ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Health Policy Analysis

A Value Framework to Assess Patient-Facing Digital Health Technologies That Aim to Improve Chronic Disease Management: A Delphi Approach

Madeleine Haig, MSc, Caitlin Main, MSc, Danitza Chávez, MSc, Panos Kanavos, PhD

ABSTRACT

Objectives: Digital health technologies (DHTs) can optimise healthcare costs and improve quality and efficiency of care. However, the fast-paced rate of innovation and varying evidence standards can make it difficult for decision-makers to assess these technologies in an efficient and evidence-based manner. We sought to develop a comprehensive framework to assess the value of novel patient-facing DHTs used to manage chronic diseases by eliciting stakeholder value preferences.

Methods: Literature review and primary data collection from a three-round web-Delphi exercise was utilized. 79 participants from 5 stakeholder groups (patients, physicians, industry, decision makers, and influencers) and 3 countries (United States of America, United Kingdom, and Germany) took part. Likert scale data were statistically analyzed to determine intergroup differences in both country and stakeholder groups, stability of results, and overall consensus.

Results: The resulting co-created framework comprised 33 stable indicators with consensus from quantitative value judgments across domains: health inequalities, data rights and governance, technical and security, economic characteristics, clinical characteristics, and user preferences. Lack of stakeholder consensus was observed on the importance of value-based care models, optimizing resources for sustainable systems, and stakeholder involvement in DHT design, development, and implementation; however, this was because of high rates of neutrality and not negative judgments. Supply-side actors and academic experts were the most unstable stakeholder groups.

Conclusion: Stakeholder value judgments revealed a need for a coordinated regulatory and health technology assessment policy response that updates laws to meet technological innovations, offers a pragmatic approach to evidence standards to assess DHTs, and involves stakeholders to understand and meet their needs.

Keywords: Delphi, digital health technology, health technology assessment, patient-facing technology solution, value framework.

VALUE HEALTH. 2023; 26(10):1474–1484

Introduction

Health systems in the post–COVID-19 world have entered an age of increased reliance on digital technology whereby patient interactions with the health system are increasingly through digital health technologies (DHTs), and their information is digitally stored, processed, and transmitted.[1] As populations continue to age and chronic diseases continue to be the leading cause of death, DHTs have emerged as a potential solution, particularly regarding improved diagnostics, remote monitoring, and disease self-management.[2] DHTs have the capacity to alleviate strains caused by rising chronic disease prevalence and the associated rise in costs.[3,4] In doing so, they can contribute to cost optimization, equity, efficiency and quality of care improvements, population health management, and improved clinical decision making.[5] DHTs also have significant variation in functionality, risk profile, and value proposition, ranging from patient-facing technologies that monitor and influence individuals' behavior to system-facing technologies that directly affect health system-level operations, data sharing, and analyses. Big data capabilities and artificial intelligence amplify several issues, including data bias, privacy, security, and governance.[5] The digital health ecosystem encompassed over 350 000 (regulated and unregulated) health-related apps available in 2020. Many of these applications, along with other digital solutions, are seeking health system integration and coverage by healthcare budgets. Such significant market growth raises questions about how to evaluate these technologies and whether existing methodologies are sufficient.

DHTs struggle to meet the same evidence standards as drugs, often because of rapid technical innovations and lack of adequate comparators.[6-8] Randomized Controlled Trials (RCTs) are considered the gold-standard for proving effectiveness; however, they present several challenges to DHTs including long timelines, measuring personalized care delivery, and developing adequate placebos. Traditional health technology assessments (HTAs) applied to pharmaceuticals and medical devices do not address

this, which can result in misjudgment of a DHT's benefits. A lack of clear DHT-specific evidence assessment guidance makes it difficult for innovators to understand evidentiary requirements and may hinder the pace of digital innovation.[9] These DHT-specific issues highlight a need for evolution in HTA methodologies. To this end, many HTA agencies have developed solutions including active stakeholders engagement to assess DHTs.[10]

Health systems and HTA bodies use a number of value frameworks to evaluate innovations and allocate resources efficiently.[11,12] Value frameworks reflect the differing preferences of key stakeholders involved in their construction and aim to communicate essential value dimensions transparently and explicitly, ultimately supporting informed decision making.[13,14] Germany is the first country in Europe with a DHT-specific pricing and reimbursement pathway for prescribed patient-facing DHTs, the DiGA directory.[15] The UK's National Institute for Health and Care Excellence has created an evidence standards framework for DHTs, to assist local decision makers in deciding which technologies to reimburse in their Integrated Care Systems.[16] Comparatively, DHTs are met with considerable skepticism in many other countries, with no formal reimbursement pathways. In Italy, for example, there are no regulations for DHT reimbursement beyond the 2017 EU medical device legislation, and there are no DHTs marketed, used, or reimbursed to date.[17]

In this study, we develop a holistic value framework for assessment of regulated patient-facing DHTs for chronic disease management by eliciting value concerns and preferences of different stakeholders. Regulated DHTs are those subjected to oversight by governmental agencies and must meet certain safety and effectiveness standards before gaining marketing authorization. Technologies considered to be medical devices, which are designed to diagnose or treat disease, and may pose a risk to patients, are likely to be subject to regulation, whereas other digital technologies, such as telehealth platforms used exclusively for video consultations, are unlikely to be subject to regulation. Patient-facing DHTs include solutions classified as "software as a medical device (SaMD)[18]" and may be used for active self-monitoring by patients as well as remote monitoring by healthcare professionals (HCPs), thus offering active data monitoring and transmission.

Our study adds to the literature in 3 ways: first, we develop a holistic framework for patient-facing DHTs, rather than a framework tailored to the assessment needs of one type of organization or a single DHT, that can be used by health systems to make value judgments. Second, by using the Delphi approach, we elicit the preferences of a large number of stakeholders across settings, which allows for widespread representation. Third, we highlight the urgent need for digital HTA pathways that promote appropriate assessment methods, utilize opinions across stakeholder groups, and encourage digital innovation within the health sector. The next section outlines the methods, and this is followed by the results. A discussion of the key issues takes place before concluding.

