

A metrológia szerepe az AI által támogatott digitális biomarker-kutatásban és -fejlesztésben

The role of metrology in AI-enabled digital biomarker research and development

Pálhalmi János

DataSenseLabs Kft. <https://datasenselabs.net/>, Budapest
janos.palhalmi@datasenselabs.net

Összefoglalás — A tanulmány a metrológia szerepéről elemzi a digitális biomarker kutatás-fejlesztésében, különösen a mesterséges intelligencia és nagy nyelvi modellek integrációja kapcsán. Rámutat arra, hogy a digitális biomarkerek technológiai fejlődése ellenére hiányoznak az egységes szabványok és értékelési protokollok, ami hátráltatja a klinikai alkalmazhatóságot. A metrológiai szemlélet alkalmazása segítheti a validálást, a nyomkövethetőséget és a szabványosítást, így támogatva a digitális biomarkerek ipari hasznosítását és egészségügyi elfogadottságát.

Kulcsszavak: digitális biomarkerek, mesterséges intelligencia, validálás, adatminőség, egészségügyi technológia

Abstract — The article explores the role of metrology in the research and development of digital biomarkers, particularly in the context of artificial intelligence and large language model integration. It highlights the lack of standardized evaluation frameworks as a major barrier to the clinical adoption of digital biomarkers. By introducing a metrology-driven approach, the study proposes systematic methods to ensure traceability, validation, and standardization, thereby fostering trust, interoperability, and industrial scalability of AI-enabled digital biomarker solutions in healthcare.

Keywords: digital biomarkers, artificial intelligence, validation, data quality, healthcare technology

1 BEVEZETÉS

A digitális biomarkerek (DBM) kutatás-fejlesztése – objektíven mért fiziológiai, viselkedési, biofizikai, általánosan biológiai adatok, amelyeket digitális eszközökkel és műszerekkel gyűjtenek és elemeznek – izgalmas határterületté vált az egészségügyi és bioinformatikai innováció terén. A tudományos, klinikai és kereskedelmi *stakeholder*-ek növekvő érdeklődése ellenére a terület fejlődését továbbra is hátráltatják a szétagolt szabványok, az inkonzisztens terminológiák és az egységes értékelési protokollok hiánya. Ez éles ellentétben áll a jól strukturált szabályozási és módszertani keretrendszerrel, melyet a hagyományos klinikai vizsgálatok területén hoztak létre, (szabályozó testületek pl. FDA és az EMA [1, 2]) ahol a molekuláris biomarkerek (pl. vérből mérhető markerek) egyértelműen

definiáltak, mind fogalmkörük, mind az értékelési és érvényesítési módszertanuk tekintetében.

A „digitális biomarker” kifejezésnek jelenleg még nincs koherens és széles körben elfogadott meghatározása. Míg bizonyos esetekben a DBM-eket a hagyományos biomarkerek digitális helyettesítőinek tekintik, mások a szenzorokból származó bizonyos mintázatokra jellemző azonosításra vagy kategorizálásra alkalmas értékeket is ide sorolják, standardizált értékelési vagy érvényesítési definíciók nélkül [3]. Ez a zűrzavar következtelenségeket eredményezhet a tudományos irodalomban, új technológiák bevezetésben és az értékesítési stratégiákban, ami akadályokat gördíthet a reprodukálhatóság, az interoperabilitás elé.

Ebben az innovációs értelemben pozitívan dinamikus, de egyben kaotikus környezetben a metrológia – a mérésstudomány – rendszerszintű megoldást kínál a K+F-folyamatok szabványosítására és harmonizálására a DBM fejlesztésében. A metrológiai szemlélet az alapkutatás és az ipari fejlesztés közötti hídként funkcionálva optimalizálhatja a költséghatékonyságot, a nyomkövethetőséget és a műszaki megvalósíthatóságot a K+F-ciklus mind a négy fázisával való együttműködésben [4]. A metrológiai módszertani megközelítések és protokollok szerepe szabványügyi szinten megalapozott a hagyományos egészségügyi területeken, például a laboratóriumi diagnosztikában és a képalkotásban, de a még csak kialakulóban lévő DBM-területhez való adaptációja számos új lehetőséget rejt magában.

