


Valós klinikai adatok a klinikai kutatásban: adatmodellek és informatikai támogató rendszerek áttekintése

Székely Orsolya¹  ■ Mészáros Ágota dr.^{1, 2, 3} ■ Kővári Eszter dr.¹
 Assabiny Alexandra dr.^{1, 4} ■ Héja Tibor¹ ■ Borbély Dóra¹
 Kovács Attila dr.^{1, 4} ■ Nemes Csaba¹ ■ Sallai András¹ ■ Bali Orsolya¹
 Kiss Loretta dr.^{1, 4} ■ Bagyura Zsolt dr.¹

¹Semmelweis Egyetem, Klinikai Adatszolgáltató Intézet, Budapest

²Semmelweis Egyetem, Egészségügyi Technológiaértékelő és Elemzési Központ, Budapest

³Semmelweis Egyetem, Farmakológiai és Gyógyszerkutatás-fejlesztési Központ, Budapest

⁴Semmelweis Egyetem, Általános Orvostudományi Kar, Városmajori Szív- és Érgyógyászati Klinika, Budapest

A valós világból származó egészségügyi adatok (real-world data – RWD) másodlagos felhasználása egyre fontosabb szerepet tölt be a bizonyítékon alapuló orvoslásban, ugyanakkor a különböző rendszerekben gyűjtött adatok közötti eltérések kihívást jelentenek a megbízható, valós világbeli klinikai bizonyítékok (real-world evidence – RWE) előállításában. Az Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) olyan nemzetközi adatmodell és eszköztár, amely egységes szerkezetet és szemantikát biztosít az eltérő struktúrájú adatforrások számára, ezáltal támogatja a nagy léptékű, reprodukálható kutatásokat és az összehasonlítható eredmények létrehozását. Az adatmodell lehetővé teszi, hogy az adatbirtokosok saját rendszereikben tárolják az adatokat, miközben egy több adatbázist összekapcsoló, ún. federált hálózaton keresztül csak az összesített elemzési eredmények kerülnek megosztásra, összhangban az adatvédelmi és engedélyezési követelményekkel. A jelen áttekintő közlemény célja, hogy bemutassa az OMOP CDM felépítését és működési logikáját, összefoglalja az adatstandardizálás és adatminőség-ellenőrzés alapelveit, illetve ismertesse azokat a nemzetközi és hazai kezdeményezéseket, amelyek ezt az adathasznosítási megközelítést alkalmazzák. Az adatmodell használatának előnye, hogy jó minőségű, összehasonlítható és klinikailag releváns bizonyítékok hozhatók létre a segítségével, ezáltal növelve a betegbiztonságot, javítva az egészségügyi ellátás minőségét és támogatva az orvosi döntéshozatalt.

Orv Hetil. 2026; 167(17): 661–672.

Kulcsszavak: standardizálás, megfigyeléses tanulmányok, rutinszerűen gyűjtött egészségügyi adatok, adatelemzés

Real-world data in clinical research: a review of data models and supporting informatics systems

Real-world data (RWD) are increasingly used for secondary purposes in evidence-based medicine; however, differences between data collected in various healthcare systems pose challenges to generating reliable real-world clinical evidence (RWE). The Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) is an international data model and toolset that provides a harmonized structure and semantics for heterogeneous data sources, thereby supporting large-scale, reproducible research and the production of comparable results. The data model enables data holders to retain their data within local systems, while only aggregated analytical outputs are shared through a federated network connecting multiple databases, in line with data protection and regulatory requirements. The aim of this narrative review is to describe the structure and operational logic of the OMOP CDM, to summarize the key principles of data standardization and data quality assessment, and to present international and national initiatives that apply this data utilization framework. The use of the data model facilitates the generation of high-quality, comparable and clinically relevant evidence, thereby enhancing patient safety, improving the quality of healthcare services and supporting clinical decision-making.

Keywords: standardization, observational studies, routinely collected health data, data analytics

Székely O, Mészáros Á, Kővári E, Assabiny A, Héja T, Borbély D, Kovács A, Nemes Cs, Sallai A, Bali O, Kiss L, Bagyura Zs. [Real-world data in clinical research: a review of data models and supporting informatics systems]. *Orv Hetil.* 2026; 167(17): 661–672.

(Beérkezett: 2025. január 8.; elfogadva: 2026. február 28.)

