0.9)	968	65.3 (1.9)	1127	60.7 (1.3)	0.303
AO	336	11.3 (0.6)	138	10.7 (1.4)	198	11.7 (0.9)	
HP	575	15.8 (0.6)	281	15.1 (1.2)	294	16.3 (1.2)	
NHB	519	10.3 (0.5)	219	8.9 (0.9)	300	11.3 (0.8)	
Education							
High School or Less	578	22.3 (1.1)	216	19.9 (1.5)	362	24.1 (1.7)	0.09
Some College	985	39.2 (1.2)	437	38.5 (1.7)	548	39.7 (1.9)	
College Graduate or More	1962	38.5 (0.7)	953	41.6 (1.7)	1009	36.3 (1.3)	
Household income							
Less than \$20,000	385	10.2 (1.1)	174	10.4 (1.4)	211	10 (1.6)	0.345
\$20,000 to <\$35,000	402	9.4 (0.7)	164	9.5 (1.2)	238	9.3 (0.9)	
\$35,000 to <\$75,000	1072	29 (1.3)	465	26.1 (1.8)	607	31.2 (2)	
\$75,000 or More	1666	51.4 (1.3)	803	54 (2.1)	863	49.5 (2.2)	
Region							
Northeast	499	16.5 (0.7)	243	19.3 (1.4)	256	14.4 (1.1)	<0.001 ***
Midwest	624	21.2 (0.7)	223	16.4 (1.3)	401	24.8 (1)	
South	1578	39 (0.8)	697	37.7 (1.8)	881	40 (1.5)	
West	824	23.3 (0.7)	443	26.6 (1.6)	381	20.8 (1.2)	

Table 2. *Cont.*

	All		TeleHealth = Yes		TeleHealth = No		<i>p</i>
	<i>n</i>	% (SE)	<i>n</i>	% (SE)	<i>n</i>	% (SE)	
Depression							
No	2479	69.4 (1.5)	971	56.8 (2.1)	1508	78.8 (1.8)	<0.001 ***
Yes	1046	30.6 (1.5)	635	43.2 (2.1)	411	21.2 (1.8)	
Lack Transportation							
No	3132	87.5 (1.3)	1403	85.4 (1.5)	1729	89 (1.8)	0.116
Yes	393	12.5 (1.3)	203	14.6 (1.5)	190	11 (1.8)	

Note: NHW, Non-Hispanic White; NHB, Non-Hispanic Black; HP, Hispanic; OA, Other and Asian.
 * $p < 0.05$, *** $p < 0.001$.

Table 3. Proportion of Sample Engaged in Digital Health and Social Media Activities by Telehealth Use.

		All		TeleHealth = Yes		TeleHealth = No		<i>p</i>
		<i>n</i>	% (SE)	<i>n</i>	% (SE)	<i>n</i>	% (SE)	
SMDDE	Low	2855	81.4 (1.3)	1277	80.7 (1.6)	1578	82 (1.9)	0.78
	Moderate	307	8.7 (0.8)	144	8.6 (1)	163	8.8 (1.3)	
	High	363	9.9 (1)	185	10.6 (1.2)	178	9.3 (1.6)	
SMHCE	Low	2715	76.5 (1.4)	1153	69.8 (2)	1562	81.5 (1.4)	<0.001 ***
	Moderate	609	17.5 (1)	319	21.1 (1.5)	290	14.8 (1.2)	
	High	201	6 (0.8)	134	9.1 (1.2)	67	3.7 (1.1)	
HISSM	Low	3221	90.9 (0.9)	1432	87.7 (1.4)	1789	93.2 (1.1)	0.007 **
	Moderate	221	6.3 (0.7)	124	7.7 (1)	97	5.2 (1.2)	
	High	83	2.8 (0.6)	50	4.6 (1.4)	33	1.5 (0.3)	
EHSI	Low	722	21.1 (1.2)	168	10.6 (1.2)	554	29 (1.8)	<0.001 ***
	Moderate	769	22.8 (1.4)	278	17.8 (1.7)	491	26.5 (2)	
	High	2034	56.1 (1.7)	1160	71.6 (1.8)	874	44.5 (2.5)	
EHIA	Low	248	8.9 (0.9)	58	4.8 (0.9)	190	12 (1.4)	<0.001 ***
	Moderate	738	21.1 (1.1)	228	14.8 (1.5)	510	25.8 (1.9)	
	High	2539	70 (1.4)	1320	80.4 (1.6)	1219	62.2 (2.4)	
mHE	Low	1294	37.3 (1.5)	429	26.9 (1.7)	865	45.2 (2.1)	<0.001 ***
	Moderate	980	26.3 (1.5)	574	34.3 (2.1)	406	20.3 (2.1)	
	High	1251	36.4 (1.5)	603	38.8 (2)	648	34.5 (2.2)	
WDTH	Low	2149	60 (1.4)	933	56 (2.1)	1216	62.9 (1.7)	0.040 *
	Moderate	277	8.5 (0.9)	134	10.3 (1.4)	143	7.2 (1.2)	
	High	1099	31.5 (1.3)	539	33.7 (2.1)	560	29.9 (1.5)	