A DBM-innováció során gyakran figyelmen kívül hagyott kulcsfontosságú elem a klinikai vizsgálatok előtti vagy a technológia-specifikus fejlesztéssel kapcsolatos megvalósíthatósági értékelés. A metrológiai szemléletű kísérlettervezéssel támogatott megvalósíthatósági értékeléssel megalapozott pre-klinikai K+F-keretrendszer nem csak a zavaró-befolyásoló tényezők azonosítására és a mérési bizonytalanságok számszerűsítésére alkalmas, hanem támogatja az erőforrás-optimalizálást és megalapozza mind az alapkutatási, mind az ipari fejlesztői és a hasznosítási oldal közötti bizalmat [5, 6, 7].

A mesterséges intelligencia (AI) és a nagy nyelvi modellek (LLM) integrálása a DBM K+F-be még szigorúbb metrológiai felügyeletet tesz szükségessé. Noha ezek a technológiák óriási lehetőségeket hordoznak a K+F szabványosítási megosztottságának áthidalására a protokollgenerálás és a valós idejű adatfeldolgozás, dokumentálás automatizálásával [8], kimeneteiket műszakilag validálni, kalibrálni és metrológiai protokollokon keresztül szabványosítani kell, hogy biztosítható legyen az interdiszciplináris szakértői bizalom, az átláthatóság és a reprodukálhatóság az egészségügyi alkalmazások területén [9].

Ebben a tudományos véleményjellegű rövid áttekintésben a teljesség igénye nélkül kerülnek bemutatásra olyan sarkpontok, melyek rávilágítanak a klasszikus metrológiai szemlélet szükségességére egy teljesen újszerű, interdiszciplináris K+F-környezetben. Bemutatásra kerül továbbá, hogy a klasszikus metrológiai megközelítés miként katalizálhatja az AI-alapú DBM-ek kísérleti prototípusokból klinikailag is hasznosítható végfejlesztéssé átalakulását.

2 MÓDSZEREK

A vonatkozó tudományos irodalomelemzés, a szabványértékelés, az ipari szabványok és az üzleti forrásokból származó technológiai áttekintések eredményeinek rövid, véleményjellegű kivonatai kerültek összesítésre. A PubMed, a Scopus és a Google Scholar adatbázisokban történt lekérdezésekhez használt kulcsszavak például „Digitális biomarkerek”, „Metrológia”, „Értékelés”, „Validálás”, „AI az egészségügyben”, „AI szabványok” és „Klinikai vizsgálati szabványok”. A jelenlegi szabályozási ökoszisztéma értékeléséhez az FDA, az EMA, az IMDRF, az ITU, az ISO és az IEEE szabványozási dokumentumai és tanulmányai kerültek felhasználásra.

Mint véleményjellegű rövid áttekintés, a kézirat nem új kísérleti vagy elemzési adatokat mutat be, hanem a jelenlegi ismereteket és szakértői meglátásokat szintetizálja. Az interpretációk és a javasolt keretrendszerek a szerző kritikai elemzését tükrözik a meglévő irodalomról, szabványokról és technológiai gyakorlatokról, céljuk pedig az interdiszciplináris párbeszéd ösztönzése és a jövőbeli kutatási irányok felvázolása.