Rövidítések

ACHILLES = (Automated Characterization of Health Information at Large-scale Longitudinal Evidence Systems) egészségügyi információk automatizált jellemzése nagy méretű longitudinális bizonyítékhálózatokban; CDM = (Common Data Model) közös adatmodell; COVID-19 = (coronavirus disease 2019) koronavírus-betegség 2019; DARWIN EU® = (Data Analysis and Real World Interrogation Network) Adatelemzés és Valós Világbeli Megfigyelési Hálózat; DQD = (Data Quality Dashboard) adatminőség-irányítópult; EFPIA = (European Federation of Pharmaceutical Industries and Associations) Európai Gyógyszeripari Szereplők és Egyesületek Szövetsége; EHDEN = (European Health Data & Evidence Network) Európai Egészségügyi Adat- és Bizonyítékhálózat; EHDS = (European Health Data Space) Európai Egészségügyi Adattér; EHR = (electronic health record) elektronikus egészségügyi dokumentációs rendszer; EMA = (European Medicines Agency) Európai Gyógyszerügynökség; ETL = (extract, transform and load) kinyerés, átalakítás és betöltés; LOINC = (Logical Observation Identifiers Names and Codes) Logikai Megfigyelési Azonosítók Nevei és Kódjai; OENO = Orvosi Eljárások Nemzetközi Osztályozása; OHDSI = (Observational Health Data Sciences and Informatics) Megfigyelésen Alapuló Egészségügyi Adattudományok és Informatika; OMOP = (Observational Medical Outcomes Partnership) Megfigyelésen Alapuló Orvosi Eredmények Társulása; PCORnet = (Patient-Centered Clinical Research Network) Betegközpontú Klinikai Kutatóhálózat; RWD = (real-world data) valós környezetben keletkezett adatok; RWE = (real-world evidence) valós környezetben keletkezett bizonyítékok; SNOMED (CT) = (Systematized Nomenclature of Medicine (of Clinical Terminologies) a gyógyászat (klinikai terminológiák) rendszerezett nomenklatúrája

Magyarzatok

ACHILLES = olyan eszközrendszer, amely segíti az egészségügyi adatok elemzését; ATLAS = olyan webes alapú szoftveres applikáció, amely az egészségügyi adatok standard elemzésének tervezését és megvalósítását segíti elő; CDM = a standard adatoknak olyan rendszere, amely tartalmazza a standardok közötti egyenértékű megfeleltetéseket és az egészségügyi fogalmak közötti hierarchiákat, így az adatok egységes rendszerben értelmezhetővé válnak; DQD = az adattérképezés során használt eszköz a térképezési eredmény minőségének ellenőrzése érdekében; ETL = az egészségügyi adatok közös adatmodellben való használatát elősegítő informatikai folyamat; fogalomtérkép = hozzárendelések, illesztések eredménye; mappelés = adattérképezés, illesztés, azaz a standard egészségügyi fogalom megfeleltetése egy másik (például nemzetközi) standard azonos jelentéstartalmú fogalmának

Az egészségügyi adatok gyűjtésének és elemzésének igénye az orvostudomány kialakulásával egyidős, ám módszerei folyamatosan változnak. Az adatok gyűjtésének elsődleges célja az ellátás, tehát a páciensállapotok megfigyelésének és nyomon követésének dokumentálása. Kutatási szempontból ezek az adatok bizonyítékként, ún. evidenciaként szolgálnak, közvetlen visszacsatolást adva a terápiák hatásáról. Az evidenciákon alapuló orvoslásnak (evidence-based medicine) köszönhetjük, hogy ma klinikai irányelvek segítik a gyógyító munkát; ezekben az irányelvekben a különböző eljárásokból összehasonlíthatóvá, értékelhetővé és prioritizálhatóvá válnak [1].

A bizonyítékokat adatgyűjtés és adatelemzés alapján állíthatjuk fel. Számos különböző adatforrás definiálható az egészségügyben, mint például klinikai vizsgálatok adatbázisai, finanszírozási adatok, elektronikus egészségügyi dokumentumok [2], betegségregiszterek [3], ellátásszervezési adatok, amelyek mind alkalmasak arra, hogy információt biztosítsanak a páciensek egészségi állapotáról, az eljárások, beavatkozások és kezelések eredményességéről, valamint az ellátórendszer működéséről.

Azokat az adatokat, amelyek a rutinellátáshoz kapcsolódóan keletkeznek az egészségügyi rendszerben, nem pedig beavatkozással járó klinikai vizsgálatok eredményeként, „real-world data”-nak (RWD) nevezzük. Az ezen adatokra épülő, beavatkozással nem járó vizsgálatok [4] alapján felállított bizonyítékokat pedig „real-world evidence”-nek (RWE) nevezzük. RWE különböző módszertanú vizsgálatok eredményeként keletkezhet, például: adatbázisok vagy kohorszok jellemzése [5], populációszintű hatásbecslések vagy személyre szabott előrejelzések készítése [6, 7]. Az RWE tehát kiegészíti a bizonyítékokon nyugvó orvoslás bizonyítékalapját, és bővíti azt a valós környezetben keletkezett adatokkal, információkkal. A továbbiakban az RWE, megfigyeléses vizsgálat alatt a beavatkozással nem járó vizsgálatok értendők [4].