Methods

Our study utilizes a literature review and primary data collection through a web-based Delphi exercise with statistical and thematic analysis to identify stakeholder value preferences in patient-facing DHTs. The web-Delphi study received ethics approval from the London School of Economics (reference: 4363). All participants completed an informed consent form confirming voluntary participation, agreeing to anonymous data usage and acknowledging private sponsorship.

Decision Context

The study countries are the United Kingdom, Germany, and the United States of America, all selected because of significant developments in regulatory frameworks and DHT assessment methods. These 3 countries also represent different archetypes of health system financing: taxation, social insurance, and a system with many private payers and a significant public sector, respectively.

The scope for primary and secondary research includes regulated patient-facing DHTs used in chronic disease management (see Appendix Table 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008). These consisted of healthcare apps and wearables considered to be a medical device, including implanted devices, that is, patient-facing technologies of significant risk level to be regulated as a medical device.

Secondary Research

A scoping review was conducted to gather evidence on existing value frameworks and thus assist with the identification of the value criteria for the starting framework. Gray and peer-reviewed literature was included if it was based on the United Kingdom, the United States of America, or Germany; additional literature was included if it had a global or pan-European scope, and thus covered the study countries. (see Appendix section 2.2, Appendix Figure 1, and Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008)

Key value domains of DHTs identified in the literature included economic, clinical and technical characteristics, user preferences, technology safety, and regulatory compliance[19–23] (see Appendix Table 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008). An initial framework was proposed based on these value domains and indicators found in the literature, comprising 34 indicators across 6 domains: (a) health inequalities (3), (b) data rights and governance (7), (c) technical and security (11), (d) economic characteristics (6), (e) clinical characteristics (9), and (f) user preferences (9) (see Appendix Tables 4 and 5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008).

The starting framework consisted of indicators and their descriptions, to ensure all participants had a base-level understanding of the concepts. An "indicator" refers to a component considered during technology assessment, whereas a "value domain" is a category within which indicators lie. The value domains, indicators, and their descriptions were discussed and altered according to feedback from experts in the HTA field, industry representatives, and academia.

Primary Research

The Delphi approach

A web-Delphi approach elicited preferences and value concerns of stakeholders. Delphi studies have been widely used to measure consensus since the emergence of the technique in the 1960s[24–28] and have been used in value framework creation in the health sector.[29] We used the online platform Welphi[30] to communicate with participants, conduct the Delphi exercise, share preliminary results in Round (R)3, and extract data after each round. Welphi[30] facilitates an online Delphi panel in which round results are collected asynchronously, allowing participants to complete the exercise at their own pace. All opinions are shared anonymously, with feedback incorporated iteratively through 3 rounds.

A total of 3 rounds of Delphi were conducted: in R1, participants commented on the indicators proposed in the starting framework and were able to propose their own; in R2, participants

rated each indicator on an "importance" Likert scale (from "not at all important" to "very important"); and in R3, participants were shown the distribution of responses across all participants for each indicator, as well as their own response, and had the opportunity to change their response.

Recruitment and eligibility

The stakeholder groups consisted of patients, HCPs, supply-side actors, decision makers, and policy influencers, (see Appendix Table 6 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008). Participants must work or have worked in 1 of the 3 study countries, aside from "Influencers," whose impact tends to be more far-reaching. The team had a target sample size of 12 to 15 individuals per group, with an overall 3-round completion goal of 60 participants.

Assumptions around individuals' willingness to participate were made at the beginning of the project that did not hold throughout the recruitment period. Supply-side actors and influencers had a high willingness to participate, whereas other stakeholder groups were less willing or were prevented from participating based on legal restrictions. Professional recruiters were hired to fill recruitment gaps and ensure equal representation across stakeholder groups including: 5 patients, 1 doctor, and 4 decision makers from Germany, as well as 6 decision makers and 6 patients from the United States. In the trade-off between potential bias considerations around professional recruiters versus limited representation from German and American patients and decision makers, researchers prioritized stakeholder representation. There is ongoing debate in the research community regarding paying for recruitment, which is further acknowledged in the limitations section, as well as Appendix section 2.2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008.

Data Analysis

The primary study outcome is stable consensus (or dissensus) to establish a set of indicators deemed important for assessing the value of patient-facing DHTs. The secondary outcome is value concerns of stakeholder groups.

Thematic and statistical analysis

Thematic analysis was conducted using NVivo[31] to identify key concepts and themes proposed by participants in R1 and include them in the R2 and R3 framework. Anonymized, participant-specific Likert scale quantitative data from R2 and R3 were analyzed using StataSE 16.1 (StataCorp LLC, College Station, TX).[32]

Several statistical tests were completed to determine stakeholder and country group differences in both R2 and R3 (Table 1[17-21,27-31]). Inter-rater agreement within each stakeholder group was tested through the Kappa statistic and Gwet's agreement coefficient, establishing whether individuals within each stakeholder and country group had a substantial likelihood of

Table 1. Summary of definitions and methods used in this study.

Definition(s)		Method	Interpretation
Agreement	The overall agreement on importance of an indicator across all participants	Percentage agreement	Agreement is classed as participants rating an indicator as "important" or "very important"
		Central tendency and level of dispersion using median and the interquartile range (IQR)[21,29,30]	Positive impact: median of "important" or "very important" No positive impact: median of "little importance" or "not at all important" Agreement: IQR ≤ 1 (ie, >50% of all opinions fall within 1 point on the scale) Lack of agreement: IQR > 1
	The likelihood at which participants independently rate a given statement the same in each round accounting for agreement occurring simply by chance */Whether participants agree with each other on the ranking they gave for each indicator in each round*	Inter-rater agreement (IRA) using Gwet's kappa coefficient applying linear weights[31]	Poor agreement: Gwet's kappa < 0.00 Slight agreement: 0.00 > Gwet's kappa ≤ 0.20 Fair agreement: 0.20 > Gwet's kappa ≤ 0.40 Moderate agreement: 0.40 > Gwet's kappa ≤ 0.60 Substantial agreement: 0.60 > Gwet's kappa ≤ 0.80 Almost perfect agreement: 0.80 > Gwet's kappa ≤ 1
	Whether the medians of ≥2 groups are statistically different and which exact groups are different / *Whether stakeholder or country groups have statistically significant median rankings per indicator*	Kruskall-Wallis H-test adjusted for ties with 4 degrees of freedom	Significant difference $P \leq .05$
		Dunn's test using pairwise multiple comparisons with the Bonferroni adjustment	Significant difference $P \leq 0.05$
Stability	The stability of group responses per indicator between rounds/ *The likelihood participants changed their opinions as a group from R2 to R3*	Non-parametric Wilcoxon matched-pairs signed-rank test[21,29]	Unstable response: $P \leq 0.05$ statistically significant change
Consensus	Consensus was considered achieved when an indicator was approved by qualified majority, meaning the indicator had an IQR ≤ 1 in R3 and showcased stability (nonstatistically significant change) between R2 and R3.		