Note: SMDDE, Social Media Health Decision and Discussion Engagement; SMHCE, Social Media Health Content Engagement; HISSM, Health Information Sharing on Social Media; EHSI, eHealth System Interaction; EHIA, eHealth Information Access; mHE, mHealth Engagement; WDTH, Wearable Device Tracking Health. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Based on the LCA results, the study population was best described by four distinct eHealth engagement (EHE) classes. Figure 2 presents the distribution of EHE levels (Low, Moderate, High) across these classes. Class 1, “Highly Digital Engagers”, actively use digital tools for health information and system interaction but engage less with social media. Class 2, “Moderate Digital Users”, utilize online health resources and mobile apps moderately, with minimal engagement in social media or wearable devices. Class 3, “Social Media and App Enthusiasts”, are highly engaged across digital health platforms, including social media, mobile health apps, and wearable technology, actively sharing and consuming

health information. Class 4, “Wearable and Health App Enthusiasts”, engage extensively with online health platforms and wearable devices but rely less on social media.

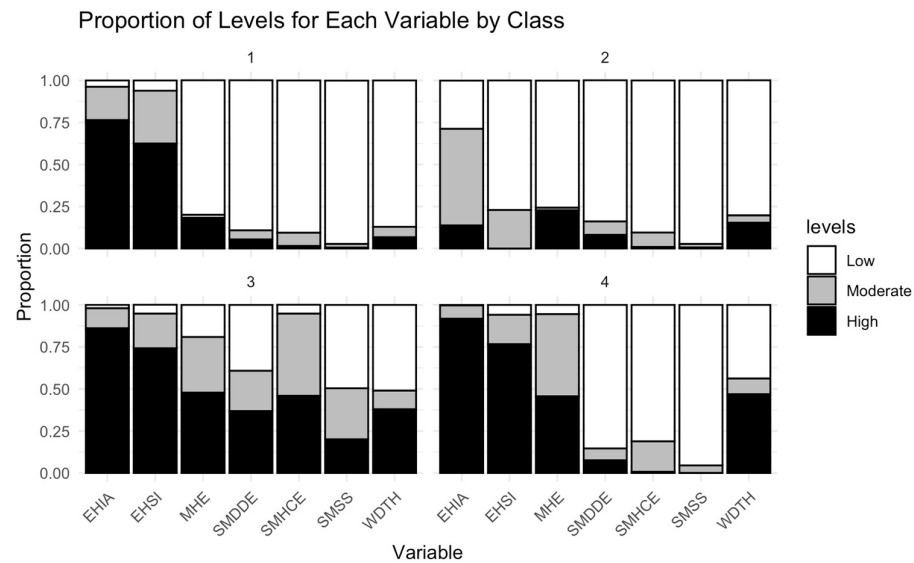


Figure 2. Distribution of eHealth engagement levels across the four identified latent classes. Numbers (1–4) denote latent eHealth engagement classes: (1) Highly Digital Engagers, (2) Moderate Digital Users, (3) Social Media and App Enthusiasts, and (4) Wearable and Health App Enthusiasts.

To understand how different patterns of digital health engagement relate to telehealth utilization, multiple logistic regression analyses (Table 4) examined the association between telehealth use, eHealth engagement (EHE) classes, and covariates. Compared to “Highly Digital Engagers”, (Class 1), “Moderate Digital Users”, (Class 2) were significantly less likely to use telehealth (OR = 0.53, 95% CI = [0.31, 0.89]), while “Social Media and App Enthusiasts”, (Class 3) and “Wearable and Health App Enthusiasts”, (Class 4) were more likely (OR = 2.32, 95% CI = [1.31, 4.12]; OR = 1.36, 95% CI = [0.95, 1.94]). Telehealth users were more likely to be female, older, have depression symptoms, and reside outside the Midwest. No significant interactions were found between EHE classes and covariates. Sub-group analyses (Supplementary Figure S3) showed depression was a significant predictor of telehealth use in all classes except Class 2. Consistent with the model with all participants, all eHealth engagement subgroups showed that depression is associated with significantly higher telehealth use. Age groups were significantly associated with increased telehealth use only in moderate digital users (class 2) and wearable and health app enthusiasts (class 4). Among the census regions, only south among class1 and Midwest among class 4 were significantly associated with less telehealth use.

