

AVALIAÇÃO DA USABILIDADE DO INSTRUMENTO SYSTEM USABILITY SCALE (SUS) EM APLICATIVOS DE SAÚDE: REVISÃO INTEGRATIVA DE LITERATURA

Tamires Santos de Melo¹ – ORCID ID - <https://orcid.org/0000-0003-3928-885X>

Marcela Melo dos Santos¹ – ORCID ID - <https://orcid.org/0000-0002-9805-3260>

Clarissa de Almeida Brandão Simão¹ – ORCID ID - <https://orcid.org/0000-0003-0843-8771>

Daniel de Moraes Telles¹ – ORCID ID - <https://orcid.org/0000-0001-9576-4342>

Guaracilei Maciel Vidigal Junior¹ – ORCID ID - <https://orcid.org/0000-0002-4514-6906>

¹ Departamento de Odontologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ.

tamiressantosdemelo@gmail.com

RESUMO

OBJETIVO: A falta de avaliação dos aplicativos leva à má usabilidade, culminando em estudos-piloto malsucedidos. Este estudo tem por objetivo a identificação dos métodos usados na avaliação da usabilidade de aplicativos em saúde. **MÉTODOS:** A estratégia de busca foi realizada no banco de dados PubMed/Medline utilizando descritores MeSH. Foram incluídos os estudos que avaliaram o uso de aplicativos em dispositivos móveis na área de saúde, e excluídos: (1) estudos em animais; (2) estudos que avaliaram tele-educação ou educação mediada por tecnologia em saúde; (3) estudos que não utilizaram a escala de usabilidade System Usability Scale (SUS). Na fase de seleção de títulos, esta revisão identificou 5 publicações, das quais 1 atendeu aos critérios de inclusão / exclusão. A qualidade dos estudos foi avaliada usando a ferramenta AMSTAR 2. **RESULTADOS:** 1 artigo preencheu os critérios de inclusão / exclusão. Após a avaliação constatou-se que a qualidade do estudo era “criticamente baixa”. **CONCLUSÃO:** Tomando como base a classificação geral da AMSTAR 2, concluiu-se que a revisão não deve ser considerada como um resumo preciso e abrangente dos estudos disponíveis. Neste momento, a literatura científica carece de estudos de qualidade que avaliem os instrumentos de análise da usabilidade.

PALAVRAS-CHAVE

Escala de Usabilidade do Sistema. Revisão Sistemática. Telemedicina.

ABSTRACT

OBJECTIVE: The lack of app evaluation leads to poor usability, culminating in unsuccessful pilot studies. This study aims to identify methods to evaluate the usability of health applications. **METHODS:** A search strategy was performed in the PubMed/Medline database using MeSH

descriptors. Studies that evaluated the use of mobile applications in healthcare were included, and also were excluded: (1) animal studies; (2) studies that evaluate eLearning or education mediated by health technology; (3) studies that did not use the System Usability Scale (SUS). In the title selection step, 5 publications were identified, of which 1 met the inclusion and exclusion criteria. Study quality was assessed using the AMSTAR 2 tool. RESULTS: 1 article met the inclusion and exclusion criteria. After applying the AMSTAR 2 to assess the quality of evidence, the included study was found to be "critically low". CONCLUSION: Based on the overall AMSTAR 2 rating, it was concluded that the review should not be considered as an accurate and comprehensive summary of available studies. At this moment, the scientific literature lacks quality studies that evaluate the SUS instrument in the usability analysis.

KEY-WORDS

System Usability Scale. Systematic Review. Telemedicine.

INTRODUÇÃO

Nos últimos anos foi observado um crescimento exponencial de aplicativos destinados a área de saúde, culminando no desenvolvimento de uma prática suportada por dispositivos móveis. A mHealth, também conhecida como eHealth ou Saúde Móvel, apresenta grande potencial na prevenção, assistência e educação continuada em saúde¹.

O desenvolvimento do segmento de eHealth foi previsto pela Organização Mundial de Saúde (OMS), na 71^a Assembleia Mundial de Saúde, em Genova². Um estudo de Allied Market Research (2021) explicitou que o mercado global de eHealth deve atingir US\$ 230,6 milhões até 2027 e a expectativa de crescimento é de 14,5% entre 2020 a 2027³.

Desenvolver soluções digitais que contemplem questões de saúde é uma tarefa complexa. Quando implementadas, muitas aplicações ficam aquém das expectativas ou acabam como estudos-piloto malsucedidos. Sendo um dos principais motivos é a falta de avaliação dos aplicativos⁴.

