

Desenvolvimento de um Método de Avaliação para Aplicativos Móveis de Saúde – O método QRAMHe

Thais Maria Yomoto Ferauche¹, Marcia Ito¹

¹ Programa de Mestrado Profissional em Sistemas Produtivos do Centro Paula Souza,
UEPEP/CEETEPS, São Paulo (SP), Brasil

{thais.ferauche, marcia.ito}@cpspos.sp.gov.br

Abstract. *With the growing supply of health applications, it is necessary to guarantee reliable information that encourages adherence to treatment recommended by health professionals and minimizes the risks associated with inaccurate or inadequate information. However, the absence of structured methods to evaluate these applications makes their selection difficult. To fill this gap, a method was developed based on scientific literature, the needs of users and professionals, best practices and scientific evidence. The approach combines systematic literature review and experimental development, resulting in a questionnaire validated by experts with Fleiss' Kappa analysis confirming its satisfactory reproducibility.*

Resumo. *Com a crescente oferta de aplicativos de saúde é necessário garantir informações confiáveis, que incentivem a adesão ao tratamento recomendado por profissionais de saúde e minimizem os riscos associados a informações imprecisas ou inadequadas. No entanto, a ausência de métodos estruturados para avaliar essas aplicações dificulta sua seleção. Para suprir essa lacuna, foi desenvolvido um método baseado em literatura científica, nas necessidades de usuários e profissionais, nas melhores práticas e em evidências científicas. A abordagem combina revisão sistemática da literatura e desenvolvimento experimental, resultando em um questionário validado por especialistas com análise Kappa de Fleiss confirmando sua reproduzibilidade satisfatória.*

1. Introdução

Os aplicativos móveis de saúde oferecem acesso a informações, monitoramento de condições e apoio ao autocuidado. A ausência de critérios claros de avaliação dificulta a seleção de aplicativos confiáveis, afetando cidadãos, pacientes, cuidadores e profissionais de saúde. A falta de métodos de avaliação específicos e a análise insuficiente de aspectos como acessibilidade e segurança de dados representam desafios adicionais(CHOI et al., 2020) (AGARWAL et al., 2021).

A Avaliação de Tecnologia em Saúde (Health Technology Assessment – HTA) é um processo sistemático que avalia a eficácia, segurança e custo das tecnologias de saúde, auxiliando em decisões políticas e melhorando desfechos clínicos. Esse processo aplica-se a diversos contextos, de medicamentos a dispositivos médicos, e diversos métodos foram desenvolvidos para garantir a qualidade dos aplicativos de saúde(WIETHOFF et al., 2023) (GARFIELD et al., 2023).

Dos métodos existentes oito deles são os mais utilizados e conhecidos e que são específicos para a área da saúde: Escala de Classificação de Aplicativos para Dispositivos Móveis (MARS), Escala de Classificação de Aplicativos para Dispositivos Móveis Adaptado (A-MARS), Escala de Classificação de Aplicativos para Dispositivos Móveis – Versão do usuário (uMARS), Escala de Usabilidade do Sistema (SUS), Questionário de Usabilidade de Aplicativos para saúde (MAUQ), Questionário de Usabilidade de Aplicativos de Saúde (PSSUQ), Relatório e Avaliação de Evidências de Saúde Móvel (mERA - OMS) e ISO/TS 82304-2 (STOYANOV et al., 2015).

Apesar da variedade de métodos, nenhum cobre todos os aspectos necessários. Um método abrangente pode proporcionar uma avaliação mais completa dessas ferramentas (STOYANOV et al., 2015).

Desta forma, foi desenvolvido um método de avaliação de aplicativos de saúde para garantir a segurança do paciente. O método avalia a confiabilidade das informações, além de outros critérios, como acessibilidade, adequação, usabilidade, segurança de dados, utilidade, interoperabilidade e infraestrutura. Oferece diretrizes para selecionar aplicativos seguros e confiáveis na área da saúde, integráveis a outros sistemas de informação a fim de melhorar o cuidado dos pacientes.

A principal contribuição deste estudo é o desenvolvimento e validação de um método de avaliação para aplicativos móveis de saúde, focado na confiabilidade das informações. O método permite identificar aplicativos confiáveis, contribuindo para uma avaliação mais segura e eficaz, alinhada aos desafios dos sistemas de informação na saúde.

2. Métodos

A pesquisa seguiu uma abordagem quali-quantitativa com uma perspectiva interpretativista, enfatizando a compreensão das experiências dos participantes sobre o protocolo de avaliação de aplicativos de saúde. O desenvolvimento desse protocolo foi baseado na revisão sistemática da literatura de Ferauche, Silva e Ito (2024), que identificou métodos existentes e suas características, e fundamentando-se na ISO/IEC 25000 ([s.d.]).

