

INNOVATÍV ÉS DIGITÁLIS EGÉSZSÉGIPARI TECHNOLOGIÁK FEJLESZTÉSE ÉS ÉRTÉKELÉSE

DIGITÁLIS ORVOSTECHNIKAI ESZKÖZÖK (DOE) ÉRTÉKELÉSE: HATÁSOSság, BIZTONSÁGOSSÁG ÉS TÁRSADALMI HASZNOSULÁS

DOE-k klinikai vizsgálatainak tervezése és értékelése

avagy hogyan hozzunk létre klinikai bizonyítékokat a DOE-k hatásosságáról és biztonságosságáról?

Péntek Márta, Gulácsi László, Zrubka Zsombor, Hölgyesi Áron, Kovács Levente



ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY

TKP-2021-NKTA-36



NEMZETI KUTATÁSI, FEJLESZTÉSI
ÉS INNOVÁCIÓS HIVATAL

AZ NKFI ALAPBÓL
MEGVALÓSULÓ
PROJEKT

AI alapú digitális orvostechnikai eszközök



Mesterséges intelligencia (AI) alkalmazási területek

I. Beültethető orvostechnikai eszközök

II. Robotsebészet

III. Segítő robotok (tapasztalatok, otthoni ápolás, kórházi ellátás)


IV. Egészségügyi chatbotok

V. Gyermekkori cukorbetegség

VI. Fogászat

VII. Kéz higiénia

2. Klinikai vizsgálatok tervezése és értékelése



2.
Klinikai
vizsgálatok
tervezése és
értékelése

Irányelvek AI-alapú klinikai vizsgálatok tervezéséhez és közléséhez

Egészségügyben alkalmazott chatbotok randomizált kontrollált vizsgálatai

AI-alapú eszközök klinikai vizsgálatai 1-es típusú gyermekkori cukorbetegségben

Cél: az orvosi mesterséges intelligencia kutatások közlési irányelveinek szisztematikus áttekintése:

- milyen felhasználók számára,
 - milyen fókusszal,
 - milyen struktúrával,
 - milyen felhasználási célra és
 - milyen módszerekkel
- lettek az egyes irányelvek kifejlesztve.

Eredmények

2023 elejéig 24 közlési irányelvet azonosítottunk. Ezek közül

- 4 narratív irányelv
- 7 általános közlési irányelv
- 4 speciális klinikai vizsgálati módszertanhoz kapcsolódó közlési irányelv
- 9 egy-egy klinikai területhez kapcsolódó közlési irányelv

Tizenhét irányelv szerzők, szerkesztők és lektorok, hét a klinikusok és felhasználók, három pedig a felhasználók számára készült.

Kilenc irányelv a tudományos publikáció struktúráját (bevezetés, módszerek, eredmények, diszkusszió), tizenkettő a gépi tanulás munkamenetet, és három egyéb logikát követett. Az irányelvek összesen 704 ajánlást fogalmaztak meg. Az irányelv fejlesztés módszertanát tizenhárom irányelv közölte, ebből egy volt teljeskörű (CONSORT–AI), ez azonban csak randomizált vizsgálatokra fókuszált.

Az orvosi mesterséges intelligencia kutatások közlési irányelveivel kapcsolatban jelenleg nincs konszenzus.

A legjobb eredmények érdekében az egyes vizsgálati típusok orvosszakmai közlési irányelveit és egy robusztus fejlesztésen átesett általános mesterséges intelligencia irányelvet célszerű kombinálni és együtt figyelembe venni.

TKP2021-NKTA-36 publikációk:

- Zrubka Z, Kovács L, Motahari-Nezhad H, Czere J, Gulácsi L, Péntek M. Artificial Intelligence in Medicine: A Systematic Review of Guidelines for the Reporting and Interpretation of Studies. Acta Polytechnica Hungarica. 22(10):125-143. DOI: 10.12700/APH.22.10.2025.10.8 https://acta.uni-obuda.hu/Zrubka_Kovacs_MotahariNezhad_Czere_Gulacsi_Pentek_162.pdf
- Zsombor, Zrubka ; László, Gulácsi ; Márta, Péntek. Time to start using checklists for reporting artificial intelligence in health care and biomedical research: a rapid review of available tools. In: Szakál, Anikó (szerk.) 26th IEEE International Conference on Intelligent Engineering Systems, INES 2022 : Proceedings. Piscataway (NJ), Amerikai Egyesült Államok : IEEE Hungary Section (2022) 273 p. pp. 15-20. , 6 p. <https://ieeexplore.ieee.org/document/9922639>

A legtöbb AI-chatbot randomizált kontrollált vizsgálat publikációja **nem ad meg elegendő információt** ahhoz, hogy a technológia megbízhatóságát és biztonságosságát egyértelműen meg lehessen ítélni. A vizsgálattervezés és közlés gyakorlatát standardizálni és javítani kell.