Uma das formas de avaliar um aplicativo é por meio de teste de usabilidade. De acordo com a Organização Internacional de Padronização (ISO), usabilidade é "a medida em que um sistema, produto ou serviço pode ser usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso"⁵.

Portanto, a avaliação de usabilidade é uma etapa fundamental no desenvolvimento de aplicações digitais, considerada essencial por diversas organizações, como a OMS^{2,6}. Sua importância é reconhecida para que as tecnologias em saúde sejam devidamente planejadas e voltadas para usuários finais, antes de serem utilizadas como possíveis ferramentas tecnológicas⁷. Desta forma, o objetivo desta revisão integrativa de literatura foi identificar e avaliar (usando a ferramenta Measuremet Tool to Assess Systematic Reviews - AMSTAR 2) estudos que abordem sobre os métodos usados na avaliação da usabilidade de aplicativos em saúde.

MÉTODOS

Foi realizada uma série de pesquisas nos seguintes bancos de dados eletrônicos: PubMed e Portal de Periódicos CAPES. As pesquisas foram feitas usando as seguintes palavras-chave: System Usability Scale; Systematic Review; Telemedicine associados ao seguinte operador booleano: and. Foram usados os seguintes critérios de inclusão/exclusão:

- Critério de inclusão: (1) Estudos que avaliem o uso de aplicativos em dispositivos móveis na área de saúde.
- Critérios de exclusão: (1) Estudos em animais; (2) Estudos que avaliam tele-educação ou educação mediada por tecnológica em saúde; (3) Estudos que não utilizaram o instrumento SUS.

No dia 29/05/2023 foi realizada uma busca no PubMed com as seguintes associações de palavras-chave e operador booleano (System Usability Scale) AND ("Systematic Review" [Publication Type]) AND (Telemedicine), e foram encontrados 5 títulos. Dos 5 títulos encontrados, foram excluídos 4 títulos, sendo: 4 títulos por não utilizarem o instrumento SUS (System Usability Scale) Usa como forma de avaliação da usabilidade. Na segunda e terceira fase, o 1 título selecionado se manteve até o final.

Após isso, foi selecionado instrumento AMSTAR 2 para a avaliação da qualidade da evidência do estudo selecionado, sendo ela segmentada em 4 domínios que separam 16 itens.

RESULTADOS

O fluxograma mostra o resultado do processo de seleção dos estudos.

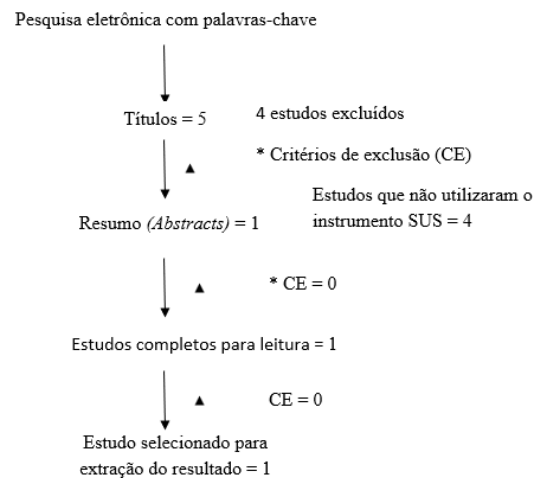


Figura 1- Fluxograma do processo de seleção dos estudos.

DISCUSSÃO / CONCLUSÃO

Esta revisão integrativa da literatura avaliou o estudo selecionado usando a partir da ferramenta AMSTAR 2 foi útil para mensurar a qualidade da evidência. Importante pontuar que, os itens 11, 12, 14 e 15 não foram discutidos pela falta da metanálise sobre revisão sistemática. Com este instrumento é possível avaliar revisões sistemáticas que incluam estudos randomizados e não randomizados em saúde.⁸

O primeiro domínio, denominado critérios de elegibilidade do estudo, contém os 3 primeiros itens. (1) "As questões de pesquisa e os critérios de inclusão para a revisão incluíram os componentes do PICO?". Este item deve ser marcado como "sim" quando os 4 elementos da estratégia PICO são descritos em algum lugar do relatório em análise. Os elaboradores da ferramenta AMSTAR 2 entendem que o uso desta estratégia é uma prática comum em revisões sistemáticas sendo, também, uma estrutura de

organização para uma questão de estudo. O estudo avaliado neste trabalho foi marcado “não”, pois não ficou explícito o uso da estratégia PICO.⁸

(2) “O relatório da revisão contém uma declaração explícita de que os métodos de revisão foram estabelecidos antes da realização da revisão e o relatório justifica quaisquer desvios significativos do protocolo?” Neste item os elaboradores da ferramenta afirmaram que a adoção de um protocolo detalhado reduz as chances de vieses na revisão, por esse motivo a ferramenta avalia sua elaboração e comparação entre o relatório e o protocolo do mesmo trabalho.⁸ Foi marcado “não” neste item porque nem todos os subitens foram contemplados para alcançar uma classificação “parcialmente sim”.