An Extreme Gradient Boosting (XGBoost) model was used to evaluate how eHealth engagement (EHE) class membership and contextual features contribute to predicted telehealth use. Figure 3A displays SHAP value distributions (violin plots) for class membership. Class 2 showed consistently negative SHAP values, indicating a subtractive contribution to telehealth prediction. In contrast, Classes 3 and 4 yielded positive contributions, enhancing predicted utilization relative to the baseline, Class 1.

Figure 3C examines geographic region, specifically the South census region. For Class 4 members, Southern residence contributed positively to telehealth predictions, suggesting that regional constraints may be mitigated by high digital engagement. However, in other classes, Southern residence contributed negatively, pointing to geographic disparities conditional on behavioral readiness.

Table 4. Logistic regression modeling telehealth use by eHealth engagement and other independent variables.

Variable	OR [LCL, UCL]	<i>p</i>
Race		
NHW	Reference	
AO	0.87 [0.54, 1.41]	0.5501
HP	1.13 [0.74, 1.71]	0.5602
NHB	0.81 [0.52, 1.26]	0.3329
Education		
High school or less	Reference	
Some College	1 [0.65, 1.55]	0.9975
College Graduate or More	1.12 [0.73, 1.72]	0.5805
Household Income		
<\$20,000	Reference	
\$20,000 to <\$35,000	0.82 [0.45, 1.47]	0.4770
\$35,000 to <\$75,000	0.71 [0.4, 1.27]	0.2349
\$75,000 or More	0.83 [0.44, 1.57]	0.5540
Urbanicity		
Metro: >1M	Reference	
Metro: 20K to 1M	0.86 [0.62, 1.19]	0.3512
Rural	0.86 [0.56, 1.31]	0.4598
Birth gender		
Male	Reference	
Female	1.22 [0.91, 1.64]	0.1711
Age group		
18–34	Reference	
35–49	2.09 [1.4, 3.13]	0.0011 **
50–64	1.69 [1.21, 2.35]	0.0036 **
65–74	1.77 [1.11, 2.82]	0.0184 *
75+	2.06 [1.15, 3.68]	0.0178 *
eHealth engagement		
Class 1	Reference	
Class 2	0.53 [0.31, 0.89]	0.0196 *
Class 3	2.32 [1.31, 4.12]	0.0062 **
Class 4	1.36 [1.05, 1.94]	0.0458 *
Census region		
Midwest	Reference	
Northeast	2.06 [1.32, 3.21] **	0.0031 **
South	1.65 [1.19, 2.28] **	0.0046 **
West	2.38 [1.59, 3.56] ***	<0.001 ***
Depression		
No	Reference	
Yes	2.44 [1.76, 3.4] ***	<0.001 ***
Lack Transportation		
No	Reference	
Yes	1.22 [0.83, 1.78]	0.3001

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In sum, SHAP analyses reveal that key predictors such as class membership, depression, and geography are context-sensitive; their influence depends on the underlying engagement profile. These results demonstrate how behavioral segmentation sharpens the interpretability of predictive models in digital health, uncovering non-obvious subgroup disparities in telehealth access.