O desenvolvimento do protocolo, baseado na ISO/IEC 25000, seguiu etapas estruturadas, incluindo a avaliação da qualidade. Essas etapas envolveram: (a) definição dos objetos e contextualização da avaliação; (b) configuração do modelo de qualidade com critérios fundamentados na literatura; (c) identificação das medidas de avaliação; (d) definição dos níveis de pontuação e critérios de julgamento; (e) elaboração da avaliação; e (f) execução, abrangendo coleta, comparação e análise dos dados (ROBERTS et al., 2021).

O plano de validação do protocolo inclui diretrizes, procedimentos e o método de coleta de dados, e entrevista como técnica de validação. Essa abordagem, em pesquisas qualitativas, visa garantir credibilidade. Na etapa final, o avaliador deve: (i) reunir as medidas de qualidade, (ii) compará-las com os critérios pré-estabelecidos e (iii) analisar os dados coletados (STOYANOV et al., 2016).

A validação do protocolo envolveu análises qualitativas para ajustes e quantitativas para avaliar sua reprodutibilidade. Foram utilizadas a análise qualitativa de conteúdo para identificar padrões e a estatística Kappa de Fleiss para medir a reprodutibilidade do instrumento.

3. Resultados e Discussão

Um mapeamento sistemático, conforme Ferauche, Silva e Ito (2024), identificou características relevantes em aplicativos de saúde, formando a estrutura inicial para o método de avaliação. Pesquisas adicionais, como as conduzidas por Stoyanov et al. (2015) (2016) e Roberts et al. (2021) orientaram os estágios do processo de desenvolvimento do método de avaliação. O protocolo foi baseado na ISO/IEC 25000 ([s.d.]), unificando as normas ISO/IEC 9126 ([s.d.]) e ISO/IEC 14598 ([s.d.]) para garantir a qualidade do software, envolvendo especificação de requisitos e avaliação da qualidade.

O método QRAMHe foi desenvolvido para analisar e comparar aplicativos de saúde existentes, avaliando sua qualidade e confiabilidade por profissionais da área. Os requisitos de qualidade do software são convertidos em características mensuráveis, estruturadas em um questionário com questões objetivas e respostas nominais. O método está disponível publicamente via a ferramenta computacional QRAMHe (<https://qramhe.org/>).

O QRAMHe tem como principal componente a “Definição dos Objetos e Contextualização”, baseada em um modelo de qualidade e confiabilidade para avaliar aplicativos de saúde. Esse modelo é estruturado em quatro grandes áreas: Fornecedor, que analisa informações do aplicativo sem necessidade de contato com o desenvolvedor; Adoção da Tecnologia, que avalia usabilidade, segurança e privacidade; Saúde, dividido em adequação, integração, utilidade e confiabilidade das informações; e Qualidade Subjetiva, que considera a percepção individual de qualidade com base em experiências e preferências pessoais. Essa estrutura hierárquica foi fundamentada na revisão da literatura. Em sequência é apresentado cada uma das categorias originadas pelas quatro áreas.

As categorias do QRAMHe estruturam a avaliação dos aplicativos de saúde a partir de diferentes aspectos. A categoria Informações Iniciais analisa a transparência e acessibilidade das informações essenciais, como contatos e políticas de privacidade. Em Adequação e Restrição verifica se o aplicativo atende às necessidades do usuário, considerando limitações etárias, diagnósticas e de uso. A Usabilidade foca na interface, garantindo que seja intuitiva e eficiente para profissionais da saúde. No aspecto de Segurança e Privacidade, são avaliadas medidas de proteção de dados e conformidade com regulamentos como LGPD, GDPR e HIPAA.

A categoria Utilidade e Funcionalidade examina se o aplicativo cumpre sua proposta, considerando funcionalidades, possíveis restrições de uso e custo-benefício. Em Integração com Sistemas de Saúde e Profissionais, avalia-se a capacidade do aplicativo de se conectar a sistemas profissionais, facilitando a troca segura de informações. A Confiabilidade das Informações assegura que os conteúdos oferecidos são baseados em evidências científicas e diretrizes médicas.

