International Journal of Engineering, Management and Humanities (IJEMH)
Volume 6, Issue 2, Mar.-Apr, 2025 pp: 328-332 ISSN: 2584-2145 www.ijemh.com

A/B Testing in UX Design: Evaluating Mobile App UI Variants

Shlok Jain[1], Dhyey Savaliya[2], Utsav Patiyat[3], Akshat Patel[4], Rozal Khan[5], Dr.RashmiPhalnikar[6]

[1,2,3,4,5] Dept. of Computer Engineering and Technology Dr.VishwanathKaradMITWorldPeace University Pune, India

[6]Mentor Dept. of Computer Engineering and Technology Dr.VishwanathKaradMITWorldPeace University Pune, India

---

Date of Submission: 08-04-2025 Date of Acceptance: 19-04-2025

***Abstract*—User experience (UX) plays a pivotal role in de- termining the success and adoption of AI-powered healthcare applications, particularly those delivering critical services like dermatological diagnostics. This paper introduces a structured, generalizable framework for evaluating the effectiveness of user interfaces (UIs) in digital health environments using A/B test-ingmethodology.Acomparativestudywasconductedontwo UI variants (A and B) of a dermatological diagnosis mobile application,aimedatassessingtheimpactofdesignelements ontaskefficiency,usersatisfaction,andtasksuccessrates. A total of 100 participants were randomly assigned to either variant and asked to complete diagnostic tasks under controlled conditions. Quantitative metrics such as task completion time, success rate, and satisfaction scores were collected alongside qualitative interaction data using heatmaps. Results indicated statistically significant improvements in Variant B, which fea- tured enhanced navigation, visual hierarchy, and streamlined workflows. Independent sample t-tests and Bayesian analyses confirmed that users of Variant B completed tasks more quickly, expressed higher satisfaction, and achieved greater success rates compared to Variant A. The study underscores the importanceofuser-centereddesignindigitalhealthcare,demonstrating that thoughtful UI enhancements can significantly influence usability, engagement, and perceived trust. Beyond this specific case, the proposed A/B testing-based framework is adaptable across various domains in digital health, including telemedicine, mental health, and chronic disease management. It offers a data- driven, replicable approach for iterative UX evaluation, helping developerscreateintuitive,accessible,andimpactfuldigitaltools. This work contributes to bridging the gap between advanced AI functionalities and real-world user adoption.**

I. INTRODUCTION

AI-powered healthcare applications are revolutionizing the way individuals access medical support by offering pre- liminary diagnoses remotely. These apps leverage machine learning, computer vision, and natural language processing to provide users with instant assessments, often through image analysis or symptom-based inputs. Particularly in domains such as dermatology, where visual inspection is crucial, AI toolshavedemonstratedstrongpotentialinbridginggaps in healthcare access—especially for users in underserved or geographically isolated areas.

However, despite significant advances in backend AI algo- rithms, the practical impact of these applications is contingent ontheirusability.AhighlyaccurateAIsystemcanberendered ineffective if users struggle to navigate the interface or inter- pret its outputs correctly. Usability and user experience (UX) are increasingly recognized as critical components of system effectiveness, particularly in sensitive domains like healthcare where user trust, clarity, and ease of use are paramount [2],[5].

Poor UI design may hinder users from completing tasks, lower engagement, and ultimately reduce the tool's credi- bility. For instance, disorganized interfaces, unclear call-to- action buttons, or overwhelming workflows may result in abandonment of the application or incorrect usage [2]. The importance of an intuitive, accessible, and responsive UI is therefore central not just to convenience but to medical safety and compliance as well.

Tosystematicallyaddressthesechallenges,wepropose a structured UX evaluation framework based on A/B test-ing. A/B testing has proven to be an effective method for comparing design alternatives and measuring the impact ofUI changes on user behavior, satisfaction, and performance metrics[1],[3],[4].Itiswidelyadoptedinbothindustry and academia for refining mobile interfaces, particularly in usability-drivendomainssuchase-commerce,digitalservices, and increasingly, healthcare [4].

This paper applies the framework to a real-world AI-driven dermatological diagnosis application by comparingtwo UI variants. Using a combination of task-based metrics, heatmapping,andstatisticalanalysis,weassesshowvariations inUIdesignaffectuserengagement,taskcompletiontime,and satisfaction. Our intent is not only to improve one application but to offer a generalizable methodology for designers and researchers seeking to enhance UX in AI-enabled healthcare environments. The broader goal is to bridge the gap between algorithmic sophistication and real-world usability, ensuring these tools are truly accessible and effective for their intended audiences [2], [5].