Source: The Authors based their chosen methodology on a search of the literature on Delphi panel methodologies.[17-21,27,28]
Note. Because "consensus" is subjective, it was used only for the purposes of the discussion section.
R2 indicates round 2; R3, indicates round 3.

making the same judgment independently. A benchmark scale was used to assess levels of agreement from Gwet's agreement coefficient with linear weights,[33] allowing the comparison of groups with internal agreement.

Stability of group responses for each indicator between R2 and R3 was examined through the Wilcoxon test, indicating whether respondents were actively changing their minds and considering new viewpoints. An indicator is unstable if ≥1 subgroup is unstable according to the Wilcoxon test. Stability was considered a prerequisite to consensus measurement, as unstable indicators warrant further exploration in additional Delphi rounds.[28,34,35]

Consensus of each indicator was determined by the interquartile range (IQR) and median responses across all respondents. Consensus is attained if an indicator is stable and IQR is ≤1.[28,36] Kruskall-Wallis' H-test was used in subgroup analysis to determine whether there were significant disagreements between groups for each indicator, and Dunn's test was used to establish which groups were disagreeing. It is possible for an indicator to have consensus based on IQR even if a significant disagreement was found between 2 groups.

Three inclusion criteria were utilized in the creation of the final value framework: first, a simple majority (>50%) of all participants rating the indicator as "important" or "very important," second, stability across groups, and, third, an IQR ≤1, indicating overall consensus among all participants. The final value framework was split into levels of "importance" and included indicators rated as "important" or "very important" according to majority ratings in R3.

Results

The collaborative framework began with 34 indicators based on a literature review, was adapted to 45 indicators based on participant contributions, and resulted in a final framework comprising 33 stable indicators with consensus split between "very important" and "important" value judgments.

Web-Delphi Panel Results

212 people were contacted to participate, of whom 129 accepted, including 25 patients, 32 HCPs, 29 supply-side actors, 19 decision makers, and 24 influencers. R1 had 101 participants, R2 had 91 participants, and R3 had 79 participants indicating a 61% overall retention rate (Appendix Table 7 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008). Most participants were male, white, and between the ages of 30 to 60 (Appendix Table 8 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008). The final value framework of 33 indicators ranked and split by importance scores across all participants (Table 2) excludes unstable indicators and those without consensus according to the IQR of all responses. Seventeen indicators are classified as "very important" and 16 as "important."

Indicator alteration from R1 thematic analysis

Of the 34 starting indicators, 15 remained the same, 13 were altered, and 6 were removed based on R1 qualitative feedback (Appendix Table 6 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008). Seventeen additional indicators were added from participant suggested themes, resulting in 45 indicators used in R2 and R3.

Consensus measurement

Agreement within groups. Inter-rater agreement was calculated for each stakeholder and country group. For stakeholders, everyone but influencers (moderate agreement with Ky = 0.57) showed substantial agreement in R2 ($0.61 < Ky \leq 0.71$), and all groups showed substantial agreement in R3 ($0.66 < Ky \leq 0.74$) (Table 3[26,33]). For country groups, the United Kingdom and the United States maintained significant agreement between the 2 rounds and Germany's participants moved from moderate (Ky = 0.56) to substantial agreement (Ky = 0.62) between R2 and R3 (Table 3[26,33]), suggesting that individuals within the same group have similar value sentiments.

Stability between rounds. Instability illustrates participants were changing their minds about the value of certain indicators between R2 and R3. A total of 9 indicators demonstrated modest, but statistically significant, instability between R2 and R3 in which value judgments shifted positively toward either "important" or "very important." All unstable indicators demonstrated high overall importance scores and did not threaten our results.

- *Stability among stakeholder groups.* Seven (out of 45) indicators were shown to lack stability in either the supply-side or influencer groups (Table 4). Three were within the user preferences domain with Z-values ranging from −1.98 to −2.00. The data rights indicator "Sharing of identifiable data…" was unstable for both influencers (Z = −2.23) and supply-side groups (Z = −1.99).
- *Stability among country groups.* Overall, most indicators were stable among country groups with only 4 (out of 45) presenting instability (Table 4). German and British participants were each unstable for 2 indicators and Americans for 1 (Table 4). See Appendix section 3.1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008 for more information.

Agreement across all respondents. IQR was used to measure consensus across all respondents. A total of 33 stable indicators had consensus based on IQR, 17 were rated "very important" and 16 rated "important" (Table 2). Three indicators had statistically significant dissensus across all respondents based on an IQR, of which 2 within the economic characteristics domain and 1 in the user preferences domain (Table 6). Importantly, dissensus was because of high ratings of neutrality and no instances of dissensus were recorded in which negative ratings exceeded 7% of overall participants. There was no dissensus between positive ("important" or "very important") and negative judgments ("not at all important" or "low importance"); therefore, overall disagreement does not result from notable controversy among participants.

The economic characteristics indicator "Pre-marketing approval, innovation incentives exist for supply-side actors …" had the lowest rated importance of the final indicators, with 61% overall positive ratings. These findings illustrate that the broad stakeholder network may not perceive innovation incentives as critical as other indicators. Conversely, the technical and security indicator "Complies with local data protection regulations" had the highest rated importance with 100% positive ratings, 81% of which were "very important."

It is important to note that the Likert scale judgments by stakeholders represent independent value scores of each indicator. Although it is possible to denote objective comparisons between indicators' overall scores, the Delphi exercise does not ask participants to directly compare each indicator's level of importance. Therefore, further exercises utilizing direct value comparisons would be required to understand each indicator's relative weights. Table 5 displays stability and consensus for all indicators.

