

Usability Evaluation of SATUSEHAT Mobile Using the System Usability Scale: A Demographic Comparison Among Users in Yogyakarta

Arif Nugroho Triutomo^{1*}
Nita Budiyan²

^{1,2}Medical Record and Health Information Diploma Program, Poltekkes Kemenkes
Yogyakarta, Yogyakarta, Indonesia

*Corresponding author email: arifngrtu@gmail.com

Abstract

SATUSEHAT Mobile provides access to digital health services and personal health information. Usability may affect users' ability to access these services effectively. To evaluate SATUSEHAT Mobile usability and compare System Usability Scale (SUS) scores across demographic groups in Yogyakarta. This descriptive-comparative cross-sectional study involved 99 purposively selected active users. Data were collected using a demographic questionnaire and the 10-item SUS. Descriptive statistics, Mann-Whitney U tests, and Kruskal-Wallis tests were performed using IBM SPSS Statistics version 25. Bonferroni adjustment was applied to pairwise age-group comparisons. The mean SUS score was 68.0 (median = 70.0; SD = 12.1). Item 10 had the lowest mean score (1.73/4), suggesting potential difficulties with initial learning. SUS scores differed significantly across age groups ($H = 8.663$, $p = 0.034$). Bonferroni-adjusted comparisons showed a significant difference between participants aged ≤ 20 years (median = 72.5) and those aged 41–50 years (median = 58.7; adjusted $p = 0.025$). No significant differences were found by gender, employment status, or educational level. SATUSEHAT Mobile usability was around the commonly cited acceptability threshold. Age-related differences and the low Item 10 score highlight the need to evaluate age-inclusive interface design and clearer onboarding guidance.

Keywords: SATUSEHAT mobile; usability; system usability scale; digital health; health information system

1. Introduction

The rapid evolution of information and communication technology has accelerated the digital transformation of the healthcare sector globally and nationally. The Ministry of Health of the Republic of Indonesia has initiated the National Digital Health Transformation agenda to establish an integrated national health data ecosystem through the SATUSEHAT platform. A vital component of this strategy is SATUSEHAT Mobile, an official health application designed to facilitate public access to personal health information and digital health services.

SATUSEHAT Mobile evolved from the PeduliLindungi application, which was previously used during the COVID-19

pandemic. Following the transition into the post-pandemic era, the application was developed to support access to various health services and information, including immunization records, laboratory results, electronic medical record information, and healthcare facility visit histories. By facilitating access to digital health information, SATUSEHAT Mobile has the potential to support patient engagement, personal health monitoring, and more patient-centered healthcare delivery.

However, the successful implementation and sustained adoption of digital health platforms depend not only on the availability of technical features but also on system usability. Usability refers to the extent to which a product can be used by specified users to achieve specified goals

with effectiveness, efficiency, and satisfaction in a specified context of use (ISO 9241 2018). In digital health interventions, usability is particularly important because users differ in their digital literacy, technological experience, and ability to interact with digital platforms (Durmus 2024). Poor usability may create barriers to technology adoption and reduce the effectiveness of digital health services (Hyzy et al. 2022).

High usability can support user acceptance and continued engagement, whereas systems that are difficult to navigate or understand may discourage use despite their potential functional benefits (Bangor, Kortum, dan Miller 2009). Consequently, empirical usability evaluation is an important component of post-implementation assessment for health information systems.

The System Usability Scale (SUS), introduced by John Brooke in 1986, is one of the most widely used standardized instruments for measuring perceived usability. It consists of ten items rated on a five-point Likert scale and produces a composite score ranging from 0 to 100. SUS provides a practical and efficient measure for evaluating perceived usability across software applications, websites, and mobile applications (Lewis 2018). Previous research has also supported the application of usability evaluation methods in digital health environments (Deniz-garcia et al. 2023). In Indonesia, SUS and other usability evaluation methods have been applied to health-related applications, including SATUSEHAT. For example, (Kumaladewi dan Saputra 2023) evaluated the SATUSEHAT application using usability testing and reported user experience problems, including difficulties in understanding system responses, unclear information, and application instability. These findings highlight the importance of evaluating the usability of digital health applications from the user's perspective.