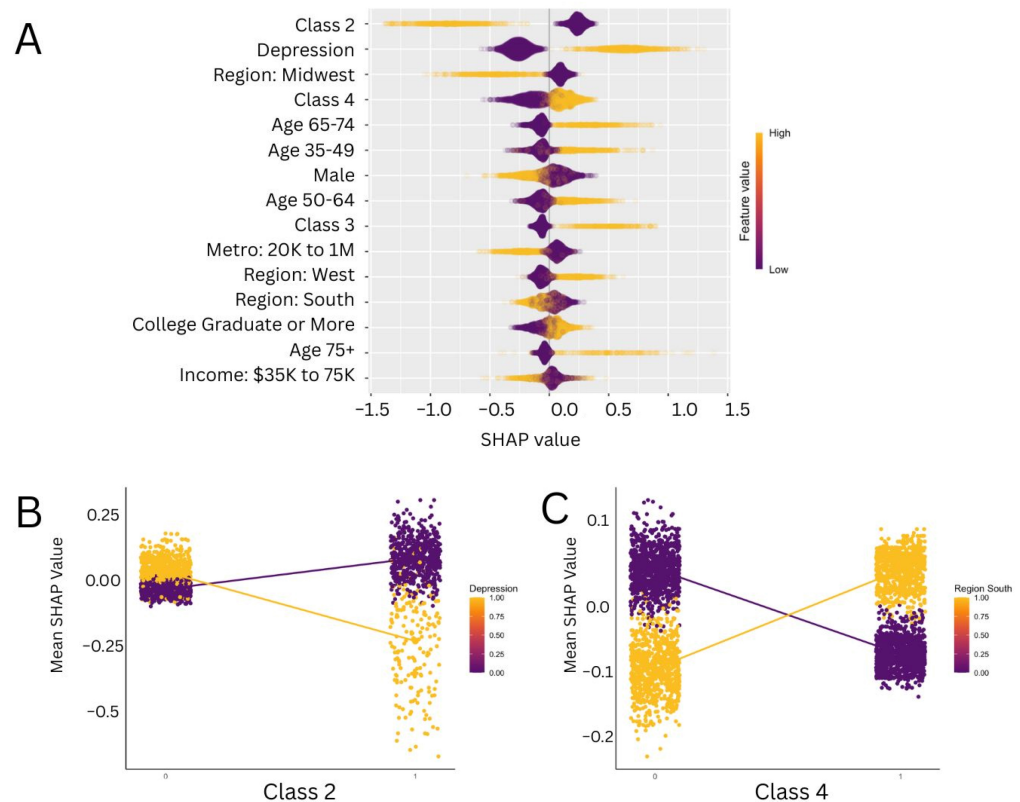


Figure 3. SHAP analysis of telehealth utilization predictors. (A) SHAP summary plot ranking the top features by importance; (B) Interaction effect between Class 2 (Moderate Digital Users) and depression status; (C) Interaction effect between Class 4 (Wearable & Health App Enthusiasts) and residence in the South census region.

4. Discussion

This study provides novel insights into how patterns of eHealth engagement influence telehealth utilization, offering practical implications for healthcare systems seeking to enhance both efficiency and equity in digital service delivery. By combining Latent Class Analysis (LCA) with interpretable predictive modeling, we demonstrate how behavioral segmentation, a well-established strategy in consumer analytics, can be adapted to healthcare contexts to guide population-level service planning, outreach strategies, and digital inclusion efforts [24,41].

A key finding with substantial policy relevance is the significantly reduced likelihood of telehealth use among individuals categorized as “Moderate Digital Users”, (Class 2). Although this group engages with core eHealth functions such as accessing health information or patient portals, they show limited participation in interactive, social, or preparatory digital activities. Notably, this aligns with prior studies documenting digital access gaps among populations with lower health literacy or confidence navigating eHealth tools [31,35]. However, this pattern does not mean a deficit in digital capability or unwillingness to engage in care. It could reflect a more transactional mode of digital engagement, in which technologies are used episodically to meet discrete needs.

Conversely, individuals in “Social Media and App Enthusiasts”, (Class 3) and “Wearable and Health App Enthusiasts”, (Class 4) were more likely to use telehealth services, consistent with prior research linking digital engagement with increased telehealth adoption [37,42]. These users may possess greater digital literacy and comfort with online health interactions, which fosters smoother transition into virtual care settings. Additionally, depression status, census region, and age emerged as significant contextual features, with

depression contributing positively to telehealth use across most engagement profiles. Older age was strongly associated with increased telehealth utilization, with odds ratios rising from 2.07 (ages 35–49) to 3.01 (ages 75+) compared to the 18–34 reference group ($p < 0.001$), likely reflecting greater healthcare needs among older adults. This finding echoes earlier work showing increased reliance on digital care pathways among individuals managing mental health needs [34]. Beyond statistical distinctiveness, the identified classes represent divergent behavioral phenotypes. Class 3 (Social Media & App Enthusiasts) appears driven by communicative and information-seeking motives, relying on peer networks and external content. In contrast, Class 4 (Wearable & App Enthusiasts) exhibits a ‘Quantified Self’ profile, driven by self-regulation and automated data tracking. Both pathways—social engagement and personal monitoring—converge to support high telehealth readiness, but they suggest different entry points for intervention.

Importantly, our results suggest that behavioral profiles offer predictive value above and beyond traditional sociodemographic variables. While past research has largely emphasized age, socioeconomic status, and access as telehealth determinants [33,36], our class-based modeling approach captures engagement complexity that is often missed in individual-level predictors.

The four engagement classes identified in this study are consistent with digital health engagement frameworks that conceptualize engagement as a dynamic, multi-dimensional process encompassing access, interaction, participation, and self-management, rather than a binary state of use or non-use. Based on this perspective, digital health engagement varies in intensity and continuity over time and is shaped by individuals’ goals, preferences, capabilities, and health contexts. Moderate Digital Users (Class 2) exhibit a selectively engaged profile characterized by basic access to digital health information and systems but limited participation in more interactive, social, or self-tracking activities. This pattern reflects a lower-intensity, task-oriented mode of engagement in which digital tools are used to meet discrete healthcare needs without sustained or integrative interaction. In contrast, Social Media and App Enthusiasts (Class 3) and Wearable & Health App Enthusiasts (Class 4) represent more active engagement pathways, marked by continuous interaction, participatory behaviors, and ongoing self-monitoring within broader digital health ecosystems [43,44]. Prior work emphasizes that lower-intensity and time-to-time engagement may still be appropriate and meaningful depending on users’ needs and care contexts [44]. Accordingly, the engagement classes help explain why particular combinations of digital behaviors co-occur and why their association with telehealth utilization differs across subgroups.