Subgroup Analysis

The Kruskall-Wallis H-test and Dunn's test were used to compare stakeholder and country group medians to locate

Table 2. Final value framework showcasing indicators ranked according to rating of majority of participants.

Value domain	Ranked stable indicators with "Very Important" rated by majority	% of participants rating "Very important"
Technical and Security	Complies with local data protection regulations	81
Data rights and governance	Adheres to strong information governance standards, including processes for data protection violations	80
Data rights and governance	The storage and processing of data corresponds to the regional legal requirements for data privacy	78
Technical and security	Data has a high degree of integrity and credible provenance	77
Clinical characteristics	Evidence of clinical benefit based on patient-centered endpoints	71
Data rights and governance	Policies on data privacy, sharing, collection, and commercialization are clearly communicated to all users	70
Clinical characteristics	Improves patient adherence to treatment	68
Economic characteristics	Affordability to the patient	67
Economic characteristics	Affordability to the system	66
Clinical characteristics	Clinical risk management in place	65
Clinical characteristics	Improves personal health engagement and patient activation, including improved decision-making abilities	57
Technical and security	Security specifications are simply and transparently communicated to all users, including detailed information about updates	56
Clinical characteristics	Integrates with and improves clinical processes	56
Technical and security	Data interoperability	51
Health Inequalities	Does not exacerbate existing health inequalities	46
User preferences	Technical and user support	44
Data rights and governance	Data are user owned	43
Value domain	**Ranked stable indicators with "Important" rated by majority**	**% of participants rating 'Important'**
Technical and Security	Systems are in place for continued product development and security updating after product release	53
User preferences	Where relevant, offers customizable integration with other solutions to facilitate management of multiple co-occurring conditions	52
User preferences	Provides an attractive/engaging experience for the end-user	52
Health inequalities	Supports digital literacy	51
User preferences	User is able to choose communication method as a result of personal preferences	51
Technical and security	Uses multifactor authentication	49
Clinical characteristics	Improves communication and information sharing	48
Clinical characteristics	Supports and sustains lifestyle changes	48
Technical and Security	Where applicable, convenient and sustainable device consumables	48
Data rights and governance	Systems are in place for health data (RWD) and its analytics (RWI) to contribute to real-world evidence (RWE) generation while adhering to privacy standards. (Including postmarketing approval)	47
Technical and Security	Where applicable, ability for patient users to input data	47
Technical and security	Capable of working and storing data offline and then syncing when internet restored, where clinically appropriate.	46
Health Inequalities	Helps reduce socioeconomic health inequalities	46
Clinical characteristics	Where appropriate, utilizes real-world data for proof of benefit	41
User preferences	Connection to peer support where appropriate	39
Economic characteristics	Pre-marketing approval, innovation incentives exist for supply-side actors, such as opportunities for managed entry and risk-sharing agreements	38

Source: The authors from analysis of Delphi data.
Note. Table only includes indicators ranked "Very important" and "Important," which were stable with consensus across all participants, according to the Wilcoxon test and IQR calculations. No indicators had a majority rating of "neutral," "little importance," "not at all important." Indicators were considered to have consensus if their IQR < 2.
IQR indicates interquartile range; RWD, real-world data; RWI, real-world insights.

Table 3. Inter-rater agreement within stakeholder and country groups across all indicators.

	Round 2				Round 3			
	Ky	95% CI		Benchmark Interval	Ky	95% CI		Benchmark Interval
Stakeholder group								
Patient users	0.71*	0.67	0.76	Substantial agreement	0.74*	0.70	0.79	Substantial agreement
Health care professionals	0.61*	0.57	0.66	Substantial agreement	0.67*	0.63	0.72	Substantial agreement
Supply-side	0.61*	0.55	0.67	Substantial agreement	0.69*	0.63	0.74	Substantial agreement
Decision makers	0.66*	0.62	0.70	Substantial agreement	0.70*	0.65	0.74	Substantial agreement
Influencers	0.57*	0.52	0.63	Moderate agreement	0.66*	0.60	0.71	Substantial agreement
Country group								
Germany	0.56*	0.50	0.62	Moderate agreement	0.62*	0.56	0.68	Substantial agreement
United Kingdom	0.67*	0.61	0.72	Substantial agreement	0.72*	0.67	0.77	Substantial agreement
United States	0.67*	0.63	0.71	Substantial agreement	0.71*	0.68	0.74	Substantial agreement

Source: The authors from analysis of Delphi data.
Note. Inter-rater agreement measured by the Gwet's agreement coefficient with linear weights. Benchmark scale of the level of agreement as suggested by Landis and Koch[33]: Coef. < 0.00 Poor agreement; 0.00 < Coef. ≤ 0.20 Slight agreement; 0.20 < Coef. ≤ 0.40 Fair agreement; 0.40 < Coef. ≤ 0.60 Moderate agreement; 0.60 < Coef. ≤ 0.80 Substantial agreement; 0.80 < Coef. ≤ 1 Almost perfect agreement.[26]
Coef. indicates coefficient.
*$P < .01$.

statistically significant disagreements between groups. No significant disagreements between stakeholder groups were found for any indicators. Significant disagreements were found between country groups for 2 indicators: "Helps reduce socioeconomic health inequalities" and "Integrates with and improves clinical processes." Germans were found to disagree with both the British and Americans regarding "… socioeconomic health inequalities," with only 67% of German participants offering positive judgments (Table 6). Germans also disagreed with the British for "… improves clinical processes," with Germans and British offering 81% and 100% positive judgments, respectively.

Table 4. Significantly unstable indicators between R2 and R3, by stakeholder and country group.