AI Chatbotok az Egészségügyben: A Bizonyítékok Minőségi Problémája

Az AI-alapú chatbotok egyre elterjedtebbek az egészségügyben, különösen a mentális egészség, a közegészségügy és a daganatos megbetegedések területén. Egy 28 randomizált, kontrollált vizsgálatot (RCT) elemző kutatás feltárta, hogy ezen tanulmányok jelentési minősége nem felel meg a CONSORT-AI nemzetközi irányelveinek, ami megkérdőjelezi az eredmények megbízhatóságát és reprodukálhatóságát.

A Probléma: Kritikus Hiányosságok a Jelentésekben

0%

Egyetlen vizsgált tanulmány sem felelt meg teljesen a CONSORT-AI irányelveknek.

A megfelelés mértéke 28,6% és 78,6% között mozgott a 28 elemzett vizsgálatban.



3.6%

A tanulmányok csupán 3,6%-a részletezte a bemeneti adatokra vonatkozó kritériumokat.

Ez alapvető fontosságú az AI-eszköz megbízhatóságának és alkalmazhatóságának megítéléséhez.

7.1%

A vizsgálatok mindössze 7,1%-a elemezte a chatbot teljesítményhibáit.

A hibaelemzés hiánya megnehezíti a chatbot biztonságosságának és megbízhatóságának értékelését.

A Megoldás: Hogyan Javítható a Minőség a CONSORT-AI Szerint?



Bemeneti Adatok Meghatározása

Pontosan rögzíteni kell a bemeneti adatokra vonatkozó befogadási és kizárási kritériumokat.

Verziókövetés és Hozáférhetőség

Az AI-algoritmus verzióját és a kód elérhetőségét egyértelműen jelezni kell az átláthatóság érdekében.

Ember-AI Interakció és Hibaelemzés

Részletezni kell a felhasználói interakciót, a szükséges szakértelmet és a teljesítményhibákat.

TKP2021-NKTA-36 publikációk:

- Czere, János Tibor ; Hölgyesi, Áron ; Péntek, Márta. Improving the Usability of Randomized, Controlled Trials of Artificial Intelligence-based Chatbots in Healthcare: Results of a Systematic Literature Review. ACTA POLYTECHNICA HUNGARICA 22 : 10 pp. 145-166. , 22 p. (2025) https://acta.uni-obuda.hu/Czere_Holgyesi_Pentek_162.pdf
- Czere, János Tibor ; Péntek, Márta. Leveraging Prompt Engineering in the Stage Gate Innovation. In: Szakál, Anikó (szerk.) 19th IEEE International Symposium on Applied Computational Intelligence and Informatics, SACI 2025 : Proceedings. Danvers (MA), Amerikai Egyesült Államok : Institute of Electrical and Electronics Engineers (IEEE) (2025) 755 p. pp. 485-490. , 6 p. <https://ieeexplore.ieee.org/document/11030180>
- Johanna, Tripo ; József, Fogarasi ; Márta, Péntek. Statistical reporting quality of randomized controlled trials on artificial intelligence chatbots for promoting healthy lifestyle: exploring the SAMPL guideline. In: Szakál, Anikó (szerk.) IEEE 20th Jubilee International Symposium on Intelligent Systems and Informatics (SISY 2022) Szabadka, Szerbia : Institute of Electrical and Electronics Engineers (IEEE) (2022) 457 p. pp. 31-36. , 6 p. <https://ieeexplore.ieee.org/document/10036314>