2.1 Adatforrások

A vélemény formálás során különös figyelmet kaptak a DBM-fejlesztői, -tesztelési és -validációs adatbázisok:

- Open mHealth, szenzoradat-standardizációhoz [10]
- PhysioNet, validált fiziológiai jeladatbázisokhoz [11]
- Digital Biomarker Discovery Project (DBDP) [12]

2.2 Sztenderdek és protokollok

- IEEE 11073 and P2413 (IoT sztenderdek)

- ISO/IEC 17025 (tesztelő laboratóriumok kompetenciája)
- ISO 14155 (orvostechikai eszközök klinikai vizsgálata)
- Az EMA minősítési dokumentumok, segédanyagok [18]

3 EREDMÉNYEK ÉS TÁRGYALÁS

3.1 Fragmentált szabványok

A DBM széles körű elterjedésének elsődleges akadálya a protokolltervezés, az adatfeldolgozás és a validációs módszerek harmonizációjának hiánya. Ellentétben a molekuláris biomarkerekkel, melyek a nemzetközileg elfogadott bioanalitikai módszerek irányelvei alapján kerülnek fejlesztésre, tesztelésre, validálásra [13], a DBM-ek gyakran technológiai szempontból esetspecifikus, szellemi-tulajdonjogi szempontból pedig egyedi környezetben kerülnek fejlesztésre. A harmonizáció elősegítésére az EURAMET olyan egészségügyi metrológiai keretrendszerek bevezetését kezdeményezi - a vonatkozó mérésbizonytalanság-elemzés, a referenciaprotokollok nyomonkövethetősége, valamint az eszközök és mérési, számítási módszerek közötti összehasonlíthatóság érdekében -, mely keretrendszerbe foglalhatja mind az AI-támogatott, mind a klasszikus számítástudományi alapokon nyugvó DBM K+F-folyamatokat [14].

3.2 Metrológiai értékelés által történő preklinikai validálás

A DBM-fejlesztés jellemzően az adatgyűjtésre és az AI-támogatott vagy klasszikus modellezésre összpontosít, gyakran figyelmen kívül hagyva a bemeneti adatfolyamok szigorú műszaki, informatikai és minőségkontroll szintű validációját. A mérési módszerek és számítási algoritmusok értékelése jellemzően statisztikai teljesítménymutatókra fókuszáló megközelítéssel történik, nem pedig szenzor-specifikus precíziós kiértékelésekkel, kalibrációs rutinokkal vagy bizonytalanságbecslési keretrendszerekkel [15].

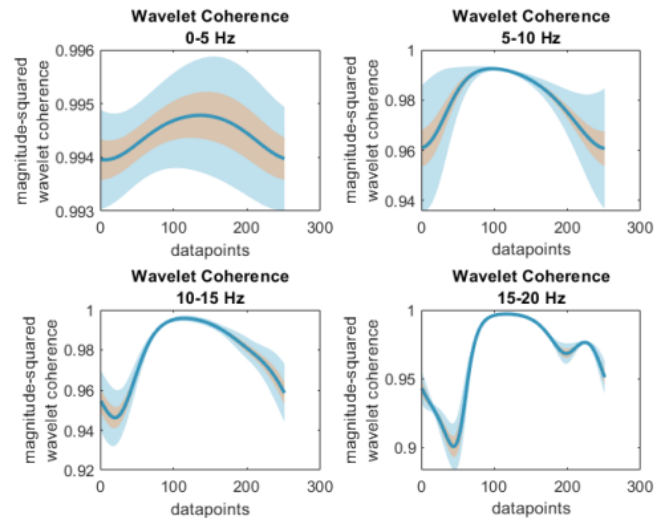
A metrológia szerepe itt a legkritikusabb: segít azonosítani a mérési vagy számítási folyamatból adódó torzításokat, a mintavételi problémákat, a nemcsak műszaki, hanem akár biológiai háttérzajból származó jeltorzulásokat és az algoritmikus inkonzisztenciákat – mindezt még a klinikai tesztelési, validációs munkafolyamatok előtt. A preklinikai metrológiai értékelésnek a DBM K+F-munkafolyamatok kötelező szakaszává kellene válnia ahhoz, hogy az alapkutatói és ipari hasznosítási folyamat megfelelően optimalizálható legyen.