Value domain	Indicator	Wilcoxon test Z value (>*P* value)	Stakeholder group
Data rights and governance	Sharing of identifiable data to outside commercial entities is not permissible	−2.230 (.0258) −1.997 (.0458)	Influencer Supply-side actor
Technical and security	Sustainable data architecture	−1.997 (.0458)	Supply-side actor
Clinical characteristics	Where applicable, improves quality of life for patients and carers	−2.000 (.0455)	Influencer
Economic characteristics	Long-term cost effectiveness to the system	−2.000 (.0455)	Influencer
User preferences	Relevance of the solution to the targeted user group	−1.977 (.0458)	Supply-side actor
	Ease of adoption and use with minimal training	−2.000 (.0455)	Supply-side actor
	Technology is offered in multiple languages	−2.000 (.0455)	Influencer
Value domain	**Indicator**	**Wilcoxon test Z value (>*P* value)**	**Country group**
Data rights and governance	Upon discontinuation of use, methods of de-registering and deleting data are clearly communicated to the user and are not difficult to achieve	−2.000 (.0455) −2.233 (.0255)	United States United Kingdom
Technical and security	Processes in place to prevent unauthorized access to patient and outcomes data	−1.988 (.0458)	United Kingdom
Economic characteristics	Long-term cost effectiveness to the system	−2.000 (.0455)	Germany
User preferences	Ease of adoption and use with minimal training	−1.988 (.0457)	Germany

Source: The authors from Delphi data analysis.
Note. Value aspects are only displayed if there is a significant difference in the median values for a stakeholder group between round 2 and 3. Stakeholder group sample sizes are between 13 and 18 participants. Country group sample sizes are between 19 and 22 participants.
R2 indicates round 2; R3, indicates round 3.

Table 5. R2 and R3 overall statistical results for all proposed indicators.

Value domain	Indicator	R3 overall rated importance, %	Stability		Whole group consensus (based on IQR)*
			Across all Stakeholder groups	Across all Country groups	
Health Inequalities	Supports digital literacy	84	√	√	√
	Does not exacerbate existing health inequalities	81	√	√	√
	Helps reduce socioeconomic health inequalities	82	√	√	√
Data rights and governance	Adheres to strong information governance standards, including processes for data protection violations	100	√	√	√
	The storage and processing of data corresponds to the regional legal requirements for data privacy	94	√	√	√
	Data are user owned	86	√	√	√
	Policies on data privacy, sharing, collection, and commercialization are clearly communicated to all users	99	√	√	√
	Upon discontinuation of use, methods of de-registering and deleting data are clearly communicated to the user and are not difficult to achieve.	95	√	X	-
	Systems are in place for health data (RWD) and its analytics (RWI) to contribute to real-world evidence (RWE) generation while adhering to privacy standards. (Including postmarketing approval)	89	√	√	√
	Sharing of identifiable data to outside commercial entities is not permissible	87	X	√	-
Technical and security	Complies with local data protection regulations	100	√	√	√
	Uses multifactor authentication	72	√	√	√
	Systems are in place for continued product development and security updating after product release	96	√	√	√
	Security specifications are simply and transparently communicated to all users, including detailed information about updates.	89	√	√	√
	Data has a high degree of integrity and credible provenance.	97	√	√	√
	Processes in place to prevent unauthorized access to patient and outcomes data	100	√	X	-
	Data Interoperability	94	√	√	√
	Where applicable, ability for patient users to input data	86	√	√	√
	Capable of working and storing data offline and then syncing when internet restored, where clinically appropriate.	89	√	√	√
	Where applicable, convenient and sustainable device consumables.	71	√	√	√
	Sustainable data architecture	85	X	√	-
Clinical characteristics	Evidence of clinical benefit based on patient-centered endpoints.	96	√	√	√

continued on next page

Table 5. Continued

Value domain	Indicator	R3 overall rated importance, %	Stability		Whole group consensus (based on IQR)*
			Across all Stakeholder groups	Across all Country groups	
	Clinical risk management in place	95	√	√	√
	Integrates with and improves clinical processes	94	√	√	√
	Improves patient adherence to treatment	92	√	√	√
	Improves personal health engagement and patient activation, including improved decision-making abilities	97	√	√	√
	Supports and sustains lifestyle changes	90	√	√	√
	Where applicable, improves quality of life for patients and carers.	91	X	√	-
	Improves communication and information sharing	91	√	√	√
	Where appropriate, utilizes real-world data for proof of benefit.	81	√	√	√
Economic characteristics	Adheres to value-based care methodology	75	√	√	X
	Pre-marketing approval, innovation incentives exist for supply-side actors, such as opportunities for managed entry and risk-sharing agreements	61	√	√	√
	Sustainable system improvements through resource optimization	73	√	√	X
	Affordability to the patient	92	√	√	√
	Affordability to the system	92	√	√	√
	Long-term cost effectiveness to the system	97	X	√	-
User preferences	Relevance of the solution to the targeted user group	100	X	√	-
	Technical and user support	92	√	√	√
	Ease of adoption and use with minimal training	96	X	X	-
	Multi stakeholder design, development, and implementation	73	√	√	X
	Connection to peer support where appropriate.	61	√	√	√
	Provides an attractive/engaging experience for the end-user	82	√	√	√
	Where relevant, offers customizable integration with other solutions to facilitate management of multiple co-occurring conditions.	85	√	√	√
	User is able to choose communication method as a result of personal preferences	78	√	√	√
	Technology is offered in multiple languages	86	X	√	-

Note. Overall rated importance is the percentage of all respondents who rated the indicator as either 'important' or 'very important' in R3. 'X' indicates that ≥1 group had significant instability (column 4) or dissensus (column 5) for that indicator. '√' indicates that all groups were stable (column 4) or had consensus (column 5) for that indicator.
IQR indicates interquartile range; RWD, real-world data; RWI, real-world insights.
*'-' indicates that consensus is unable to be determined for unstable indicators.

Table 6. Indicators with no consensus across all respondents based on IQR.