From a Health Information

Management perspective, the usability of SATUSEHAT Mobile is relevant not only to access but also to how individuals interact with and utilize their personal health information. Users need to be able to locate relevant health records, understand the information presented, and use it to support personal health monitoring and communication with healthcare professionals. Difficulties in navigating the application or interpreting the information provided may limit the practical use of available health information. Nevertheless, usability represents the user-interaction dimension of digital health information management and does not, by itself, establish the accuracy, completeness, security, or interoperability of the underlying data.

Access to personal health information through digital platforms may support patient engagement, self-management, communication with healthcare professionals, and participation in healthcare decisions. A systematic review by (Alomar et al. 2024) found that patient access to electronic health records was associated with several dimensions of healthcare engagement, including patient involvement, self-management, healthcare communication, and the use of healthcare resources. Similarly, (Kessel et al. 2024) identified technological factors, individual capabilities, and social and demographic characteristics as factors that may influence patient engagement with health data through digital platforms. These findings suggest that the benefits of digital health services depend not only on the availability of health information but also on the interaction between system characteristics and users' capabilities.

In the context of SATUSEHAT Mobile, clear navigation, understandable labels, and an efficient interaction process may help users locate and utilize available information and features. Conversely, login difficulties, confusing menus, unclear instructions, and slow system responses

may hinder access and discourage continued use. From an HIM perspective, evaluating usability can therefore provide insight into users' experiences when accessing and utilizing digital health information. Nevertheless, usability should be distinguished from other dimensions of health information management, such as data accuracy, completeness, confidentiality, security, and interoperability, which require separate assessments. Usability evaluation focuses primarily on the effectiveness, efficiency, and satisfaction of users when interacting with a system.

Previous studies have evaluated the usability and user experience of SATUSEHAT through different approaches. For example, (Kumaladewi dan Saputra 2023) assessed the application using the System Usability Scale (SUS), identified user experience problems, and proposed recommendations for improvement. Other studies have examined SATUSEHAT usability using mixed methods and alternative evaluation instruments, while more recent research has assessed user experience through the H.E.A.R.T. framework. However, findings from these evaluations do not necessarily establish whether perceived usability, as measured by SUS, differs across demographic groups. A focused examination of demographic variation among SATUSEHAT Mobile users in Yogyakarta is therefore relevant, particularly to determine whether users from different age groups report different levels of perceived usability. This study addresses this question by comparing SUS scores across demographic groups, including age, gender, educational level, and employment status, thereby extending the assessment beyond an overall usability score.

Therefore, this study evaluates the perceived usability of SATUSEHAT Mobile using the System Usability Scale (SUS) and compares SUS scores across demographic groups among users in Yogyakarta. Beyond reporting an overall SUS score, the principal

contribution of this study is its examination of demographic variation in perceived usability, with particular emphasis on age-related differences. The analysis provides empirical evidence of differences in SUS scores across age groups and identifies a significant pairwise difference between users aged ≤ 20 years and those aged 41–50 years. These findings offer a more differentiated understanding of user experience and may inform the development of age-inclusive interfaces and user support strategies for digital health services. As the study employs a cross-sectional design, the findings demonstrate differences in perceived usability but do not establish that age directly causes these differences.

2. Method

2.1 Study Design and Setting

A quantitative descriptive-comparative cross-sectional design was employed to evaluate the perceived usability of SATUSEHAT Mobile and examine differences in usability scores across demographic groups. The descriptive component was used to summarize users' perceived usability through the System Usability Scale (SUS), providing an overall assessment of the application's usability among the study participants. The comparative component enabled the examination of differences in SUS scores across demographic characteristics, including age, gender, educational level, and employment status. This design was appropriate because it allowed the researchers to describe overall perceived usability and statistically compare SUS scores across demographic groups without manipulating variables or introducing an intervention. The cross-sectional approach provided a snapshot of users' perceived usability during the study period, with each participant assessed at a single point in time. However, the findings do not establish temporal relationships or causal effects. Data were collected from March to

November 2025 in the Special Region of Yogyakarta, Indonesia. The target population comprised active SATUSEHAT Mobile users residing in Yogyakarta.