3.3 Az adatbázisépítés és adatminőség problémája az AI-támogatott DBM K+F területén

A mesterséges intelligencia által támogatott algoritmusok több rétege képes feldolgozni akár strukturálatlan, szenzorokból származó nyers adatokat, akár előfeldolgozott, származtatott adatokat, melyekből egészségügyi indikátorok kerülnek definiálásra. Egy ilyen komplex munkafolyamatban keletkező eredmény megbízhatóságának interdiszciplináris elfogadottsága csak akkor alakulhat ki, ha a fejlesztés, tesztelés, értékelés, érvényesítés iteratív folyamatának támogatására specifikus adatbázisokat hozunk létre. Az adatbázisstervezés és adatbázis generáló keretrendszer kialakítására a metrológia robusztus eszközöket kínál, melyek implementálhatóak mind a szimulációs és laboratóriumi mérési folyamatok szintjén [19].

Akár szimulációval létrehozott, akár szakértői vélemény alapján annotált mérési adatbázisról van szó, a „kevesebb az több” elv alkalmazása lehetővé teszi, hogy célorientált, felhasználási terület szerint tervezett és adatminőség-kontrollmetrikával asszociált adatokkal tanított algoritmusok teljesítménymutatói felülmúlják az akár nagyságrendekkel nagyobb, de minőségben alacsonyabb értéket képviselő adatbázisokkal tanított algoritmusok képességeit. Egy lehetséges adat- és jelminőségkontroll módszer a bioszignál metrológiai eljárás, amely a fiziológiai, biológiai jelek heterogén környezetben történő mérésének, osztályozásának és integrálásának kihívásaival foglalkozik [6], és biztosítja a digitális biomarker-számításhoz használt adatok alapvető integritását. A megfelelő adatbázisépítési stratégia és az azt támogató adat-jelminőség-kontroll eljárás DBM-munkafolyamatokba való beépítése megerősíti a nyomomonkövetőségi láncot, így növeli az AI-támogatott algoritmusok által létrehozott egészségügyi indikátorok, markerek elfogadottságát, megbízhatóságát.

A viselhető elektronikai és „okos” egészségügyi szenzor- és algoritmusfejlesztések területén gyakran használt periodikus kardiovaszkuláris biojelek számos esetben nem kerülnek adat- és jelminőségi vizsgálat alá. Az így keletkező mérési bizonytalansági komponens hatása nehezen számszerűsíthető és csökkenti a bizalmat mind a társadalom, mind a professzionális felhasználók (orvos, egészségügyi szakemberek) felé. Egyik lehetséges megoldás az összehasonlító bioszignál metrológiai eljárással képzett, frekvencia specifikus Wavelet koherenciaanalízis [6], mellyel lehetőség van az adott kardio-vaszkuláris biojel (EKG, PPG, stb.) esetleges minőségi problémáiból származó mérési bizonytalanság hullám- (frekvencia-)specifikus számszerűsítésére, mely támpontot jelent az akár AI-támogatott, akár klasszikus jelfeldolgozási és jelkomponens-detekciós algoritmusok megbízhatósági intervallumainak definiálása terén (1. ábra).



1. ábra: Wavelet koherenciaanalízis [6]

A nagy nyelvi modelleket egyre gyakrabban használják az adatharmonizáció, a protokolltervezés és az irodalomszintézis segítésére a DBM-kutatásban. Bár eredményesek, kimeneteikben hajlamosak a hallucinációkra, a kontextuális eltérések generálására és a valódi összefüggések feltárásának hiányára. A metrológiai validációs szemlélettel alkotott keretrendszerek és adatbázisok alkalmazása segítheti célorientált alkalmazásukat a vonatkozó egészségügyi munkafolyamatokban [16].