Az RWE létrehozása kihívásokkal teli folyamat. Elvben több adatbázison kutatva megbízhatóbb eredményre juthatunk, azonban az egészségügyi adatokat különböző célokból, különböző módszerekkel és különböző nyilvántartási rendszerekben gyűjtjük, ezáltal a formátumuk és minőségük is eltérhet egymástól [5, 6]. Az alábbiakban, az adatok sokféleségének szemléltetésére, áttekintést nyújtunk az elektronikus egészségügyi adatterről

szóló rendeletben (EHDS-rendelet) meghatározott, másodlagos célból rendelkezésre bocsátandó adatforrásokról és adattípusokról:

- a) „az EHR-ekből származó adatok”,
- b) az egészséget befolyásoló tényezőkre vonatkozó adatok, beleértve az egészséget meghatározó társadalmi-gazdasági, környezeti és viselkedésbeli tényezőket,
- c) összesített adatok az egészségügyi szükségletekről, az egészségügyi ellátásra elkülönített forrásokról, az egészségügyi ellátás biztosításáról és az ahhoz való hozzáférésről, az egészségügyi kiadásokról és finanszírozásról,
- d) az emberi egészségre hatást gyakorló kórokozókra vonatkozó adatok,
- e) az egészségügyi ellátással kapcsolatos adminisztratív adatok, beleértve a gyógyszerkiadásokra, biztosítási igényekre, megtérítési igényekre és megtérítésekre vonatkozó adatokat is,
- f) emberi genetikai, epigenomikai és genomikai adatok,
- g) egyéb emberi molekuláris adatok, mint például proteomikus, transzkriptomikai, metabolikus, lipidomikai és egyéb omikus/omikai adatok,
- h) orvostechnikai eszközök segítségével automatikusan generált személyes elektronikus egészségügyi adatok,
- i) jólléti alkalmazásokból származó adatok,
- j) a természetes személyek kezelésében részt vevő egészségügyi szakemberek szakmai státuszára, valamint szakterületére és intézményére vonatkozó adatok,
- k) lakosság alapú egészségügyi adat-nyilvántartásokból, így például népegészségügyi nyilvántartásokból származó adatok,
- l) orvosi nyilvántartásokból és halálozási nyilvántartásokból származó adatok,
- m) az 536/2014/EU rendelet, az (EU) 2024/1938 európai parlamenti és tanácsi rendelet, az (EU) 2017/745 és az (EU) 2017/746 rendelet hatálya alá tartozó klinikai vizsgálatokból és klinikai kutatásokból származó adatok,
- n) az orvostechnikai eszközökből származó egyéb egészségügyi adatok,
- o) gyógyszerek és orvostechnikai eszközök nyilvántartásaiból származó adatok,
- p) az egészséggel kapcsolatos kutatási kohorszokból, kérdőívekből és felmérésekből származó adatok, a vonatkozó eredmények első közzétételét követően,
- q) biobankokból és a kapcsolódó adatbázisokból származó adatok [8].

Amennyiben általános érvényű, több különböző populáció esetében is helytálló tudományos megállapításokat kívánunk tenni, az adatok diverzitásából adódó problémák növekednek, különösen akkor, amikor különböző országok egészségügyi rendszereiből származó adatot

vonunk be a kutatásba. Ennek ellenére a különböző adatforrások nagyobb mértékű generalizálhatóságot eredményeznek.

A különböző forrásokból származó adatok elemzéséhez az információforrások és az adatok mélyreható technikai ismeretére van szükség. Amikor az eltérő források miatt ezek halmozódnak, a szükséges szakértelem is hatványozódik, ami a gyakorlatban akár megoldhatatlan problémát jelenthet a kutatóknak. Ezért sokkal előnyösebb egy olyan megközelítés, amely interdiszciplináris együttműködésekben és adatkezelésben jártas csapatok kollaborációján alapul. Ilyen kollaboráció például az Observational Medical Outcomes Partnership (OMOP), amely eredetileg az Egyesült Államokban a US Food and Drug Administration, a Foundation for the National Institutes of Health, gyógyszercegekéből álló konzorciumok, akadémiai kutatók és ún. adatpartnerek között jött létre 2008-ban [9] azzal a céllal, hogy a megfigyelésen alapuló egészségügyi adatokat felhasználva fejleszteni tudja a gyógyszerbiztonsági felügyelet tudományos támogatottságát [10]. A különböző adatforrások kezelésével kapcsolatos kihívásokat felismerve a közösség létrehozta az OMOP Common Data Modelt (CDM), amely megteremtette a lehetőséget az adatok struktúrájának, tartalmának, szemantikájának standardizálására [11].