2.2 Sampling and Eligibility Criteria

Because the population size of active SATUSEHAT Mobile users in the Special Region of Yogyakarta was unknown, Cochran's formula for an unknown or large population was used to estimate the minimum sample size. The calculation used a 95% confidence level ($Z = 1.96$), an expected proportion (p) of 0.50, its complement (q) of 0.50, and a margin of error (e) of 0.10. Using the formula $n = Z^2pq/e^2$, the minimum sample size was calculated as 96.04 and rounded up to 97 participants.

Purposive sampling was used to recruit participants who met the predefined eligibility criteria. Data were collected from March to November 2025 across various locations in the Special Region of Yogyakarta, Indonesia. Recruitment was conducted by approaching potential participants directly in public places and distributing the questionnaire link online via WhatsApp. Data were collected using Google Forms and paper-based questionnaires.

Participants were eligible if they (1) were aged 17 years or older, (2) resided in the Special Region of Yogyakarta, (3) had an active SATUSEHAT Mobile account, and (4) had used the application at least twice before participating. This minimum-use criterion was intended to ensure that participants had prior experience with the application and could evaluate its usability based on actual use. Although no upper age limit was imposed, no participants aged over 50 years were included in the final sample.

Responses were screened for eligibility and completeness, and invalid or incomplete questionnaires were excluded. A total of 99 valid responses were included in the final analysis.

2.3 Data Collection Instrument

Primary data were collected using a self-administered questionnaire consisting of demographic questions (age, gender, educational attainment, and employment status) and the Indonesian-language version of the standardized 10-item System Usability Scale (SUS) adapted by (Zahra dan Santoso 2016). The existing Indonesian-language version was used rather than independently translating or adapting the instrument for this study. Each SUS item used a five-point response scale ranging from 1 (strongly disagree) to 5 (strongly agree). The standard SUS item structure and scoring procedure were retained to calculate the overall perceived usability score.

2.4 Data Scoring and Statistical Analysis

SUS scores were calculated according to the standard scoring procedure. For positively worded odd-numbered items, one was subtracted from each response, whereas for negatively worded even-numbered items, each response was subtracted from five. The converted item scores were summed and multiplied by 2.5 to obtain a composite score ranging from 0 to 100. Overall SUS scores were interpreted using established SUS adjective-rating and acceptability guidance, whereas individual item means were interpreted according to the constructs measured by each item rather than assigned unsupported labels such as "Good" or "Poor."

Statistical analyses were performed using IBM SPSS Statistics version 25. Frequencies and percentages were used to summarize participants' demographic characteristics. Overall SUS scores were described using means and standard deviations, as well as medians and interquartile ranges (IQRs), as appropriate. The Shapiro-Wilk test indicated a statistically significant departure from normality ($p = 0.026$); therefore, non-parametric tests were selected for group comparisons. Mann-Whitney U tests were used to compare overall SUS scores between male and female participants and between

employed and unemployed/student participants. Kruskal-Wallis H tests were used to compare overall SUS scores across educational attainment (senior high school/equivalent, Diploma III, Bachelor's degree/Diploma IV, and Master's degree) and age groups (≤ 20 , 21–30, 31–40, and 41–50 years).

Following a statistically significant Kruskal-Wallis H test for age, Dunn's post hoc pairwise comparisons with Bonferroni adjustment were performed to identify which age groups differed significantly. Six pairwise comparisons were possible among the four age groups. The Bonferroni adjustment was applied to account for multiple comparisons and control the family-wise error rate. Statistical significance was set at $p < 0.05$ for the primary group comparisons, whereas Bonferroni-adjusted p -values below 0.05 were considered statistically significant for post hoc comparisons.

Effect sizes were calculated to complement p -values and quantify the magnitude of the observed differences. Rank-biserial correlation was used for Mann-Whitney U comparisons, whereas epsilon-squared was used for Kruskal-Wallis H tests. Effect sizes for significant post hoc pairwise comparisons were reported where applicable.

2.5 Ethical Considerations

This study was conducted with consideration for ethical principles applicable to research involving human participants. Before completing the questionnaire, participants were informed about the study objectives, procedures, voluntary nature of participation, and measures to protect the confidentiality of their responses. Informed consent was obtained from all 99 participants, either electronically or in person, before data collection. Participation was voluntary, and participants were free to decline participation. The collected data were anonymized for analysis and reporting and were used solely for scientific purposes. The researchers take responsibility for the ethical

conduct of the study, including respect for participant autonomy, confidentiality, and data protection. Following consultation with the institutional ethics committee, this statement was prepared by the researchers to document the ethical safeguards implemented in the study.