(3) “Os autores da revisão explicaram sua seleção dos desenhos de estudo para inclusão na revisão?”. Nele os elaboradores da AMSTAR 2 afirmam que os autores da revisão devem indicar qual estratégia seguiram para selecionar os desenhos de estudos. Especialmente quando há a inclusão de estudos que não sejam ensaio clínico randomizado.⁸ Há uma tabela demonstrando os estudos excluídos e seus motivos, contudo a explanação não considera o desenho de estudos, ou seja, não comenta sobre a seleção exclusiva de estudos clínicos randomizados e não randomizados, por esse motivo a classificação para este item foi “não”

O segundo domínio, denominado de “identificação e seleção dos estudos”, é composto pelos itens (4) “Os autores da revisão usaram uma estratégia abrangente de pesquisa de literatura?”, deseja-se saber se os autores da revisão realizaram uma busca abrangente da literatura, tendo pelo menos dois bancos de dados bibliográficos

incluídos na busca, e (5) “Os autores da revisão realizaram a seleção do estudo em duplicata?”, onde é importante sinalizar se a revisão foi realizada por dois autores durante o processo de determinação da elegibilidade dos estudos.⁸ Para esses itens foi marcado “Parcialmente sim” e “Sim”, respectivamente, pois os autores da revisão sistemática explicam bem sua estratégia de busca e comentam sobre a participação de mais de dois autores para dirimir conflitos.

O terceiro domínio, denominado coleta de dados e avaliação do estudo, é composto pelos itens: (6) “Os autores da revisão realizaram a extração de dados em duplicata?”. Assim como o 5, este item avalia se o processo de extração dos dados foi conduzido com pelo menos 2 avaliadores independentes e se um método de consenso foi usado quando surgiram divergências.⁸ Nesse estudo o item (6) foi marcado com “Não” por não ficar clara a informação solicitada pelo item.

(7) “Os autores da revisão forneceram uma lista de estudos excluídos e justificaram as exclusões?” exige que os autores apresentem uma lista completa dos estudos potencialmente relevantes com as justificativas para cada exclusão realizada.⁸ O item 7 foi marcado com “Parcialmente sim”, pois foram informadas justificativas de exclusão, contudo, não de cada estudo excluído de forma detalhada.

(8) “Os autores da revisão descreveram os estudos incluídos em detalhes adequados?” os autores devem detalhar apropriadamente os estudos incluídos de forma a permitir que um avaliador faça julgamentos sobre a escolha adequada do estudo em relação a estrutura PICO.⁸ A classificação dada para este item foi “Parcialmente sim” pois não houve

detalhamento suficiente para atingir a classificação superior.

(9) "Os autores da revisão usaram uma técnica satisfatória para avaliar o risco de viés (RoB) em estudos individuais que foram incluídos na revisão?". Este é um item muito importante na avaliação da qualidade dos estudos, especialmente para as revisões que incluem estudos não randomizados. Aqui os autores levam em consideração o risco de viés apresentado pelos estudos incluídos que é gerado pela aplicação de ferramentas como a ROB 2.0 da Cochrane.⁸ No estudo avaliado não houve a utilização da RoB 2.0 para análise de viés, apesar da utilização de uma ferramenta para avaliar a qualidade das evidências, por esse motivo foi marcado com "Não".

(10) "Os autores da revisão relataram as fontes de financiamento para os estudos incluídos na revisão?", sugere que os autores forneçam informações sobre fontes de financiamento para cada estudo incluído na revisão.⁸ Esta informação foi fornecida pelos autores e por isso classificou-se o item (10) com "Sim".

O quarto domínio, denominado síntese e achados, contém os 6 itens dos quais apenas 2 foram tratados nessa seção, sendo eles:

O item (13) "Os autores da revisão consideraram RoB 2.0 em estudos individuais ao interpretar/discutir os resultados da revisão?" avalia se os autores da revisão incluíram na discussão os impactos do risco de viés na interpretação dos resultados, mesmo que a metanálise não tenha sido feita.⁸ Para este estudo, a RoB 2.0 não foi conduzida na revisão e por isso marcou-se "Não".