II. HYPOTHESIS

Awell-structuredandintuitiveuserinterfaceplaysacrucial role in determining the usability and effectiveness of AI-poweredhealthcareapplications.WhileAImodelscanprovide accurate predictions, their real-world impact is dependent on how easily users can interact with them.

H1:EnhancedUIdesign(VariantB)leadstofastertask completion.

H2:Improvednavigationandvisualhierarchyincreaseuser satisfaction.

H3:Streamlinedworkflowsraisetasksuccessrates.

III. METHODOLOGY

A. *Participants*

Fifty participants (n = 100) were recruited from diverse professionalandeducationalbackgrounds.Themeanage was 26.4 (SD = 4.7), with balanced gender representation. Participants were randomly assigned to either UI Variant A or Variant B.

B. *Procedure*

Participants completed a set of diagnostic tasks using the assigned UI. Tasks included image-based diagnosis, symptom entry, and interpreting the results. Interaction data, including tasktimes,successrates,andclickdistribution,werecollected.

C. *TestingEnvironment*

Tests were conducted in a controlled environment using identical smartphones (Android v13, 6.5 inch display). Ex- ternal variables such as lighting and connectivity were stan- dardized.

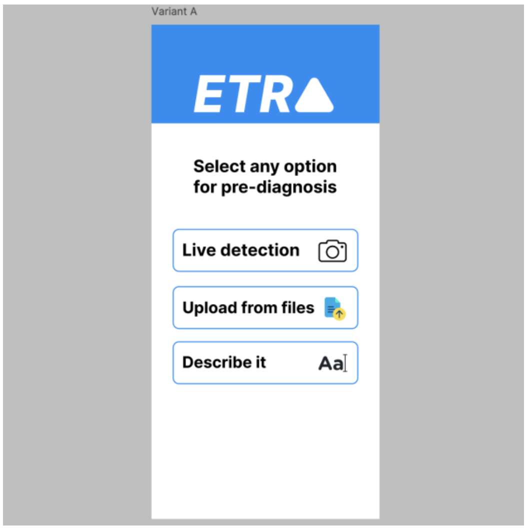

Fig.1.UIVariantAinterfaceshowingnavigationlayout

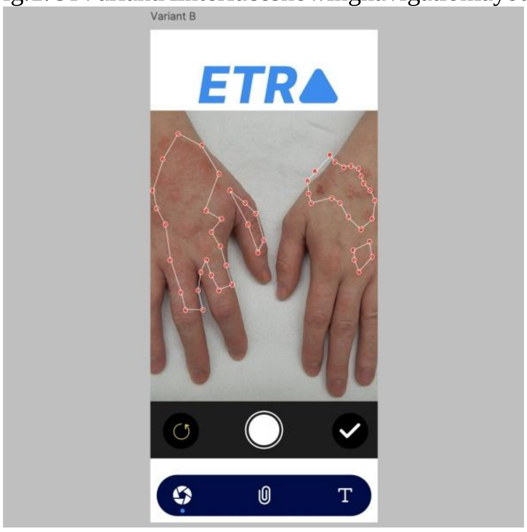

Fig.2.UIVariantBinterfaceshowingnavigationlayout

D. *DataAnalysis*

Significance testing was performed using independent sam- plet-tests(alpha=0.05).Confidenceintervalsandeffectsizes (Cohen’sd)werecalculated.Bayesiananalysiswasperformed to estimate the probability of performance superiority.

IV. RESULTS

A. *TaskCompletiontime*

VariantB(M=2.80,SD=0.60)significantlyoutperformed VariantA (M = 3.60, SD = 0.70), $t(48) = 4.32$, $p < .001$, $d = 1.20$, 95% CI [0.45, 1.15].

B. *UserSatisfaction*

Usinga5-pointLikertscale,satisfactionscoreswerehigher forVariantB(M=4.20,SD=0.50)thanforVariantA(M=

Fig.3.HeatmapshowinginteractioninVariantA

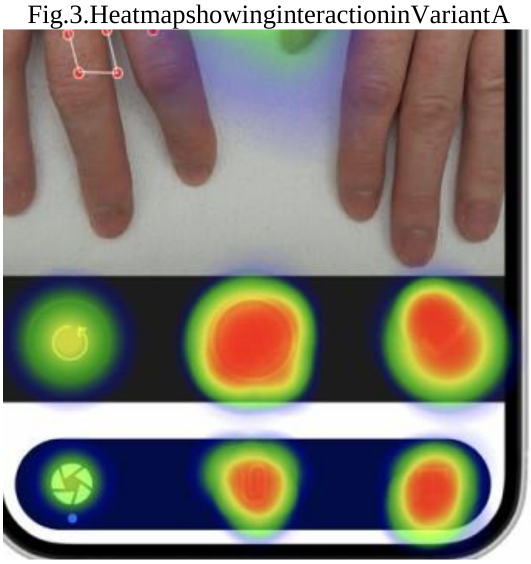

Fig.4.HeatmapshowinginteractioninVariantB

3.50,SD=0.60),t(48)=4.89,$p<.001$,d=1.30,95%CI [0.47,1.07].