3. Result and Discussion

3.1 Respondent Demographic Profiles

A total of 99 active SATUSEHAT Mobile users completed the study.

Table 1. Outlines the frequency and percentage distributions of respondent demographic characteristics.

Demographic Category	Frequency	Percentage (%)
Educational Level		
Senior High School / Equivalent	9	9.1
Diploma III (D3)	35	35.4
Bachelor / Diploma IV (D4/S1)	48	48.5
Master's Degree (S2)	7	7.1
Gender		
Male	29	29.3
Female	70	70.7
Age Group		
≤ 20 years	19	19.2
21–30 years	59	59.6
31–40 years	9	9.1
41–50 years	12	12.1
Employment Status		
Employed	59	59.6
Unemployed / Student	40	40.4

Table 1 demonstrates that the study sample was predominantly female (70.7%), aged 21–30 years (59.6%), holding a Bachelor's/D4 degree (48.5%), and currently employed (59.6%).

The item-level SUS results are reported as converted item means on a 0–4 scale. These item means help identify which aspects of perceived usability may warrant further investigation; they are not equivalent to the overall SUS score and should not be categorized using overall SUS adjective-rating thresholds.

Analysis of individual SUS items (scored 0–4 scale prior to composite transformation) provides targeted insights into user experience strengths and weaknesses.

Table 2. Lists item-level mean scores across all 10 statements

Item No.	SUS Statement Focus	Mean Score (0–4)	Construct-level interpretation
1	Frequent use intention (Positive)	3.07	Intention to use frequently
2	Unnecessary complexity (Negative)	2.71	Perceived unnecessary complexity
3	Ease of use (Positive)	3.01	Perceived ease of use
4	Need for technical support (Negative)	2.57	Perceived need for technical support
5	Feature integration (Positive)	3.01	Perceived integration of functions
6	System inconsistency (Negative)	2.55	Perceived system consistency
7	Quick learnability for most people (Positive)	2.76	Perceived ease of learning
8	Cumbersome interface (Negative)	2.84	Perceived cumbersome interaction
9	Confidence in usage (Positive)	2.90	Confidence when using the system
10	Need to learn many	1.73	Perceived learning

Item No.	SUS Statement Focus	Mean Score (0–4)	Construct-level interpretation
	things before use (Negative)		burden before use
-	Overall Average Item Score	2.72	Overall item mean (descriptive only)

Item-level mean scores ranged from 1.73 to 3.07. The highest mean scores were observed for intention to use the application frequently (Item 1, mean = 3.07), perceived ease of use (Item 3, mean = 3.01), and perceived integration of functions (Item 5, mean = 3.01). These results describe participants' responses regarding their intention to use SATUSEHAT Mobile, its ease of use, and the integration of its functions. In contrast, Item 10, which assesses the perceived need to learn many things before getting started, had the lowest mean score (1.73). This finding highlights the learning burden before initial use as a potential area for further investigation. However, these item-level descriptive results should be interpreted according to the constructs measured by each statement and do not, by themselves, establish the underlying reasons for participants' responses.

3.3 Overall System Usability Scale Score and Distribution

The overall composite SUS score was calculated for all 99 participants.

Table 3. Summary of Descriptive Statistics for SUS Scores

SUS Score Parameter	Statistical Value
Total Sample Size (n)	99
Maximum SUS Score Obtained	95.0
Minimum SUS Score Obtained	37.5
Mean SUS Score	68.0
Standard Deviation (SD)	12.1
Median SUS Score	70.0

Interquartile Range (IQR)	12.5
---------------------------	------

Table 4. Distribution of Respondents by SUS Score Category

SUS Category	Score Benchmark Range	Frequency (n)	Percentage (%)
Excellent	> 80.3	14	14.1
Good	68.0 – 80.3	32	32.3
OK / Fair	51.0 – < 68.0	12	12.1
Poor	< 51.0	41	41.4
Total	-	99	100.0