Methodologically, this study contributes by integrating XGBoost with SHAP explanations to balance predictive power and interpretability. Unlike traditional black-box models, SHAP provides clear, case-level insights into how each feature influences a prediction, an essential capability in healthcare, where decisions must be transparent, clinically sound, and ethically justified [26,28]. By offering interpretable outputs that align with stakeholder needs, SHAP supports trustworthy and actionable analytics in digital health planning. While SHAP analysis identified key predictors consistent with the logistic regression, it also revealed complex non-linear associations and threshold effects that the linear model did not capture [45].

The study demonstrates how population-level digital health engagement profiles can be translated into operationally meaningful insights for healthcare systems. Specifically, the findings show how routinely collected digital engagement data can be used to distinguish qualitatively different patterns of interaction with digital health tools, thereby informing where and how supportive resources might be deployed within existing care infrastructures. One illustrative example is the group characterized as Moderate Digital Users (Class 2) those who exhibit consistent use of core eHealth functions but limited engagement with

more interactive or preparatory digital tools. In applied settings, individuals with similar engagement patterns could be identified using existing portal usage logs or intake questionnaires, without the need for additional clinical screening instruments. Rather than physician-led intervention, support for this group could be provided by non-clinical roles, such as digital navigators or care coordinators, with a focus on onboarding, visit preparation, and clarifying when telehealth may be appropriate or beneficial within a patient's broader care plan. Potential support for this group may include simplified appointment workflows, telehealth visit preparation checklists, or brief asynchronous guidance embedded within patient portals. These examples are intentionally illustrative rather than prescriptive and are conceptually low-cost and scalable. However, their feasibility depends on local infrastructure, staffing models, and organizational readiness. Importantly, our study recognizes that engagement patterns may reflect patient personal preferences, clinical appropriateness, or contextual constraints, instead of a deficit. Accordingly, the results show how engagement profiles can inform targeted outreach and population-level resource planning, without implying individual-level prescriptions and clinical effectiveness.

Limitations and Future Directions

Several limitations should be acknowledged when interpreting the findings. First, the cross-sectional nature of the HINTS dataset precludes causal inference. While associations between digital engagement patterns and telehealth use were observed, temporal directionality cannot be established. Furthermore, these relationships may be bidirectional or confounded by unmeasured variables such as overall healthcare utilization intensity or disease severity. High utilizers of healthcare may naturally engage more across all channels, potentially inflating the observed associations. Future research should consider longitudinal designs to disentangle these effects and better understand how digital behavior influences sustained telehealth adoption over time.

Second, the reliance on self-reported data introduces the potential for recall bias and social desirability effects, particularly in measures of digital health engagement and mental health status. While the survey is nationally representative and rigorously validated, self-reporting may still underestimate or overestimate actual behaviors. Third, the 12-month telehealth utilization window may underestimate digital readiness in younger or healthier populations who, despite high eHealth literacy, lacked clinical necessity for care during the study period. Thus, 'non-user' status may reflect a lack of recent need rather than a lack of digital capability. Future studies should distinguish between the capacity to use telehealth and actual utilization to better assess digital readiness across the lifespan.

Finally, the engagement typologies were constructed using categorized composite indicators to support interpretable, population-level segmentation. While this approach facilitates practical sensemaking and aligns with health system planning objectives, it may obscure within-class heterogeneity and overlook more nuanced behavioral variation. Accordingly, the identified engagement profiles should be viewed as a starting point for understanding broad patterns of digital health behavior rather than as definitive or exhaustive classifications.

Future research could build on this work by applying finer-grained measurement approaches, such as item-level latent variable modeling or item response theory, to examine within-class differences or to refine engagement typologies for specific clinical populations or care contexts. Future research would prioritize longitudinal or experimental designs to disentangle the temporal order of these behaviors and determine whether increasing eHealth engagement causally leads to greater telehealth adoption.

5. Conclusions

This study contributes to the literature on data mining in healthcare by applying behavioral segmentation and interpretable machine learning to examine population-level patterns of telehealth utilization. Using Latent Class Analysis, we identified distinct digital engagement profiles that capture real-world heterogeneity in how patients interact with health technologies, insights that are not visible through traditional demographic models. By pairing these profiles with SHAP-enhanced XGBoost modeling, we introduce a transparent, stakeholder-aligned approach to predictive analytics in digital health. SHAP not only improves model interpretability but also reveals how behavioral traits interact with clinical and geographic factors to shape care access.