Indicators with no consensus across all respondents based on IQR			
Value domain	Indicator	Median	IQR (SD)
Economic characteristics	Adheres to value-based care methodology	Important	2 (±0.816099)
	Sustainable system improvements through resource optimization	Important	2 (±0.9305657)
User preferences	Multistakeholder design, development, and implementation	Important	2 (±0.8396225)
Significantly different country group indicator responses			
Value domain	Indicator	Kruskal-Wallis H-test*	Dunn's test for pairwise comparisons† and % rated "very important"
Clinical characteristics	Integrates with and improves clinical processes	7.871‡	Germany (38%) - UK (73%)§
User preferences	Ease of adoption and use with minimal training	6.631‡	UK (91%) - US (53%)‡

Source: The authors from analysis of Delphi data.
Note. The Likert scale assumes that the strength of importance is linear, ie, the strength increases in equal intervals of a numerical value of 1. Participants rated indicators on a Likert scale of importance (1 = not at all important and 5 = very important). For overall consensus, indicators were only included if their IQR > 1, that is, there was dissensus among the group regarding that indicator. For group differences, only stable significantly different group differences were showed. Sample sizes range from 19 to 27. Percentages display the percentage of the percentage of each significantly different country group which rated the indicator as "very important."
IQR indicates interquartile range; UK, United Kingdom; US, United States.
*Chi-square adjusted for ties with 2 degrees of freedom.
†Dunn's test (pairwise multiple comparisons using the Bonferroni adjustment)
‡$P < .05$.
§$P < .01$.

Discussion

HTA value indicators continue to evolve from focusing on medical technologies' safety, efficacy, and costs to including a broader set of criteria that include societal aspects, patient perspectives, and equity in access,[37,38] focusing on promoting value-based care.[39,40] In this study, and as part of a multistakeholder co-creation process, Delphi participants suggested indicators that assess aspects of value for DHTs that go beyond traditional HTA indicators and those included in the study countries' frameworks. Naturally, the biggest differences in HTA for DHTs compared with standard HTA approaches are inherent to digital technology and data collection. The data rights and governance, technical and security, and user preferences domains encompass indicators very specific to digital health.

During R1, numerous responses were received in the user preferences and data rights and governance domains, which were very strongly worded and primarily from patients, illustrating how much patients value owning and inputting their health information. The data rights and governance indicator "Data are user owned" was created because of these R1 responses and had consensus, but this is not reflected in any study country's regulations or value assessment frameworks. Indeed, there is no consensus around who owns certain medical data[41] or whether it can be owned at all.[42] In the United States, the Health Insurance Portability and Accountability Act (HIPAA)[43,44] classifies data rights by the entity who collects and uses the data, instead of the ability of those data to reveal sensitive information about the data subject. Conversely, the General Data Protection Regulation, which relates to Germany and the United Kingdom, covers all personally identifiable data regardless of who collects, stores, or owns them because data governance rules are classified by "rights of the data subject.[43]" Data protection regulations tend to focus on a patient's right to access their health information, rather than the right to own and custody the data themselves. It is well established in the study countries that patients have a right to access their health data, but ownership is a different matter. This highlights yet another reason why HTA for DHTs needs to be approached differently to standard HTA: the inability to meet traditional evidence standards increases the need to use real-world evidence (RWE) to prove value; therefore, whose intellectual property are the collected data, and, further, who can benefit from using those data, regardless of their identifiability?

The differences identified between stakeholders' value preferences and the laws that govern them reflect a need for a comprehensive policy approach involving a mix of regulation and value assessment incentives. Although some indicators can be part of HTA appraisals, others may require institutional intervention with data ownership being a clear example. Many question whether digital health-specific frameworks are necessary, or whether existing pharmaceutical frameworks, particularly those which include social value elements,[45] can be adopted with minor adaptations. Nevertheless, the inherent differences of DHTs versus pharmaceuticals means that additional indicators cannot simply be added to existing pharmaceutical frameworks; each value domain and its indicators would need adaptation to suit the digital context. Among other issues, widespread data collection, the use of remote monitoring, and RWE to meet evidence standards illustrate that DHTs not only bring forward new domains of value to be assessed but also point out gaps in regulatory needs. By understanding what different stakeholders value, healthcare decision makers can introduce policies

promoting the creation of solutions that can meet the needs of multiple stakeholders. To do this, facilitation of multistakeholder discussions is essential.

Multi-stakeholder involvement in policy creation is one proven way to address varying opinions on value.[46] In our study, the user-preferences indicator "Multi-stakeholder design, development, and implementation" did not have consensus based on IQR because of high neutrality ratings. This reflects a notable juxtaposition: despite participating in a study based on multi-stakeholder co-creation, not all respondents believe that adoption of such an approach is an important criterion with which to assess DHTs. In a patient-facing DHT context, this is surprising because one might think multistakeholder input in creating digital solutions would be particularly valuable. In the National Institute for Health and Care Excellence (NICE) Evidence Standards Framework, having relevant clinical/social care professionals involved in the design, development and testing is a minimum evidence standard.[7]

Two indicators in the economic characteristics domain also had dissensus: "… value-based care methodology" and "Sustainable system improvements …" This is surprising considering DHTs are well suited to support a shift to value-based care and offer resource optimization benefits.[47] One might argue dissensus highlights resistance to change because people tend to be wary of changes they may not entirely understand. In the subgroup analysis, significant differences were found between country groups but not stakeholder groups. It is not surprising that disagreement seem to stem from country-specific issues and differences in system archetypes. The disagreements demonstrate a lack of shared vision about how DHTs should be assessed and, potentially, a lack of understanding about potential health system benefits. Policy changes around incorporation of value-based care methodologies and increased patient involvement are widespread[39,40,48,49] and will require a shift in stakeholder mindset to build rewarding market pathways for such initiatives.

Digital HTA policies and regulations, although formed based on today's technologies, will shape the path of innovation for future DHTs. Within HTA policies, aligned incentives are needed to improve digital health infrastructure and enable market entry of innovative DHTs, which will allow for technical solutions that can meet the needs of key stakeholder groups. Policies are needed that enable the creation of technical innovations that can, for example, allow patients to input information credibly, solve issues around data custody, ownership, and privacy, and facilitate RWE creation for HTAs and population health management initiatives. Keeping in mind that DHTs require a different approach to assessment than other medical technologies, it is important to align stakeholder opinions on value characteristics to create regulations and assessment methods that will pave the way for continued digital innovation that promotes value-based healthcare for all.