C. TaskSuccessRate

To evaluate how interface design influenced task success, we compared the proportion of users who were able to complete all assigned tasks in each variant.

variant B users completed 94% of tasks successfullycomparedto76%invariantA,$\chi^2(1,N=50)=6.81,p=.009$.

D. BayesianAnalysis

Bayes Factor for task completion times exceeded 10, indicating strong evidence in favor of variant B.

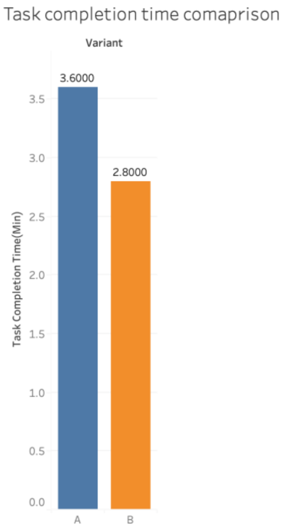

Fig.5.Taskcompletiontime(inmin)comparisonineach variant

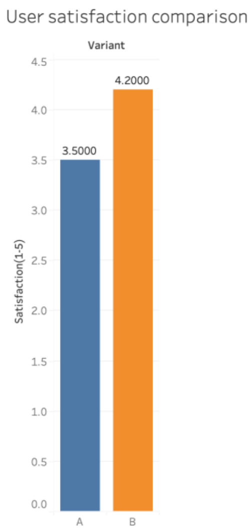

Fig.6.Usersatisfactionscaleofbothvariants

V. DISCUSSION

The findings of this study demonstrate that UI Variant B significantlyimprovesuserperformance,satisfaction,andtask successwhencomparedtoVariantA.Theseimprovementsval- idate the effectiveness of incorporating structured design prin- ciples, such as optimized call-to-action placements, reduced cognitive load, and enhanced visual hierarchy, within AI- driven healthcare applications. More importantly, the results underscorethecriticalroleofusercentereddesigninensuring thattechnologicalinnovations—nomatterhowadvanced—are accessible, intuitive, and practical for end users.

These outcomes lend strong support to the UX evaluation framework proposed in this study. By integrating controlled A/B testing with statistical validation and interaction-based heat mapping, the framework offers a comprehensive and replicablemethodforassessinginterfacequality.Thisap-

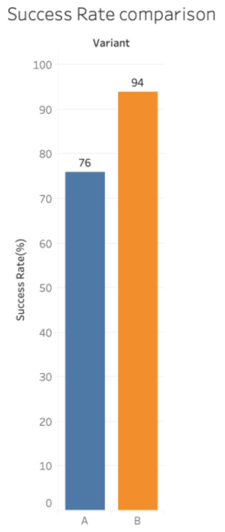

Fig.7.Tasksuccessofbothvariants

proach moves beyond subjective usability impressions and allows developers and designers to make informed, evidence- based decisions that directly impact user outcomes.

Crucially,whilethisframeworkwastestedinthecontext of anAI-powereddermatologicaldiagnosistool,itisnotdomain-specific. The structure and methods used here can be applied across a broad range of digital health platforms, including telemedicine interfaces, mental health applications, chronic diseasemanagementtools,andbeyond.Theunderlying princi- ple—continuous, data-driven UI evaluation—remains univer- sally applicable wherever usability and trust are essential to system success.

The combination of heat mapping, quantitative perfor- mance metrics, and inferential statistics provides a robust, multi-dimensional understanding of user interaction patterns. Heatmaps help uncover behavioral bottlenecks, while statisti-calanalysesensurethatobservedimprovementsarenotdueto chance. This synergy offers a powerful blueprint for iterative designrefinement,enablingcontinuousenhancementofdigital products based on real user data.

Future researchers and practitioners can adopt this framework as a practical template for iterative UX evaluation. By systematically testing, measuring, and optimizing user interfaces,theycancreatedigitalexperiencesthatare not only efficient and functional but also engaging and empowering—particularlyin critical fieldssuch ashealthcare, where user experience has direct implications on well-being and decision-making.

VI. DesignRecommendationsandImplications

Basedonthefindings,practicalUXrecommendations are proposed to enhance the effectiveness of AI-powered healthcare apps:

ClearandConsistentNavigation: Implementlogical flow and consistent navigation elements to reduce user confusion.

Visual Hierarchy: Use visual cues to prioritize important information, facilitating quick comprehension.