Our most important finding highlights the lower prevalence of telehealth use among individuals classified as Moderate Digital Users, who engage with core eHealth functions but less frequently with interactive or preparatory tools. Importantly, this pattern should not be interpreted as evidence of digital deficit or unmet need, but rather as one mode of engagement that may reflect preferences, care contexts, or clinical appropriateness. The value of identifying such profiles lies in supporting population-level understanding and planning. These findings highlight how differences in digital engagement patterns may be considered in telehealth planning, without implying that increased utilization is universally desirable or appropriate. As healthcare systems adopt AI and big data tools, our study emphasizes the value of combining behavioral pattern discovery with transparent, explainable models to support a more equitable and data-driven digital health future.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/info17020215/s1>, Table S1: Criteria for Categorizing Composite Digital Health Engagement Variables into Levels; Table S2: Average Posterior Probabilities (AvePP) for the 4-Class Model; Figure S1: Flowchart of Sample Selection and Exclusion Criteria; Figure S2: Methodological Workflow of the Study; Figure S3: Multiple Logistic Regression Results in Subgroups Defined by Latent Class Membership.

Author Contributions: Conceptualization, N.Y.; methodology, N.Y.; software, X.Y.; validation, N.Y.; formal analysis, X.Y.; investigation, X.Y.; resources, X.Y.; data curation, N.Y.; writing—original draft preparation, N.Y.; writing—review and editing, X.Y.; project administration, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The study used data derived from the 2022 U.S. Health Information National Trends Survey (HINTS). The dataset used in this study can be downloaded from the project GitHub repository: <https://github.com/tybuz2021/Uncovering-eHealth-Engagement-Patterns-Through-Latent-Class-Analysis-and-SHAP> (accessed on 26 January 2026).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bashshur, R.L.; Howell, J.D.; Krupinski, E.A.; Harms, K.M.; Bashshur, N.; Doarn, C.R. The Empirical Foundations of Telemedicine Interventions in Primary Care. *Telemed. E-Health* **2016**, *22*, 342–375. [CrossRef] [PubMed]
2. Bashshur, R.L.; Shannon, G.W.; Smith, B.R.; Alverson, D.C.; Antoniotto, N.; Barsan, W.G.; Bashshur, N.; Brown, E.M.; Coye, M.J.; Doarn, C.R.; et al. The Empirical Foundations of Telemedicine Interventions for Chronic Disease Management. *Telemed. E-Health* **2014**, *20*, 769–800. [CrossRef]
3. Hailey, D.; Roine, R.; Ohinmaa, A. Systematic Review of Evidence for the Benefits of Telemedicine. *J. Telemed. Telecare* **2002**, *8*, 1–7. [CrossRef] [PubMed]
4. Hatef, E.; Wilson, R.F.; Hannum, S.M.; Zhang, A.; Kharrazi, H.; Weiner, J.P.; Davis, S.A.; Robinson, K.A. *Use of Telehealth During the COVID-19 Era*; Agency for Healthcare Research and Quality: Rockville, MD, USA, 2023.