The overall mean System Usability Scale (SUS) score was 68.0 (SD = 12.1; median = 70.0; IQR = 12.5), placing it near the commonly cited SUS benchmark of 68.0. Under the classification adopted in this study, the mean score was at the lower boundary of the “Good” category. However, the overall mean should be interpreted alongside the distribution of individual scores. Of the 99 respondents, 41 (41.4%) fell into the Poor category (<51.0), 12 (12.1%) into the OK/Fair category (51.0–<68.0), 32 (32.3%) into the Good category (68.0–80.3), and 14 (14.1%) into the Excellent category (>80.3). The scores ranged from 37.5 to 95.0, indicating variation in participants’ perceived usability of SATUSEHAT Mobile. Notably, the proportion of respondents in the Poor category suggests that the overall mean does not fully capture the range of usability experiences within the sample. Therefore, the mean score should not be interpreted as evidence that all users perceive the application as easy to use. These findings can be contextualized by previous evaluations of digital health applications. Hyzy et al. (2022) examined SUS benchmarks for digital health applications, providing a relevant reference for interpreting usability scores in this context. In Indonesia, Kumaladewi and Saputra (2023) reported usability challenges in SATUSEHAT, including slow or unclear application responses, unclear information, and unexpected application closure.

Similarly, Dyayu et al. (2023) reported a mean SUS score of 56.0 for PeduliLindungi, indicating lower perceived usability in that study. Although these studies involved different evaluation approaches and contexts, they provide relevant background for understanding usability challenges in Indonesian public health applications. Nevertheless, differences in study populations, evaluation methods, and application characteristics mean that the scores should not be interpreted as a direct comparison of application performance.

Beyond the overall score, the distribution of SUS scores in the present study highlights the importance of examining differences in usability perceptions across demographic groups. In particular, the observed age-related differences suggest that usability experiences may not be uniform across age groups. This finding supports the need to consider users’ characteristics when evaluating digital health applications. However, the results do not establish that age itself causes differences in perceived usability. Further investigation using task-based usability testing or qualitative interviews could help identify specific interface, navigation, or information-related challenges experienced by different groups of users. Such investigations may provide more actionable evidence for improving the usability of SATUSEHAT Mobile.

3.4 Bivariate Analysis: Gender and Employment Status

To examine whether perceived usability varied across binary demographic attributes, Mann-Whitney U tests were conducted.

Table 5. Summarizes the test results for gender and employment status.

Characteristic	Subgroup	n	Median SUS	IQR	Mann-Whitney U	p-value	Effect size
Gender	Male	29	67.5	13.8	876.	0.28	0.13

Characteristic	Subgroup	n	Median SUS	IQR	Mann-Whitney U	p-value	Effect size
	Female	70	70.0	13.1	5	5	6
Employment Status	Employed	59	70.0	15.0	1164.5	0.912	0.013
	Unemployed	40	67.5	11.9			

Female respondents had a median SUS score of 70.0 (IQR = 13.1), compared with 67.5 (IQR = 13.8) among male respondents; the difference was not statistically significant ($U = 876.5$, $p = 0.285$). Employed respondents had a median of 70.0 (IQR = 15.0), compared with 67.5 (IQR = 11.9) among unemployed respondents/students, also not statistically significant ($U = 1164.5$, $p = 0.912$). These results indicate that this study did not detect statistically significant differences for gender or employment status; they do not demonstrate that the groups are equivalent. The results are consistent with the possibility that broad demographic categories do not always correspond to differences in perceived usability, although the present sample and analysis cannot determine which user or interface factors explain individual ratings. Usability may shape how confidently users access personal health information and navigate digital services, but those outcomes were not directly measured in this study.

3.5 Bivariate Analysis: Educational Level and Age Groups

Kruskal-Wallis H tests were used to compare SUS scores across demographic variables with more than two categories.

Table 6. The comparative statistical metrics across educational backgrounds and age group.