Além disso, a Infraestrutura analisa aspectos essenciais para o funcionamento adequado do aplicativo, como segurança, backup, sincronização de dados e compatibilidade com diferentes dispositivos. A Acessibilidade busca garantir que o aplicativo seja utilizável por todas as pessoas, promovendo inclusão e igualdade no acesso aos serviços de saúde digital. Por fim, a Qualidade Subjetiva refere-se à percepção individual do usuário sobre o aplicativo, considerando fatores como usabilidade, design e o quanto ele atende às suas expectativas.

O componente "Definição e Configuração do Modelo de Qualidade" possibilita uma avaliação mais precisa do desempenho dos aplicativos de saúde, mapeando categorias em atributos mensuráveis conforme a ISO/IEC 25000 ([s.d.]). Para isso, foi desenvolvido o questionário QRAMHe, baseado em métodos de avaliação existentes e adaptado ao contexto da saúde. Com 50 questões distribuídas nas 10 categorias do modelo, o questionário utiliza a Escala de Likert de 5 pontos, permitindo que os respondentes expressem seu nível de concordância ou discordância em relação às afirmações apresentadas (BARUA, 2013).

O componente "Identificação das Medidas da Avaliação do Protocolo" define critérios para determinar a qualidade e confiabilidade dos aplicativos, convertendo as respostas do questionário em dados quantitativos. As perguntas são pontuadas em uma escala de 1 a 5, onde 1 indica um nível muito baixo e 5 um nível muito elevado. Cada questão foi estruturada com opções de resposta que permitem uma análise clara e objetiva das avaliações.

O componente "Definição de Níveis de Pontuação e Critérios de Julgamento" estabelece critérios para avaliar os aspectos do protocolo, com a pontuação final sendo calculada pela média aritmética de cada categoria. Os valores de 1 a 5 classificam as categorias como ineficiente (até 1), deficiente (1 a 2), regular (2 a 3), aceitável (3 a 4) e eficiente (4 a 5). O nível de qualidade geral do aplicativo é determinado pela média das categorias, sendo classificado como 'Superior' ($\geq 80\%$), 'Médio' (50% a 80%) ou 'Baixo' ($< 50\%$).

Um aplicativo com nível superior indica excelente qualidade, sendo recomendado para uso sem restrições. Um nível médio sugere que o aplicativo pode ser usado, mas com restrições. Já um nível baixo indica que o aplicativo não atinge padrões satisfatórios de qualidade e não deve ser recomendado ou adotado.

Para a avaliação inicial do protocolo do método, optou-se pela entrevista com especialistas da área, antes de avaliá-la com uma experimentação de campo. A validação por entrevista é uma técnica usada em pesquisas qualitativas para obter credibilidade (BARUA, 2013). Esta pesquisa foi aprovada pelo Comissão de Ética do Programa de Mestrado Profissional em Gestão e Tecnologia em Sistemas Produtivos.

O planejamento da entrevista, adaptado de Mann (2016), incluiu etapas como definição do objetivo, seleção de participantes, preparação, realização e pós-entrevista. O objetivo foi validar o protocolo QRAMHe com três participantes doutores especializados em aplicativos móveis, tecnologia e saúde. As entrevistas, realizadas à distância e com duração de 30 a 60 minutos, foram guiadas por um questionário de dez perguntas.

Os entrevistados opinaram sobre a clareza, abrangência, sequência lógica, terminologia e tempo de execução do questionário. Com base nas respostas dos entrevistados, foram sugeridas melhorias específicas para o questionário, como substituir a escala Likert por *checkboxes* em algumas questões, ajustar terminologia, adicionar perguntas específicas, informar sobre o objetivo de cada categoria, tornar as perguntas mais sintéticas e por vezes dividí-las, enfocar a confiabilidade dos dados e considerar novos aplicativos.

Para avaliar a reprodutibilidade do protocolo QRAMHe, foi elaborado um questionário com dez itens abordando fatores como entendimento das questões, relevância, abrangência, organização e tempo de resposta. A estatística Kappa de Fleiss

foi utilizada, com o auxílio do software IBM SPSS Statistics (versão 28.0), para avaliar a confiabilidade das respostas. O Kappa de Fleiss mostrou uma confiabilidade geral razoável entre os avaliadores ($k = 0,298$ [IC 95%: 0,011-0,586]; $z = 2,032$; $p < 0,043$). Para a resposta "SIM", a confiabilidade foi moderada ($k = 0,441$ [IC 95%: 0,083-0,799]; $z = 2,415$; $p < 0,017$). A análise revelou uma reproduzibilidade significativa ($p < 0,017$), indicando que o instrumento é satisfatoriamente reproduzível.