Accessibledesign: Ensure that the application is accessibleto users with varying abilities by incorporating features suchas adjustable text sizes and high contrast color schemes.

VII. EthicalConsiderations

All participants provided their informed consent. No identifiable data was collected. The study complied with the institutional ethical guidelines for user research.

VIII. Limitations

While this study yielded valuable insights into the impactof UI design on user experience in healthcare applications, several limitations must be acknowledged:

Participant Diversity: Although the study included arelatively large sample (n = 100), the participant pool wasstill somewhat limited in terms of age and cultural diversity. Most participants were tech-savvy individuals within the 18–35agegroup,whichmaynotreflectthebehaviorof older adults or individuals less familiar with mobile health technologies.

Contextual Constraints: The usability testing was conductedin a controlled environment with consistent devices, lighting, andinternetconditions.Whilethishelpedreduceconfounding variables, it may not fully represent the variability of real- world usage scenarios.

Short-term evaluation: The study focused on short-term interactions and immediate usability outcomes. It did not capture long-term behavioral patterns such as user retention, repeated engagement, or changes in trust and confidence over time. Longitudinal studies would be needed to evaluate the sustained impact of UI improvements.

IX. CONCLUSION

This study presented a structured and generalizable frame- workforevaluatinguserexperience(UX)inAI-drivenhealth- care applications through A/B testing. By systematically com- paring two interface variants of a dermatological diagnosis app, we demonstrated that thoughtful UI design improve- ments—suchasoptimizednavigation,clearvisuals,andbetter call-to-action placements can lead to statistically significant gains in task performance, user satisfaction, and task success rates.
The results confirm that even small design adjustments can makeameasurabledifferenceinhoweffectivelyusersinteract

with digital healthcare tools. UI Variant B, developed in line withUXbestpractices,consistentlyoutperformedthebaseline design. These findings validate the proposed methodology as both effective and replicable for researchers and developers aimingto improve user-centered outcomes in similar contexts.

Beyond this specific case study, the framework's core components—A/B testing, heatmap-based behavior analysis, andstatisticalevaluation—canbeappliedacrossawide range of digital health interfaces. This includes applicationsin remote patient monitoring, mental health support, fitness tracking, and chronic disease management. As healthcare increasingly transitions toward digital-first solutions, ensuring a seamless and supportive user experience is no longer optional; it is a fundamental requirement for adoption, engagement, and trust.

In conclusion, the proposed UX evaluation framework offers a practical, scalable, and data-driven approach for refining digital healthcare systems. As the demand for accessible, AI-enabled health tools continues to grow, such methodologies will be essential in creating products that are notonly intelligent but alsoinclusive, intuitive, and impactful.

X. FUTUREWORK

Buildingontheoutcomesofthisstudy,futureworkcan focus on expanding the demographic diversity of participants to include older adults, individuals with limited digital literacy, and users of different cultural backgrounds. Furthermore, conducting longitudinal studies will help to assess the long-term effects of UI improvements on user retention, trust, and behavioral outcomes.

Further research can also explore integrating AI-driven personalization into the interface - adapting layouts and workflows based on real-time user behavior - to enhance individual engagement. Finally, applying this framework to otherhealthcaredomains,suchasmentalhealth,chronic care, or teleconsultation platforms, can help validate its generalizability and drive broader improvements in UX for digital health.

REFERENCES

[1] S. Sundaresan and M. Skoglund, ”A/B testing: A systematic literaturereview,” *J. Syst. Softw.*, vol. 204, p. 111619, 2024. [Online]. Available:https://doi.org/10.1016/j.jss.2024.111619

[2] P. Kumar and S. Gupta, ”Enhancing user engagement through adaptiveUI/UX design: A study on personalized mobile app interfaces,” *Int. J.Hum.-Comput. Interact.*, vol. 39, no. 7, pp. 1123-1140, 2023. [Online]. Available:https://doi.org/10.1080/10447318.2023.2195748

[3] J.NielsenandK.Brown,”Improvinglibraryuserexperiencewith A/Btesting:Principlesandpractices,” *Weave:J.Libr. User Exp.*, vol. 1, no. 1, 2022. [Online]. Available:https://quod.lib.umich.edu/w/weave/12535642.0001.101

[4] L.GrecoandT.Z.Zlateva,”Evaluatingmobile appswithA/BandquasiA/B tests,” in *Proc. 18th Int. Conf. Mobile Softw. Eng. Syst.*, 2016, pp.54-61. [Online]. Available: https://doi.org/10.1145/2998572.2998583

[5] D. Martinez and L. Chen, ”Usability testing of mobile applications: Amethodological framework,” *Appl. Sci.*, vol. 14, no. 5, p. 1792, 2024.[Online]. Available: https://doi.org/10.3390/app14051792