Characteristic	Subgroup Category	n	Median SUS	IQR	Kruskal-Wallis H	p-value	Effect size
Educational Level	Senior High School	9	65.0	8.8	3.726	0.293	0.008
	Diploma III (D3)	35	70.0	15.0			
	Bachelor / D4 (S1)	48	70.0	14.4			
	Master's (S2)	7	70.0	20.0			
Age Group	≤ 20 years	19	72.5	10.0	8.663	0.034*	0.060
	21–30 years	59	70.0	15.0			
	31–40 years	9	65.0	10.0			
	41–50 years	12	58.7	28.8			

Educational attainment was not significantly associated with differences in SUS scores ($H = 3.726$, $p = 0.293$). The median SUS score was 65.0 among respondents with a senior high school education and 70.0 among those with Diploma III (D3), Bachelor's/D4 (S1), and Master's (S2) education.

In contrast, age was the only demographic characteristic that showed a statistically significant difference in SUS scores ($H = 8.663$, $p = 0.034$, $\epsilon^2 = 0.060$). The median SUS score was highest among respondents aged ≤20 years (72.5; IQR = 10.0), followed by those aged 21–30 years (70.0; IQR = 15.0), 31–40 years (65.0; IQR = 10.0), and 41–50 years (58.7; IQR = 28.8). The difference in median scores between the youngest and oldest groups was 13.8 points. This pattern indicates lower perceived usability scores among older respondents in this sample. However, the Kruskal-Wallis test identifies only an overall difference across age groups;

post-hoc pairwise comparisons are required to determine which specific groups differ significantly. These findings should not be interpreted as evidence that age caused the observed differences.

3.6 Post-Hoc Analysis of Age Groups

Because the overall Kruskal-Wallis test for age was statistically significant, post hoc pairwise comparisons were conducted using Dunn's procedure with Bonferroni adjustment across the six age-group pairs. The Bonferroni adjustment was applied by multiplying each unadjusted p-value by six, with adjusted p-values capped at 1.0. Adjusted p-values below 0.05 were considered statistically significant.

Table 7. Dunn's pairwise comparisons of SUS scores across age groups with Bonferroni adjustment.

Pair No.	Age Group Pairwise Comparison	Adjusted p-value	Statistical Significance
1	≤ 20 years vs. 21–30 years	1.000	Not Significant
2	≤ 20 years vs. 31–40 years	1.000	Not Significant
3	≤ 20 years vs. 41–50 years	0.025	Statistically Significant*
4	21–30 years vs. 31–40 years	1.000	Not Significant
5	21–30 years vs. 41–50 years	0.120	Not Significant
6	31–40 years vs. 41–50 years	1.000	Not Significant

After Bonferroni adjustment, the only statistically significant pairwise difference was observed between respondents aged ≤20 years and those aged 41–50 years (adjusted $p = 0.025$). The median SUS score was 72.5 in the ≤20-year group and 58.7 in

the 41–50-year group, representing a 13.8-point difference in median scores. The pairwise effect size was $r = 0.514$, indicating a relatively large difference according to conventional benchmarks. The other five pairwise comparisons were not statistically significant after adjustment (adjusted p-values ≥ 0.120). The overall Kruskal-Wallis effect size was $\epsilon^2 = 0.060$. These results indicate an age-group difference in perceived usability in this sample, particularly between the youngest and oldest groups; they do not establish that age caused the difference or identify the mechanisms underlying it.

3.7 Discussion and Design Implications

The overall System Usability Scale (SUS) score was 68.0, providing an overall assessment of perceived usability of SATUSEHAT Mobile among the study participants. Although this score summarizes users' perceptions of the application, it does not identify specific usability barriers. Therefore, the overall SUS result should be interpreted alongside item-level findings and demographic comparisons to provide a more comprehensive understanding of users' experiences. Previous research evaluating SATUSEHAT Mobile using mixed methods and the Post-Study System Usability Questionnaire has also highlighted the importance of evaluating the application's usability using approaches beyond a single overall score (Kumaladewi dan Saputra 2023).

The lowest item mean concerned the perceived need to learn many things before getting started (Item 10, 1.73/4), suggesting that onboarding and user guidance warrant further evaluation. This finding indicates that the initial user experience deserves closer examination, particularly regarding the clarity of instructions and the ease of learning the application's functions. However, the item score alone does not establish which aspects of the interface create difficulties. Further evaluation through task-based usability testing could help identify specific barriers encountered

by users when learning to navigate the application.