Limitations

Our study is not without limitations. First, some indicators demonstrated dissensus and instability: dissensus was because of high neutrality ratings, and conclusions should not be made when there is participant instability, necessitating additional Delphi rounds.[28,34,35] Second, in R1 of the Delphi exercise, participants could only alter indicators not value domains, which may limit creativity of participants and introduce bias. Nevertheless, the starting framework was the result of an extensive literature review coupled with expert input from stakeholders. Third, results from this exercise pertain to 3 specific health systems and are not generalizable; nonetheless, they provide a useful starting point for similar exercises in a wider range ot settings. Fourth, STATA does not support powered analysis for non-parametric tests; therefore, effect sizes were not able to be calculated. Finally, during recruitment, some, not all, participants were paid by professional recruiters to participate. This was done to ensure an equal number and therefore a comparable sample of participants from our 3 study countries and 5 stakeholder groups. There is an ongoing debate in the research community about the type of bias introduced by paying or not paying participants, please see Appendix section 2.2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.06.008 for further information.

Conclusions

By conducting a scoping review and a web-based Delphi exercise, we have identified several indicators for the assessment of regulated patient-facing DHTs that incorporate the value concerns of stakeholders. DHTs continue to highlight the need for unique assessment methods, such as multistakeholder collaborative value frameworks and multicriteria approaches, because their differing abilities to produce evidence require a different and innovative approach.[39,40] A move toward comprehensive assessment frameworks for these technologies has led to the introduction of new value domains and indicators, which help to address these challenges, some of which may require regulatory intervention. This technology-agnostic and collaborative value framework, comprising 33 indicators, aims to assist not only decision makers in effectively assessing DHTs but also innovators creating technologies that deliver benefit to patients and, more generally, to health systems and society.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2023.06.008.

Article and Author Information

Accepted for Publication: June 20, 2023

Published Online: July 27, 2023

doi: https://doi.org/10.1016/j.jval.2023.06.008

Author Affiliations: Department of Health Policy and LSE Health - Medical Technology Research Group, London School of Economics and Political Science, London, England, UK (Haig, Main, Chávez, Kanavos).

Correspondence: Madeleine Haig, MSc, Department of Health Policy and LSE Health - Medical Technology Research Group, London School of Economics and Political Science, London, England, UK. Email: m.haig@lse.ac.uk

Author Contributions: *Concept and design:* Haig, Main, Kanavos
Acquisition of data: Haig, Main
Analysis and interpretation of data: Haig, Main, Chávez
Drafting of the article: Haig, Main, Chávez
Critical revision of the article for important intellectual content: Haig, Kanavos
Provision of study materials or patients: Haig, Main, Chávez, Kanavos
Obtaining funding: Kanavos
Administrative, technical, or logistical support: Haig, Main, Kanavos
Supervision: Kanavos

Conflict of Interest Disclosures: Funding for this work was provided by Sanofi. Dr Kanavos reported contracts with Novartis, Takeda, Merck Sharp & Dohme, and Amgen to the LSE; consulting fees from PhRMA; and honoraria from EFPIA, Merck Sharp & Dohme, GlaxoSmithKline, Janssen, and Amgen across other research efforts in the last 36 months. No other disclosures were reported.

Funding/Support: This work was supported by funding from Sanofi S.A. to the LSE.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the article; or decision to submit the article for publication.

Acknowledgment: We would like to thank our Delphi participants for providing their time and expertise in this study. Additionally, we would like to thank participation and feedback we received from participants at ISPOR US 2022 and 2023 and an LSE Health Executive Masters Seminar at the London School of Economics and Political Science where we presented early-stage results and discussion points. Thanks are also due to two anonymous referees for providing helpful comments. All outstanding errors are our own.