5. Peters, G.M.; Kooij, L.; Lenferink, A.; Van Harten, W.H.; Doggen, C.J. The Effect of Telehealth on Hospital Services Use: Systematic Review and Meta-Analysis. *J. Med. Internet Res.* **2021**, *23*, e25195. [\[CrossRef\]](#)
6. Snoswell, C.L.; Chelberg, G.; De Guzman, K.R.; Haydon, H.M.; Thomas, E.E.; Caffery, L.J.; Smith, A.C. The Clinical Effectiveness of Telehealth: A Systematic Review of Meta-Analyses from 2010 to 2019. *J. Telemed. Telecare* **2023**, *29*, 669–684. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Wade, V.A.; Elliott, J.A.; Hiller, J.E. Clinician Acceptance Is the Key Factor for Sustainable Telehealth Services. *Qual. Health Res.* **2014**, *24*, 682–694. [\[CrossRef\]](#)
8. Choi, W.S.; Park, J.; Choi, J.Y.B.; Yang, J.-S. Stakeholders' Resistance to Telemedicine with Focus on Physicians: Utilizing the Delphi Technique. *J. Telemed. Telecare* **2019**, *25*, 378–385. [\[CrossRef\]](#)
9. Sagaro, G.G.; Battineni, G.; Amenta, F. Barriers to Sustainable Telemedicine Implementation in Ethiopia: A Systematic Review. *Telemed. Rep.* **2020**, *1*, 8–15. [\[CrossRef\]](#)
10. Hall, J. 67 Million Medicare Recipients Face 'Chaos' If Congress Cuts Telehealth Benefits. Available online: <https://www.marketwatch.com/story/67-million-medicare-recipients-are-facing-chaos-if-congress-cuts-telehealth-benefits-b7a5c00e> (accessed on 26 January 2025).
11. Barelo, S.; Triberti, S.; Graffigna, G.; Libreri, C.; Serino, S.; Hibbard, J.; Riva, G. eHealth for Patient Engagement: A Systematic Review. *Front. Psychol.* **2016**, *6*, 2013. [\[CrossRef\]](#)
12. Norman Burrell, D. Telehealth Technologies, Disruptions, and Health Literacy. *Health Econ. Manag. Rev.* **2024**, *5*, 33–46. [\[CrossRef\]](#)
13. Petretto, D.R.; Carrogu, G.P.; Gaviano, L.; Berti, R.; Pinna, M.; Petretto, A.D.; Pili, R. Telemedicine, e-Health, and Digital Health Equity: A Scoping Review. *Clin. Pract. Epidemiol. Ment. Health CPEMH* **2024**, *20*, e17450179279732. [\[CrossRef\]](#)
14. World Health Organization. *Global Strategy on Digital Health 2020–2027*; World Health Organization: Geneva, Switzerland, 2025.
15. Doumen, M.; De Cock, D.; Van Lierde, C.; Betraains, A.; Pazmino, S.; Bertrand, D.; Westhovens, R.; Verschueren, P. Engagement and Attrition with eHealth Tools for Remote Monitoring in Chronic Arthritis: A Systematic Review and Meta-Analysis. *RMD Open* **2022**, *8*, e002625. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Stefanicka-Wojtas, D.; Kurpas, D. eHealth and mHealth in Chronic Diseases—Identification of Barriers, Existing Solutions, and Promoters Based on a Survey of EU Stakeholders Involved in Regions4PerMed (H2020). *J. Pers. Med.* **2022**, *12*, 467. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Renzi, E.; Baccolini, V.; Migliara, G.; De Vito, C.; Gasperini, G.; Cianciulli, A.; Marzuillo, C.; Villari, P.; Massimi, A. The Impact of eHealth Interventions on the Improvement of Self-Care in Chronic Patients: An Overview of Systematic Reviews. *Life* **2022**, *12*, 1253. [\[CrossRef\]](#)
18. Ch, A.S.; Pervaiz, F.; Nawaz, H.; Rafiq, S.; Bibi, Y.; Almagharbeh, W.T.; Mobeen, M. Health equity in telemedicine: Addressing disparities in digital healthcare access—a narrative review-s. *Insights-J. Life Soc. Sci.* **2025**, *3*, 56–62. [\[CrossRef\]](#)
19. Rush, K.L.; Singh, S.; Seaton, C.L.; Burton, L.; Li, E.; Jones, C.; Davis, J.C.; Hasan, K.; Kern, B.; Janke, R. Telehealth Use for Enhancing the Health of Rural Older Adults: A Systematic Mixed Studies Review. *Gerontol.* **2022**, *62*, e564–e577. [\[CrossRef\]](#)
20. Phuong, J.; Ordóñez, P.; Cao, J.; Moukheiber, M.; Moukheiber, L.; Caspi, A.; Swenor, B.K.; Naawu, D.K.N.; Mankoff, J. Telehealth and Digital Health Innovations: A Mixed Landscape of Access. *PLoS Digit. Health* **2023**, *2*, e0000401. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Arias López, M.D.P.; Ong, B.A.; Borrat Frigola, X.; Fernández, A.L.; Hicklent, R.S.; Obeles, A.J.; Rocimo, A.M.; Celi, L.A. Digital Literacy as a New Determinant of Health: A Scoping Review. *PLoS Digit. Health* **2023**, *2*, e0000279. [\[CrossRef\]](#)
22. Bitar, H.; Alismail, S. The Role of eHealth, Telehealth, and Telemedicine for Chronic Disease Patients during COVID-19 Pandemic: A Rapid Systematic Review. *Digit. Health* **2021**, *7*, 20552076211009396. [\[CrossRef\]](#)
23. Collins, L.M.; Lanza, S.T. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
24. Chen, R.; Sun, J.; Dittus, R.S.; Fabbri, D.; Kirby, J.; Laffer, C.L.; McNaughton, C.D.; Malin, B. Patient Stratification Using Electronic Health Records from a Chronic Disease Management Program. *IEEE J. Biomed. Health Inform.* **2016**. [\[CrossRef\]](#)
25. Deo, R.C. Machine Learning in Medicine. *Circulation* **2015**, *132*, 1920–1930. [\[CrossRef\]](#)
26. Dey, S.; Chakraborty, P.; Kwon, B.C.; Dhurandhar, A.; Ghalwash, M.; Saiz, F.J.S.; Ng, K.; Sow, D.; Varshney, K.R.; Meyer, P. Human-Centered Explainability for Life Sciences, Healthcare, and Medical Informatics. *Patterns* **2022**, *3*, 100493. [\[CrossRef\]](#)
27. Saraswat, D.; Bhattacharya, P.; Verma, A.; Prasad, V.K.; Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. Explainable AI for Healthcare 5.0: Opportunities and Challenges. *IEEE Access* **2022**, *10*, 84486–84517. [\[CrossRef\]](#)
28. Yang, C.C. Explainable Artificial Intelligence for Predictive Modeling in Healthcare. *J. Health Inf. Res.* **2022**, *6*, 228–239. [\[CrossRef\]](#)
29. Zhang, Y.; Wang, J.; Zong, H.; Singla, R.K.; Ullah, A.; Liu, X.; Wu, R.; Ren, S.; Shen, B. The Comprehensive Clinical Benefits of Digital Phenotyping: From Broad Adoption to Full Impact. *npj Digit. Med.* **2025**, *8*, 196. [\[CrossRef\]](#)
30. Jung, H.W.; Kim, D.Y.; Lee, I.; Kim, O.; Lee, S.; Lee, S.; Chung, U.S.; Kim, J.-H.; Kim, S.; Kim, J.W. Key Features of Digital Phenotyping for Monitoring Mental Disorders: Systematic Review. *J. Med. Internet Res.* **2025**, *27*, e77331. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Massey, P.M. Where Do U.S. Adults Who Do Not Use the Internet Get Health Information? Examining Digital Health Information Disparities From 2008 to 2013. *J. Health Commun.* **2016**, *21*, 118–124. [\[CrossRef\]](#) [\[PubMed\]](#)