Perceived usability differed significantly across age groups (Kruskal-Wallis $H = 8.663$, $p = 0.034$, $\varepsilon^2 = 0.060$). The median SUS score was highest among participants aged ≤ 20 years (72.5), followed by those aged 21–30 years (70.0), 31–40 years (65.0), and 41–50 years (58.7). Dunn's post-hoc analysis with Bonferroni adjustment identified a statistically significant difference between participants aged ≤ 20 and 41–50 years (adjusted $p = 0.025$), with a median difference of 13.8 points. The effect-size estimate complements the statistical test by indicating the magnitude of the overall age-group difference. These findings suggest variation in perceived usability across age groups in this sample, rather than a uniform usability experience among all users (Khamaj dan Ali 2024). The observed age-related differences should be interpreted within the age range represented in this study. Because no participants aged over 50 years were included, the findings do not establish whether similar usability patterns occur among older users.

The observed age-group difference may be related to differences in digital literacy, prior experience with mobile technologies, or familiarity with digital health applications. However, these factors were not directly measured in the present study and cannot be established as explanations for the observed difference. Previous research has not consistently found the same relationship between age and SUS scores across settings. For example, Orfanou et al. (2015) reported no significant relationship between age and SUS scores in a learning management system (Orfanou, Tselios, dan Katsanos 2015). Differences between studies may reflect the application context, interface design, prior user experience, or sample characteristics, although the present study cannot determine which factors account for the observed age-group difference.

Research specifically examining mobile health applications among older adults has also identified potential usability and accessibility challenges, including small text,

complex navigation, and unintuitive interface design. These findings provide a relevant context for considering whether interface accessibility may contribute to differences in perceived usability across age groups. Nevertheless, the findings from other applications cannot be assumed to explain the age-group differences observed in SATUSEHAT Mobile. Further investigation is required to determine whether similar challenges occur among its users. Although older adults were not represented in this study, age-inclusive design remains an important consideration for digital health applications. Future evaluations should specifically examine the accessibility and usability needs of older users.

From a Health Information Management perspective, usability is relevant to how users interact with digital health information and services. A usable interface may help individuals navigate an application, locate their personal health information, understand available functions, and complete tasks within a digital health system. Conversely, difficulties with navigation, terminology, or task flow may create barriers to accessing health information and interacting with digital health functions (Sanjaya, Kusuma, dan Ristya 2025).

This perspective is particularly relevant to SATUSEHAT Mobile because it functions as a channel for users to access personal health records and other digital health services. Research using the HEART framework to evaluate SATUSEHAT Mobile has identified challenges involving technical performance and account verification, alongside opportunities to improve engagement, adoption, and task success. These findings suggest that digital health implementation requires attention not only to perceived usability but also to the practical barriers users encounter when accessing and using application functions.

Practical implications of the present findings include evaluating clearer onboarding instructions, simplifying navigation and labels, improving text legibility, and

providing adjustable font options. These should be treated as potential design priorities rather than interventions proven effective by the present study. Importantly, this study measured perceived usability using the SUS and did not directly assess information-access success, actual service utilization, digital literacy, or cognitive and visual factors. Therefore, the implications for access to personal health information and use of digital health services should be interpreted as potential consequences of usability difficulties, not as outcomes demonstrated by this study.

Future usability evaluations should involve users from different age groups and combine SUS assessments with task-based usability testing. Future studies should purposively recruit participants aged over 50 years and include a broader range of older users. Combining SUS assessments with task-based usability testing and qualitative interviews may help identify age-related accessibility barriers and specific usability challenges. Relevant measures could include task completion, errors, time on task, and users' ability to locate and understand personal health information. Such evaluations would help determine whether perceived usability differences correspond to observable difficulties in interacting with digital health information systems.

4. Conclusion

This study found that the usability of SATUSEHAT Mobile was near the commonly cited acceptability threshold. The lowest mean score was observed for Item 10, indicating that initial learning and user guidance warrant further evaluation. Usability scores differed significantly across age groups, with lower scores observed among participants aged 41–50 years than among those aged ≤20 years. No significant differences were found by gender, employment status, or educational attainment. These findings highlight the importance of age-inclusive interface design and clearer onboarding to improve user

experience. Future studies should include participants from a broader age range, particularly those aged over 50 years, and employ additional usability testing to identify specific accessibility barriers and evaluate potential improvements.

