

# Good evaluation to warrant better digital health

## Gute Evaluation für eine bessere digitale Versorgung

Ursula Hübner<sup>1</sup>

<sup>1</sup> Research Center for Health  
and Social Informatics,  
Hochschule Osnabrück  
University of Applied  
Sciences, Osnabrück,  
Germany

### Editorial

This special issue of MIBE comprises contributions submitted to the Special Topic Conference “Good Evaluation – Better Digital Health” of the European Federation of Medical Informatics (EFMI). EFMI Special Topic Conferences (STC) are scientific events focussing on a dedicated topic in biomedical and health informatics and are addressing particularly young and early career scientists. This year, the STC is hosted by the German Association of Medical Informatics, Biometry and Epidemiology (GMDS) and organised by Osnabrück University of Applied Sciences. Publishing contributions to a European conference illustrates MIBE’s general offer to colleagues from Europe and beyond to publish their work in this journal. As a first step, posters submitted as a three-page poster contribution to the STC were considered for publication and underwent the same rigorous peer review process as all contributions to the conference before they were proposed for publication in MIBE. Here they were peer reviewed again.

Nine out of 20 poster contributions submitted were accepted and published in this special issue covering topics of the theme “Good Evaluation – Better Digital Health”. Evaluation is a centre piece of the software and systems life cycle and starts – in contrast to usual anticipations – not at the end when the system has been rolled out but at the very beginning of the development process. It thus allows the integration of empirical patterns against which the testing can be performed such as user expectations and requirements. The articles “Co-creating a cancer screening dashboard with screening invitees and experts” by Oldhoff-Nuijsink et al. [1] and “Building a digital platform for collaborative second opinions in rare disease: Integrating AI and healthcare networks for improved care” by Lima et al. [2] illustrate the meaningfulness of early evaluations. Conversely and even a step earlier, evaluations of patients’ behaviour can stimulate the needs to

foster long-term medication adherence through digital methods as was shown by Kim et al. in their study “Adherence to hormonal therapy in breast cancer patients: EHR-based retrospective data analysis” [3].

Evaluations and tests that are derived from frameworks, theories and concepts have a different background, however, need not contradict more empirically grounded evaluation targets. It is therefore useful to identify new frameworks as in the case of Hülsmann et al. in “Identifying assessment frameworks for digital public health interventions: First results of a scoping review” [4] or to augment existing ones as was proposed by Grashof et al. in “A new perspective on eHealth acceptance: Combining health-related factors with the Technology Acceptance Model” [5].

The emergence of more and more AI models for use in healthcare increases the necessity to conduct proper evaluation studies before such models can be integrated into systems for clinical use. The study by Galland-Decker et al. “Performance evaluation of Meditron3-70B in medical coding: Current limitations and integration perspectives for clinical practice” [6] investigated the performance of a medical large language model to generate SNOMED and ICD codes from vignettes and hinted at current flaws.

Evaluating a system often means evaluating the context in which the system is embedded in. This also applies to AI systems and their rootedness in models that have been developed on a specific set of data to be applied to fresh clinical data. The study by Slob et al. “FAIVOR – a push-button system for AI validation within the hospital” [7] describes a tool that serves as an AI model repository and allows the evaluation of a given model on local data. Evaluating applications and systems is often accompanied by educational measures to really obtain better digital health which means nothing less than better use from better digital systems. Two articles address education and training: “Developing competencies in health inform-

atics: Blended teaching method “by Mannevaara et al. [8] and “AI and AR for inclusive health education“ by Focsa [9].

The studies presented in this special issue give rise to longitudinal measures to evaluate the output, outcome and impact of digital systems to finally appraise whether an application really contributed to better digital health.

## Notes

## Competing interests

The author declares that she has no competing interests.

## References

1. Oldhoff-Nuijsink C, Buitendijk F, Rolink M, Derksen ME, Peute LWP, Fransen MP. Co-creating a cancer screening dashboard with screening invitees and experts. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc20. DOI: 10.3205/mibe000292
2. Lima V, Bernardi F, Ferraz V, Alves D. Building a digital platform for collaborative second opinions in rare disease: Integrating AI and healthcare networks for improved care. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc19. DOI: 10.3205/mibe000291
3. Kim S, Park YE, Lee Y, Lee JW. Adherence to hormonal therapy in breast cancer patients: EHR-based retrospective data analysis. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc18. DOI: 10.3205/mibe000290
4. Hülsmann H, Swoboda W, Holl F. Identifying assessment frameworks for digital public health interventions: First results of a scoping review. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc17. DOI: 10.3205/mibe000289
5. Grashof R, Breil B, Lipprandt M. A new perspective on eHealth acceptance: Combining health-related factors with the Technology Acceptance Model. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc16. DOI: 10.3205/mibe000288
6. Galland-Decker C, Usenbacher M, Nunes C, Bouche F, Carra G, Boillat-Blanco N, Hartley MA, Sallinen A, Raisaro JL, Bastardot F. Performance evaluation of Meditron3-70B in medical coding: Current limitations and integration perspectives for clinical practice. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc15. DOI: 10.3205/mibe000287
7. Slob D, Akhmad E, Garcia González J, Amiri S, Kasalica V, Georgievska S, Choudhury A, Lobo Gomes A, Dekker A, van Soest J. FAIVOR – a push-button system for AI validation within the hospital. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc14. DOI: 10.3205/mibe000286
8. Mannevaara P, Saranto K, Kinnunen UM. Developing competencies in health informatics: Blended teaching method. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc13. DOI: 10.3205/mibe000285
9. Focsa M. ARISE: AI and AR for inclusive health education. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc12. DOI: 10.3205/mibe000284

### Corresponding author:

Prof. Dr. Ursula Hübner  
Hochschule Osnabrück University of Applied Sciences,  
Albrechtstr. 30, 49076 Osnabrück, Germany  
u.huebner@hs-osnabrueck.de

### Please cite as

Hübner U. Good evaluation to warrant better digital health. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc21.  
DOI: 10.3205/mibe000293, URN: urn:nbn:de:0183-mibe0002936

### This article is freely available from

<https://doi.org/10.3205/mibe000293>

**Published:** 2025-10-17

### Copyright

©2025 Hübner. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.