32. Schmeida, M.; McNeal, R.S. The Telehealth Divide: Disparities in Searching Public Health Information Online. *J. Health Care Poor Underserved* **2007**, *18*, 637–647. [[CrossRef](#)] [[PubMed](#)]
33. Shih, H.-J.; Min, H.; Chang, J. Assessing Health Disparities in Digital Services and Technologies During and After the COVID-19 Pandemic: A Pooled Cross-Sectional Analysis Using HINTS Data. *Patient Prefer. Adherence* **2025**, *19*, 87–96. [[CrossRef](#)]
34. Hartoonian, N.; Ormseth, S.R.; Hanson, E.R.; Bantum, E.O.; Owen, J.E. Information-Seeking in Cancer Survivors: Application of the Comprehensive Model of Information Seeking to HINTS 2007 Data. *J. Health Commun.* **2014**, *19*, 1308–1325. [[CrossRef](#)]
35. Jacobs, W.; Amuta, A.O.; Jeon, K.C. Health Information Seeking in the Digital Age: An Analysis of Health Information Seeking Behavior among US Adults. *Cogent Soc. Sci.* **2017**, *3*, 1302785. [[CrossRef](#)]
36. Rutten, L.J.F.; Squiers, L.; Hesse, B. Cancer-Related Information Seeking: Hints from the 2003 Health Information National Trends Survey (HINTS). *J. Health Commun.* **2006**, *11*, 147–156. [[CrossRef](#)]
37. Yang, X.; Yang, N.; Lewis, D.; Parton, J.; Hudnall, M. Patterns and Influencing Factors of eHealth Tools Adoption among Medicaid and Non-Medicaid Populations from the Health Information National Trends Survey (HINTS) 2017–2019: Questionnaire Study. *J. Med. Internet Res.* **2021**, *23*, e25809. [[CrossRef](#)]
38. Linzer, D.A.; Lewis, J.B. polLCA: An R Package for Polytomous Variable Latent Class Analysis. *J. Stat. Softw.* **2011**, *42*, 1–29. [[CrossRef](#)]
39. Meng, Y.; Yang, N.; Qian, Z.; Zhang, G. What Makes an Online Review More Helpful: An Interpretation Framework Using XGBoost and SHAP Values. *J. Theor. Appl. Electron. Commer. Res.* **2020**, *16*, 466–490. [[CrossRef](#)]
40. Olive, D.J. Software for Data Analysis: Programming with R. *Technometrics* **2010**, *52*, 261.
41. Li, X.; Wong, K.-C. Multiobjective Patient Stratification Using Evolutionary Multiobjective Optimization. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1619–1629.
42. Kontos, E.; Blake, K.D.; Chou, W.-Y.S.; Prestin, A. Predictors of eHealth Usage: Insights on the Digital Divide from the Health Information National Trends Survey 2012. *J. Med. Internet Res.* **2014**, *16*, e172. [[PubMed](#)]
43. Ratcliff, C.L.; Krakow, M.; Greenberg-Worisek, A.; Hesse, B.W. Digital Health Engagement in the US Population: Insights From the 2018 Health Information National Trends Survey. *Am. J. Public Health* **2021**, *111*, 1348–1351. [[CrossRef](#)]
44. Torous, J.; Michalak, E.E.; O'Brien, H.L. Digital Health and Engagement—Looking behind the Measures and Methods. *JAMA Netw. Open* **2020**, *3*, e2010918. [[CrossRef](#)]
45. Valdes, G.; Luna, J.M.; Eaton, E.; Simone, C.B.; Ungar, L.H.; Solberg, T.D. MediBoost: A Patient Stratification Tool for Interpretable Decision Making in the Era of Precision Medicine. *Sci. Rep.* **2016**, *6*, 37854. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.