3.4 Iparosítás: a felfedező DBM-ektől a szabályozott eszközökig

A metrológiai módszerek, protokollok DBM K+F-munkafolyamatokba való integrálása lehetővé teszi a K+F életciklus-minőségellenőrzését, nyomomonkövetőséget biztosítva az érzékelő kalibrálásától az algoritmusos kimenetvalidáláson át a szabályozásnak való megfelelésig. A metrológia által vezérelt gondolkodásmód egyesíti az akadémiai innovációt az ipari robusztussággal és a szabványosítási törekvésekkel, mely serkentheti a harmonizációt az EU AI-direktívái és az orvosi eszközökre vonatkozó szoftverfejlesztési keretrendszerekkel [17].

4 KONKLÚZIÓ

A DBM K+F fordulóponthoz érkezett: az AI-támogatott technológiák forradalmi lehetőségeket kínálnak, de egyben kockázatot is jelenthetnek. A metrológia-vezérelt eljárásokat a DBM K+F központi pilléreként szükséges újrapozicionálni, mely nemcsak a műszaki-informatikai megvalósíthatóságot támogatja, hanem az érdekelt felek bizalmát is megerősíti, valamint felgyorsítja az innováció befogadását és támogatja a fenntartható értékteremtést a digitális egészségügyben.

IRODALOMJEGYZÉK

- [1] Coravos, A. et al. (2019). Digital biomarkers for health monitoring. NPJ Digital Medicine.. <https://www.nature.com/articles/s41746-019-0090-4>
- [2] FDA (2018). Bioanalytical Method Validation Guidance. <https://www.fda.gov/media/70858/download>
- [3] Powell, D. (2024). Walk, talk, think, see and feel: harnessing the power of digital biomarkers in healthcare. NPJ Digital Medicine. <https://www.nature.com/articles/s41746-024-01023-w>
- [4] Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. <https://www.nature.com/articles/sdata201618>
- [5] DiMe Society (2023). The Playbook: Digital Clinical Measures. <https://playbook.dimesociety.org/>
- [6] Pálhalmi, J. et al. (2019). A Statistical Metrology Approach to Compare the Quality of Optical Pulse Wave Signals. Computing in Cardiology , Vol 46. DOI: 10.22489/CinC.2019.165 <https://www.analog.com/media/en/technical-documentation/tech-articles/optical-integration-without-compromises.pdf>
- [7] Tretter, M. et al. (2025). AI-produced certainties in health care: current and future challenges. AI and Ethics. Springer Nature. <https://link.springer.com/article/10.1007/s43681-023-00374-6>
- [8] Bommasani, R. et al. (2022). On the Opportunities and Risks of Foundation Models. arXiv <https://arxiv.org/pdf/2108.07258>
- [9] Chen, S.-Y. et al. (2024). Perceptions of ChatGPT in healthcare: usefulness, trust, and risk. Frontiers in Public Health. <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2024.1457131/full>
- [10] Open mHealth. <https://www.openmhealth.org/>
- [11] PhysioNet. <https://physionet.org/>
- [12] <https://www.dbdp.org/>
- [13] FDA Guidance on Clinical Trials. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents>
- [14] EURAMET EMN for Health. <https://www.euramet.org/metrology-for-societys-challenges/metrology-for-health>
- [15] Mathews, S. C. et al. (2019). Digital health: a path to validation. NPJ Digital Medicine. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6550273/>
- [16] Large Language Models for Medical Applications. Frontiers in Medicine. 2025. <https://www.frontiersin.org/research-topics/64598/large-language-models-for-medical-applications>
- [17] IMDRF SaMD Definitions. <https://www.imdrf.org/documents/software-medical-device-samd-key-definitions>
- [18] European Medicines Agency: <https://www.ema.europa.eu/en/human-regulatory-overview/research-development/scientific-advice-protocol-assistance/opinions-letters-support-qualification-novel-methodologies-medicine-development>
- [19] Pálhalmi, J. et al. (2024). AI-Powered Microscopy Platform for Airborne Biothreat Detection. Springer Nature. https://link.springer.com/chapter/10.1007/978-3-031-62083-6_10