5. Acknowledgements

The authors express their sincere gratitude to Poltekkes Kemenkes Yogyakarta for financial support and institutional facilitation provided throughout this research. Special appreciation is extended to all field data collectors, administrative assistants, and the 99 study respondents in Yogyakarta whose participation made this research possible.

6. References

- Alomar, Dalia, Maryam Almashmoum, Iliada Eleftheriou, Pauline Whelen, dan John Ainsworth. 2024. "The Impact of Patient Access to Electronic Health Records on Health Care Engagement: Systematic Review." *JMIR Publication* 26. doi: <https://doi.org/10.2196/56473>.
- Bangor, A., P. Kortum, dan J. Miller. 2009. "Determining what individual SUS scores mean; adding an adjective rating." *Journal of usability studies* 4(3):114–23.
- Deniz-garcia, Alejandro, Himar Fabelo, Antonio J. Rodriguez-almeida, Las Palmas, dan De Gran Canaria. 2023. "Quality , Usability , and Effectiveness of mHealth Apps and the Role of Artificial Intelligence : Current Scenario and Challenges Corresponding Author :". *JOURNAL OF MEDICAL INTERNET RESEARCH* 25. doi: 10.2196/44030.
- Durmus, Ayhan. 2024. "The influence of digital literacy on mHealth app usability : The mediating role of patient expertise." *DIGITAL HEALTH* 10:1–13. doi: 10.1177/20552076241299061.
- Dyayu, Arine Lupita, Beny, dan Herti Yani. 2023. "Evaluasi Usability Aplikasi PeduliLindungi Menggunakan Metode

- Usability Testing dan System Usability Scale (SUS) Jurnal Manajemen Teknologi dan Sistem Informasi (JMS)." *Jurnal Manajemen Teknologi dan Sistem Informasi (JMS)* 3:395-404.
- Hyzy, Maciej, Raymond Bond, Maurice Mulvenna, Lu Bai, Alan Dix, Simon Leigh, dan Sophie Hunt. 2022. "System Usability Scale Benchmarking for Digital Health Apps: Meta-analysis." *JMIR mHealth and uHealth* 10(8):1-12. doi: 10.2196/37290.
- ISO 9241. 2018. "Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts." in Vol. 2018. Switzerland.
- Kessel, Robin Van, Sruthi Ranganathan, Michael Anderson, Brian Mcmillan, dan Elias Mossialos. 2024. "International Journal of Medical Informatics Exploring potential drivers of patient engagement with their health data through digital platforms: A scoping review." *International Journal of Medical Informatics* 189(November 2023):105513. doi: 10.1016/j.ijmedinf.2024.105513.
- Khamaj, Abdulrahman, dan Abdulelah M. Ali. 2024. "Examining the usability and accessibility challenges in mobile health applications for older adults." *Alexandria Engineering Journal* 102(May):179-91. doi: 10.1016/j.aej.2024.06.002.
- Kumaladewi, Nia, dan Faiz Rizki Saputra. 2023. "Usability evaluation of satusehat application using usability testing method." *Jurnal Perangkat Lunak* 5:240-50.
- Lewis, James R. 2018. "The System Usability Scale : Past , Present , and Future The System Usability Scale : Past , Present , and Future." *International Journal of Human-Computer Interaction* 34(7):577-90. doi: 10.1080/10447318.2018.1455307.
- Orfanou, Konstantina, Nikolaos Tselios, dan Christos Katsanos. 2015. "Perceived Usability Evaluation of Learning Management Systems: Empirical Evaluation of the System Usability Scale." *International Review of Research in Open and Distributed Learning* 16(2):227-46.
- Sanjaya, Guardian Yoki, Nandani Kusuma, dan Annisa Ristya. 2025. "Unlocking the Potential of Personal Health Records: A H . E . A . R . T . Framework Evaluation of SATUSEHAT Mobile." (March 2023):2023-24. doi: 10.3233/SHTI251309.
- Zahra, Sharfina, dan Harry Budi Santoso. 2016. "An Indonesian adaptation of the System Usability Scale (SUS)." *IEEE* 145-48. doi: 10.1109/ICAC SIS.2016.7872776.