REFERENCES

1. Fahy N, Williams G. Use of digital health tools in Europe: before, during and after COVID-19. COVID-19 health system response monitor network. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/345091/Policy-brief-42-1997-8073-eng.pdf?sequence=1&isAllowed=y. Accessed August 5, 2022.
2. Hoffman D. Commentary on chronic disease prevention in 2022. National Association of Chronic Disease Directors https://chronicdisease.org/wp-content/uploads/2022/04/FS_ChronicDiseaseCommentary2022FINAL.pdf. Accessed May 19, 2023.
3. Hajat C, Stein E. The global burden of multiple chronic conditions: a narrative review. *Prev Med Rep*. 2018;12:284–293.
4. Kelly BB, Narula J, Fuster V. Recognizing global burden of cardiovascular disease and related chronic diseases. *Mt Sinai J Med*. 2012;79(6):632–640.
5. Kruse CS, Goswamy R, Raval Y, Marawi S. Challenges and opportunities of big data in health care: a systematic review. *JMIR Med Inform*. 2016;4(4):e38.
6. Guo C, Ashrafian H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions—a call for innovative evidence generation approaches. *Npj Digit Med*. 2020;3(1):110.
7. Evidence standards framework for digital health technologies. National Institute for Health and Care Excellence. https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf. Accessed January 24, 2022.
8. Manteghinejad A, Javanmard SH. Challenges and opportunities of digital health in a post-COVID19 world. *J Res Med Sci*. 2021;26:11.
9. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. *Npj Digit Med*. 2019;2(1):38.
10. Evidence standards framework for digital health technologies: user guide. National Institute for Health and Care Excellence. https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/user-guide.pdf. Accessed January 24, 2022.
11. Faulkner E, Holtorf AP, Walton S, et al. Being precise about precision medicine: what should value frameworks incorporate to address precision medicine? A report of the personalized precision medicine Special Interest Group. *Value Health*. 2020;23(5):529–539 [published correction appears in *Value Health*. 2020;23(11):1522].
12. Kidholm K, Ekeland AG, Jensen LK, et al. A model for assessment of telemedicine applications: mast. *Int J Technol Assess Health Care*. 2012;28(1):44–51.
13. Augustovski F, Alfie V, Alcaraz A, García Martí S, Drummond MF, Pichon-Riviere A. A value framework for the assessment of diagnostic technologies: a proposal based on a targeted systematic review and a multistakeholder deliberative process in Latin America. *Value Health*. 2021;24(4):486–496.
14. Phillips KA. Methods for moving evaluation of precision medicine into practice and policy. *Value Health*. 2020;23(5):527–528.
15. Jendl R., Wild C. Framework to support reimbursement decisions for digital health applications. AIHTA. https://eprints.aihta.at/1279/1/HTA-Projektbericht_Nr.134.pdf. Accessed July 6, 2022.
16. McLaren R, Stocks Allen F, Mobasser O, Patel S, Krepchev M. Digital health. Lexology; 2021. https://www.lw.com/admin/upload/SiteAttachments/2021_Digital%20Health_united_kingdom.pdf. Accessed July 13, 2023.
17. Towards harmonised EU landscape for digital health: summary of the roundtable discussions in selected EIT health InnoStars countries. EIT Health. https://eithealth.eu/wp-content/uploads/2023/02/EIT_Health_DiGA_report_Jan2023.pdf. Accessed May 25, 2023.
18. Software as a medical device (SaMD). Food and Drug Administration https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd. Accessed May 31, 2023.
19. Bashi N, Fatehi F, Mosadeghi-Nik M, Askari MS, Karunanithi M. Digital health interventions for chronic diseases: A scoping review of evaluation frameworks. *BMJ Health Care Inform*. 2020;27(1):e100066.
20. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. *Lancet Digit Health*. 2019;1(2):e52–e54.
21. von Huben A, Howell M, Howard K, Carrello J, Norris S. Health technology assessment for digital technologies that manage chronic disease: a systematic review. *Int J Technol Assess Health Care*. 2021;37(1):e66.
22. Kolasa K, Kozinski G. How to value digital health interventions? A systematic literature review. *Int J Environ Res Public Health*. 2020;17(6):2119.
23. Kowatsch T, Otto L, Harperink S, Cotti A, Schlieter H. A design and evaluation framework for digital health interventions. *It - Inf Technol*. 2019;61(5-6):253–263.
24. Beiderbeck D, Frevel N, von der Gracht HA, Schmidt SL, Schweitzer VM. Preparing, conducting, and analyzing Delphi surveys: cross-disciplinary practices, new directions, and advancements. *MethodsX*. 2021;8:101401.
25. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014;67(4):401–409.
26. Belton I, MacDonald A, Wright G, Hamlin I. Improving the practical application of the Delphi method in group-based judgment: a six-step prescription for a well-founded and defensible process. *Tech Forecast Soc Change*. 2019;147:72–82.
27. Meijering JV, Kampen JK, Tobi H. Quantifying the development of agreement among experts in Delphi studies. *Tech Forecast Soc Change*. 2013;80(8):1607–1614.
28. von der Gracht HA. Consensus measurement in Delphi studies review and implications for future quality assurance. *Tech Forecast Soc Change*. 2012;79(8):1525–1536.
29. Dimitrovová K, Vieira A, Oliviera M, Angelis A, Kanavos P, Bana e Costa C. Exploring HTA stakeholders' views on value aspects in the evaluation of new medicines. Progressive.
30. Home. Welphi. https://www.welphi.com/en/Home.html. Accessed November 30, 2022.
31. NVivo. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home. Accessed November 24, 2022.
32. STATA. https://www.stata.com. Accessed November 24, 2022.
33. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
34. Holey EA, Feeley JL, Dixon J, Whittaker VJ. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Med Res Methodol*. 2007;7:52.
35. Hirschhorn F. Reflections on the application of the Delphi method: lessons from a case in public transport research. *Int J Soc Res Methodol*. 2019;22(3):309–322.
36. Culot G, Orzes G, Sartor M, Nassimbeni G. The future of manufacturing: A Delphi-based scenario analysis on Industry 4.0. *Tech Forecast Soc Change*. 2020;157:120092.
37. Sorenson C, Chalkidou K. Reflections on the evolution of health technology assessment in Europe. *Health Econ Policy Law*. 2012;7(1):25–45.
38. Angelis A, Kanavos P. Multiple Criteria Decision Analysis (MCDA) for evaluating new medicines in health technology assessment and beyond: the advance value framework. *Soc Sci Med*. 2017;188:137–156.
39. Bethke M, Gordon R, Elsner N, Varia H. Equipping physicians for value-based care. Deloitte Insights. https://www2.deloitte.com/uk/en/insights/industry/health-care/physicians-guide-value-based-care-trends.html. Accessed August 22, 2022.
40. Barada S, Awada S. Value based healthcare. PWC. https://www.pwc.com/m1/en/publications/documents/value-based-healthcare.pdf. Accessed August 22, 2022.
41. Galvin HK, DeMuro PR. Developments in privacy and data ownership in mobile health technologies, 2016-2019. *Yearb Med Inform*. 2020;29(1):32–43.
42. Liddell K, Simon DA, Lucassen A. Patient data ownership: who owns your health? *J Law Biosci*. 2021;8(2):lsab023.
43. Bari L, O'Neill DP. Rethinking patient data privacy in the era of digital health. Health affairs. https://www.healthaffairs.org/do/10.1377/forefront.20191210.216658/full/. Accessed August 11, 2022.
44. CFR § 160.103 - definitions. Cornell Law School,Legal Information Institute. https://www.law.cornell.edu/cfr/text/45/160.103. Accessed August 22, 2022.
45. Lakdawalla DN, Doshi JA, Garrison Jr LP, Phelps CE, Basu A, Danzon PM. Defining elements of value in health care—a health economics approach: an ISPOR Special Task Force Report [3]. *Value Health*. 2018;21(2):131–139.
46. Co-creation for policy. Participatory methodologies to structure multi-stakeholder policymaking processes. European Institute of Innovation & Technology (EIT). https://eit.europa.eu/library/co-creation-policy-participatory-methodologies-structure-multi-stakeholder-policymaking. Accessed August 22, 2022.
47. Return on health: moving beyond dollars and cents in realizing the value of virtual care. American Medical Association. https://www.ama-assn.org/system/files/2021-05/ama-return-on-health-report-may-2021.pdf. Accessed August 22, 2022.
48. Kamal RN, Lindsay SE, Eppler SL. Patients should define value in health care: a conceptual framework. *J Hand Surg Am*. 2018;43(11):1030–1034.
49. Lloyd-Williams F, Hyseni L, Guzman-Castillo M, et al. Evaluating stakeholder involvement in building a decision support tool for NHS health checks: co-producing the WorkHORSE study. *BMC Med Inform Decis Mak*. 2020;20(1):182.