4. Conclusão

A pesquisa desenvolvida teve como ponto de partida a ausência de métodos completos para avaliar aplicativos móveis voltados à área da saúde. Diante disso, propôs-se o desenvolvimento de um protocolo de avaliação que contemplasse dimensões como acessibilidade, usabilidade, segurança, privacidade, utilidade, interoperabilidade, adequação, confiabilidade das informações, funcionalidade e infraestrutura tecnológica.

O principal resultado alcançado com este estudo foi a elaboração de um método de avaliação aberto, centrado especialmente na verificação da confiabilidade das informações médicas fornecidas por aplicativos. Este método está acessível por meio de uma plataforma digital (<https://qramhe.org/>), e tem como finalidade apoiar profissionais da saúde na tarefa de selecionar soluções tecnológicas seguras e adequadas ao uso por pacientes e cuidadores.

Durante o processo de validação, foi realizada a análise das respostas de especialistas, a qual evidenciou pontos de melhoria no instrumento inicial. Os participantes destacaram a importância de tornar o questionário mais claro, direto e relevante, especialmente diante da diversidade de públicos-alvo que interagem com esses aplicativos. As recomendações recebidas reforçaram a necessidade de desenvolver um instrumento mais acessível e efetivo, capaz de avaliar criticamente os aplicativos sob múltiplas perspectivas.

O modelo proposto busca facilitar a identificação de aplicativos que sejam não apenas funcionais, mas também confiáveis do ponto de vista clínico e informacional. Dessa forma, espera-se que sua aplicação contribua para experiências mais seguras e positivas, tanto para os usuários quanto para os profissionais que indicam tais tecnologias.

Ainda como parte do processo de refinamento do protocolo, destaca-se a necessidade de ajustes no questionário para conclusão da etapa de validação. Está previsto, como trabalho futuro, um experimento de campo com a aplicação do método em situações reais, permitindo verificar sua eficácia prática na avaliação de aplicativos de saúde.

Em síntese, esta pesquisa oferece uma contribuição relevante ao propor um instrumento que responde a uma falta de instrumentos que avaliem aplicativos em saúde digital que não são considerados dispositivos médicos no campo da saúde digital. Ao apoiar decisões baseadas em critérios técnicos e científicos, o método tem o potencial de reduzir o uso de aplicativos inadequados e mitigar riscos associados à disseminação de informações médicas imprecisas.

References

AGARWAL, P. et al. Assessing the quality of mobile applications in chronic disease management: a scoping review. **NPJ digital medicine**, v. 4, n. 1, p. 46, 2021.

BARUA, A. METHODS FOR DECISION-MAKING IN SURVEY QUESTIONNAIRES BASED ON LIKERT SCALE. **Journal of Asian Scientific Research**, 2013.

CHOI, W. et al. A systematic review of mobile health technologies to support self-management of concurrent diabetes and hypertension. **Journal of the American Medical Informatics Association**, v. 27, n. 6, p. 939–945, 2020.

FERAUCHE, T. M. Y.; SILVA, J. J.; ITO, M. **Proposal for the Protocol for Assessment of mHealth: QRAMHe**. CONTECSI-International Conference on Information Systems and Technology Management [Internet]. **Anais...**2024.

GARFIELD, S. et al. Are Health Technology Assessments Keeping Pace with Health Equity Priorities: A Review of Existing Approaches and Discussion of Emerging Practices. **Medical Research Archives**, v. 11, n. 6, 2023.

ISO/IEC 9126-1:2001. Disponível em: <<https://www.iso.org/standard/22749.html>>. Acesso em: 3 fev. 2025.

ISO/IEC 14598-5:1998. Disponível em: <<https://www.iso.org/standard/24906.html>>. Acesso em: 3 fev. 2025.

ISO/IEC 25000:2014. Disponível em: <<https://www.iso.org/standard/64764.html>>. Acesso em: 3 fev. 2025.

MANN, S. **The Research Interview**. London: Palgrave Macmillan UK, 2016.

ROBERTS, A. E. et al. Evaluating the quality and safety of health-related apps and e-tools: Adapting the Mobile App Rating Scale and developing a quality assurance protocol. **Internet interventions**, v. 24, p. 100379, 2021.

STOYANOV, S. R. et al. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. **JMIR mHealth and uHealth**, v. 3, n. 1, p. e3422, 2015.

STOYANOV, S. R. et al. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). **JMIR mHealth and uHealth**, v. 4, n. 2, p. e5849, 2016.

WIETHOFF, I. et al. An introduction to health technology assessment and health economic evaluation: an online self-learning course. **Netherlands Heart Journal**, v. 31, n. 6, p. 219–225, jun. 2023.