AI-alapú eszközök klinikai vizsgálatai 1-es típusú gyermekkori cukorbetegségben



Vizsgálat

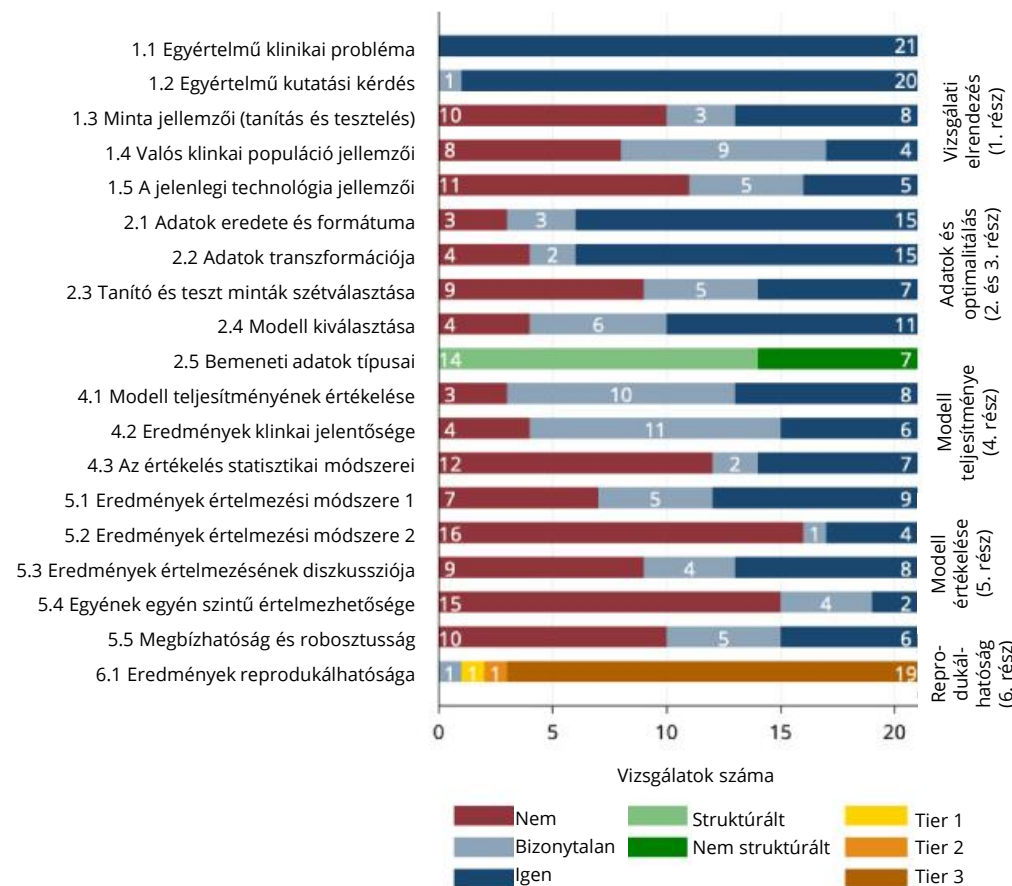
Cél: cukorbeteg gyermekeken mesterséges intelligencia alapú orvosi kutatások közlési minőségének vizsgálata a MI-CLAIM közlési ellenőrzőlista segítségével.

A vizsgálatba olyan kutatások kerültek bevonásra, melyek 2016-2020 között 2-18 éves gyermekeken alkalmaztak mesterséges intelligenciát a cukorbetegség vagy komplikációinak kezelésére, szűrésére, vagy megelőzésére.

Eredmények

28 vizsgálatot azonosítottunk, melyből 21 esetén lehetett a MI-CLAIM ellenőrzőlistát alkalmazni. A MI-CLAIM 17 eleméből átlagosan 7.4 esetén volt megfelelő a vizsgálatok közlési minősége. A legjobb minőségű közlés 12, míg a leghiányosabb csak 4 szempont esetén tartalmazott mind orvos, mind számítástudományi szakértő számára kellően pontos és részletes információt.

A gyermekkori cukorbetegségben a mesterséges intelligencia kutatások **közlési minősége általában alacsony** volt. A klinikusok számára fontos információk, mint a betegcsoportok tulajdonságai, az elérhető technológiákkal való összehasonlítás, és az eredmények értelmezése gyakran nem volt kielégítő. **Az eredmények klinikai alkalmazáshoz vizsgálatok közlési minőségén és az algoritmusok transzparenciáján, reprodukálhatóságán javítani kell.**



TKP2021-NKTA-36 publikációk:

- Zrubka, Zsombor ✉ ; Kertész, Gábor ; Gulácsi, László ; Czere, János ; Hölgyesi, Áron ; Nezhad, Hossein Motahari ; Mosavi, Amir ; Kovács, Levente ; Butte, Atul J ; Péntek, Márta. The Reporting Quality of Machine Learning Studies on Pediatric Diabetes Mellitus: Systematic Review JOURNAL OF MEDICAL INTERNET RESEARCH 26 p. e47430 , 22 p. (2024) <https://www.jmir.org/2024/1/e47430>

Főbb eredmények röviden

DOE klinikai vizsgálatok tervezését és közlését segítő irányelvek

DOE klinikai vizsgálatainak tervezésekor javasolt tájékozódni és figyelembe venni az AI vizsgálatokra vonatkozó irányelveket (reporting guidelines, checklists). Jelenleg több mint 20 ilyen irányelv érhető el.

Rendelkezésre álló irányelvek a klinikai vizsgálat és közlés minőségének (transzparens, megfelelően részletgazdag, áttekinthető) garantálását és standardizálását hivatottak elősegíteni, de nincs konszenzusos verzió. Az AI-ra vonatkozó irányelvek mellett javasolt használni a klinikai vizsgálat típusának (pl. CONSORT) megfelelő irányelveket kiegészítőleg.

Randomizált kontrollált vizsgálatok

Esettanulmányunk AI alapú chatbotok terén kimutatta, hogy a legmagasabb klinikai bizonyítékokat adó randomizált kontrollált vizsgálatok (randomized controlled trial, RCT) publikációinak módszertani megfelelősége közepes vagy gyenge.