E por fim, o (16) "Os autores da revisão relataram quaisquer fontes potenciais de conflito de interesse, incluindo qualquer financiamento que receberam para conduzir a revisão?" pergunta se os autores declaram qualquer fonte potencial de conflito de interesse no relatório. Este domínio recebeu "sim" pois os autores informaram não haver conflito de interesse.⁸

Tomando como base a classificação geral da AMSTAR 2, que foi "criticamente baixo", conclui-se que a revisão tem mais de uma falha crítica e não deve ser considerada como um resumo preciso e abrangente dos estudos disponíveis

REFERÊNCIAS

1. Carlos DAO, Magalhães TO, Vasconcelos Filho JE, Silva RM, Brasil CCP. Concepção e avaliação de tecnologia mHealth para promoção da saúde vocal. Revista Ibérica de Sistemas e Tecnologias de Informação. 2016;19(1):46
2. World Health Organization. Seventy-first World Health Assembly: Geneva, 21-26 May 2018: summary records of committees, reports of committees. 2018.
3. Allied market research. E-Health Market Size & Share | Global ehealth Market Opportunities and forecasts 2020-2027 [Internet]. 2021.
4. Schlieff M, Saunders KR, Appleton R, Barnett P, San Juan, NV, Foye U, et al. Synthesis of the Evidence on What Works for Whom in Telemental Health: Rapid Realist Review. Interact J Med Res. 2022;11(2):e38239.
5. ISO 9241-11. ISO 9241-11:2018 - Ergonomics of human-system

- interaction — Part 11: Usability: Definitions and concepts.
6. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability Evaluation of MHealth Apps for Elderly Individuals: A Scoping Review. *BMC Med Inform Decis Mak.* 2022;22(1):317.
7. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. *Int J Med Inform.* 2019;126:95–104. doi:10.1016/j.ijmedinf.2019.03.018
8. SHEA, BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *Brit Med J.* 2017;358.

ANEXOS

Tabela 1- Tabela descritiva do estudo selecionado

| Autor/Ano | Desenho do estudo | Resultados | Comentários |
|--------------------|---------------------|--|--|
| Hasnan et al. 2022 | Revisão Sistemática | <p>-Artigos incluídos: 14</p> <p>-Barreiras mais prevalentes: falta de autoeficácia, conhecimento, suporte, funcionalidade e fornecimento de informações sobre os benefícios da e-saúde para adultos mais velhos.</p> <p>-Facilitadores: engajamento ativo dos usuários finais, apoio para superar as preocupações com a privacidade e aumentar a auto-eficácia no uso da tecnologia e integração de programas de e-saúde nos serviços de saúde.</p> | Os autores apontam como limitações: a exclusão de estudos que não estivesse escritos na língua inglesa; o protocolo não foi registrado; uso de termos de pesquisa amplos (levando a possível perda de estudos focados em áreas de assunto mais específicas). |

Tabela 2 – Avaliação da qualidade da evidência, segundo a ferramenta AMSTAR 2

| Autor/Ano | Domínio | AMSTAR 2 - Domínios | Qualidade da evidência |
|-----------|---------|--------------------------------|------------------------|
| | 1 - | 1 - Critérios de inclusão PICO | Não |
| | | 2 - Presença do protocolo | Não |
| | | 3 - Seleção desenhos do estudo | Não |

| | | | |
|------------------------|---|---|------------------------------|
| Hasnan et al., 2022 | Critérios de elegibilidade e do estudo | | |
| | 2 - Identificação e seleção dos estudos | 4 - Estratégia pesquisa de literatura | Parcialmente sim |
| | | 5 – Seleção dos estudos por dois revisores independentes | Sim |
| | 3 - Coleta de dados e avaliação do estudo | 6 - Extração dos resultados por dois revisores independentes | Não |
| | | 7 – Lista dos estudos incluídos e as razões | Parcialmente sim |
| | | 8 – Descrição detalhada dos estudos incluídos | Parcialmente sim |
| | | 9 – Uso de técnica satisfatória para avaliação do risco de vieses | Não |
| | | 10 – Relato das fontes de financiamento dos estudos incluído | Sim |
| | 4 - Síntese e achados | 11 – Métodos estatísticos adequados para a combinação dos resultados | Nenhuma metanálise realizada |
| | | 12 – Se a metanálise foi realizada, avaliação do impacto do risco de vieses sobre o estudo individual | Nenhuma metanálise realizada |
| | | 13 – Houve consideração sobre o risco de vieses na interpretação/discussão dos resultados da revisão? | Não |
| | | 14 – Houve explicação satisfatória e uma discussão sobre heterogeneidade observada nos resultados? | Sim |
| | | 15 – Os autores avaliaram o viés de publicação e seu impacto nos resultados da revisão | Nenhuma metanálise realizada |
| | | 16 – Houve relato de fontes potenciais de conflito de interesse (financiamento)? | Sim |