Hasonló hiányosságokat észleltünk az 1-es típusú cukorbetegségben alkalmazott AI klinikai vizsgálatoknál is.

Ahhoz, hogy bizonyítékokon alapuló orvosi döntéshozatal történjen az AI alapú DOE-k alkalmazásával kapcsolatban, javítani kell a hatásosságra és biztonságosságukra vonatkozó klinikai vizsgálatok minőségét, és azok közlésének minőségét is.

Főbb TKP publikációk a témában

2 szakkikk, Scimago Journal Ranking (SJR): D1: 1; Q1: 1

Szakkikkek

1. Zrubka Z, Kovács L, Motahari-Nezhad H, Czere J, Gulácsi L, Péntek M. Artificial Intelligence in Medicine: A Systematic Review of Guidelines for the Reporting and Interpretation of Studies. Acta Polytechnica Hungarica. 22(10):125-143. DOI: 10.12700/APH.22.10.2025.10.8 https://acta.uni-obuda.hu/Zrubka_Kovacs_MotahariNezhad_Czere_Gulacsi_Pentek_162.pdf **SJR Q2**
2. Zrubka, Zsombor ✉ ; Kertész, Gábor ; Gulácsi, László ; Czere, János ; Hölgyesi, Áron ; Nezhad, Hossein Motahari ; Mosavi, Amir ; Kovács, Levente ; Butte, Atul J ; Péntek, Márta. The Reporting Quality of Machine Learning Studies on Pediatric Diabetes Mellitus: Systematic Review JOURNAL OF MEDICAL INTERNET RESEARCH 26 p. e47430 , 22 p. (2024) <https://www.imir.org/2024/1/e47430> **SJR D1**

Konferenciaközlemények

1. Zsombor, Zrubka ; László, Gulácsi ; Márta, Péntek. Time to start using checklists for reporting artificial intelligence in health care and biomedical research: a rapid review of available tools. In: Szakál, Anikó (szerk.) 26th IEEE International Conference on Intelligent Engineering Systems, INES 2022 : Proceedings. Piscataway (NJ), Amerikai Egyesült Államok : IEEE Hungary Section (2022) 273 p. pp. 15-20. , 6 p. <https://ieeexplore.ieee.org/document/9922639>
2. Czere, János Tibor ; Hölgyesi, Áron ; Péntek, Márta. Improving the Usability of Randomized, Controlled Trials of Artificial Intelligence-based Chatbots in Healthcare: Results of a Systematic Literature Review. ACTA POLYTECHNICA HUNGARICA 22 : 10 pp. 145-166. , 22 p. (2025) https://acta.uni-obuda.hu/Czere_Holgyesi_Pentek_162.pdf
3. Czere, János Tibor ; Péntek, Márta. Leveraging Prompt Engineering in the Stage Gate Innovation. In: Szakál, Anikó (szerk.) 19th IEEE International Symposium on Applied Computational Intelligence and Informatics, SACI 2025 : Proceedings. Danvers (MA), Amerikai Egyesült Államok : Institute of Electrical and Electronics Engineers (IEEE) (2025) 755 p. pp. 485-490. , 6 p. <https://ieeexplore.ieee.org/document/11030180>
4. Johanna, Tripo ; József, Fogarasi ; Márta, Péntek. Statistical reporting quality of randomized controlled trials on artificial intelligence chatbots for promoting healthy lifestyle: exploring the SAMPL guideline. In: Szakál, Anikó (szerk.) IEEE 20th Jubilee International Symposium on Intelligent Systems and Informatics (SISY 2022) Szabadka, Szerbia : Institute of Electrical and Electronics Engineers (IEEE) (2022) 457 p. pp. 31-36. , 6 p. <https://ieeexplore.ieee.org/document/10036314>

A kutatás a Nemzeti Kutatási, Fejlesztési és Innovációs alap Tématerületi Kiválóság Program (TKP2021-NKTA-36) támogatásával az „Innovatív és digitális egészségipari technológiák fejlesztése és értékelése - Digitális orvostechikai eszközök értékelése: hatásosság, biztonságosság és társadalmi hasznosulás” kutatási projekt keretében jött létre az Óbudai Egyetemen.

Köszönetünket fejezzük ki a kutatásokban résztvevő társszerzőknek, kollégáknak, pácienseknek és családtagjaiknak.

Levelezés: Prof. Dr. Péntek Márta, az MTA doktora
alprojekt-vezető
pentek.marta@uni-obuda.hu



NEMZETI KUTATÁSI, FEJLESZTÉSI
ÉS INNOVÁCIÓS HIVATAL

AZ NKFI ALAPBÓL
MEGVALÓSULÓ
PROJEKT