Ennek köszönhetően alakult ki az a módszertan, amelynek célja elősegíteni, hogy a különböző forrásokból származó adatok egységes elvek mentén hasznosíthatóvá váljanak tudományos kutatások céljára. Ebben a cikkben áttekintést nyújtunk az OMOP CDM-ről és a hozzá kapcsolódó további együttműködésekről, projektekről. Emellett ismertetjük az ilyen vizsgálatok módszertani alapjait, beleértve a vizsgálatok és adatfeldolgozási folyamatok minőségbiztosítási szempontjait. Célunk az, hogy megismertessük a valós környezetben keletkezett egészségügyi adatokra alapuló vizsgálatok módszertanát. Ezek az ismeretek biztosítják a kutatók számára az RWE megfelelő értelmezését. Ezáltal a hazai kutatói közösség számára is bővülnek az OMOP által nyújtott lehetőségek.

OMOP-alapokon működő RWE-közösségek

Az OMOP tehát gyógyszerbiztonsági projektként indult, de amikor teljesültek a célkitűzések, és lezárult a projekt, továbbra is fennmaradt az igény az interdiszciplináris együttműködésekre a megfigyeléses kutatások területén, így létrejött az Observational Health Data Sciences and Informatics (OHDSI) együttműködés [12]. Ennek célja, hogy összefogja az érdekelt feleket, lehetőséget teremtve a kollaborációra az egészségügyi adatok hasznosításához. Az együttműködés keretében létrejött egy olyan, kutatókból, adatgazdákból és ipari szereplőkből álló önkéntes nemzetközi hálózat, amely összeköti a megfigyelési adatokat használó szakértőket nagyszabású elemzések megvalósítása érdekében. Az így

létrejött közösség olyan nyílt forráskódú megoldásokat (OHDSI-eszközök) és módszereket (OHDSI-standardok) hozott létre és tart fenn kollaborációs platformok segítségével, amelyek támogatják az RWE-adatbázisok létrehozását és hasznosítását [12, 13].

Az OMOP CDM-alapú, tehát nyílt felhasználású adatmodell, amelynek lényege, hogy akadémiai kapcsolatoktól függetlenül bárki használhatja, illetve javaslatot tehet a továbbfejlesztésére. A közösség ezeket a javaslatokat rendszeres időközönként egy kidolgozott eljárásrend szerint elbírálja. A fejlesztések új irányelvek, szoftverek és szoftververziók formájában a felhasználói közösség számára is elérhetőek lesznek [14].

A nemzetközi kezdeményezések folytatásaként az Európai Unióban is létrejött egy olyan nyitott tudományos közösség, amely összevont hálózatot képez az evidencia-alapú orvoslás támogatására. A European Health Data & Evidence Network (EHDEN) a Horizon 2020 program és az Európai Gyógyszeripari Szereplők és Egyesületek Szövetsége (European Federation of Pharmaceutical Industries and Associations, EFPIA) közös finanszírozásával létrehozott projektként indult 2018-ban [15]. A projekt három pillér mentén valósult meg: 1) az OMOP CDM alapú összevont infrastruktúra kiépítése, 2) kutatási lehetőségek megteremtése és 3) képzési programok megvalósítása. A célkitűzés összességében az volt, hogy a részt vevő felek rendelkezésre álló egészségügyi adatai illesztésre kerüljenek az OMOP CDM-hez, szükség esetén úgy, hogy megvalósítják az adatok lefordítását, szakszóval mappelését (adattérképezés, illesztés) egy OMOP CDM-ben szereplő nemzetközi terminológiához. A projekt keretében fenntartható, nyitott, tudományos együttműködési modell jött létre Európában [16].

Magyarországon öt nyertes pályázó vett részt az EHDEN projektben, köztük a Semmelweis Egyetem is [17]. A projekt keretében, egyebek mellett, a hazai egészségügyi eljárások kódjainak egy nemzetközi kódrendszer – a Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) [18] – terminológiájára való térképezése is megvalósult a Pécsi Tudományegyetemmel közös munkában [19]. Emellett megalakult a Magyarországi Nemzeti OHDSI Csomópont (OHDSI National Node Hungary), amelynek vezető szervezete a Semmelweis Egyetem Klinikai Adatszolgáltató Intézet, tagjai között pedig további akadémiai és piaci szereplők vannak jelen [20]. A Csomópont létrehozásának köszönhetően az EHDEN projekt lezárását követően is folytatódnak a törekvések a további hazai OMOP CDM-adaptációk megvalósítására.

Adatpartner csatlakozása az OMOP-hoz, az OMOP komponensei

Egy OMOP CDM adatbázis kialakításához olyan folyamatra van szükség, amely az adatok kivonásából, átalakításából és betöltéséből áll (extract, transform and load – ETL). Ennek eredményeként a nyers adatok egységes,

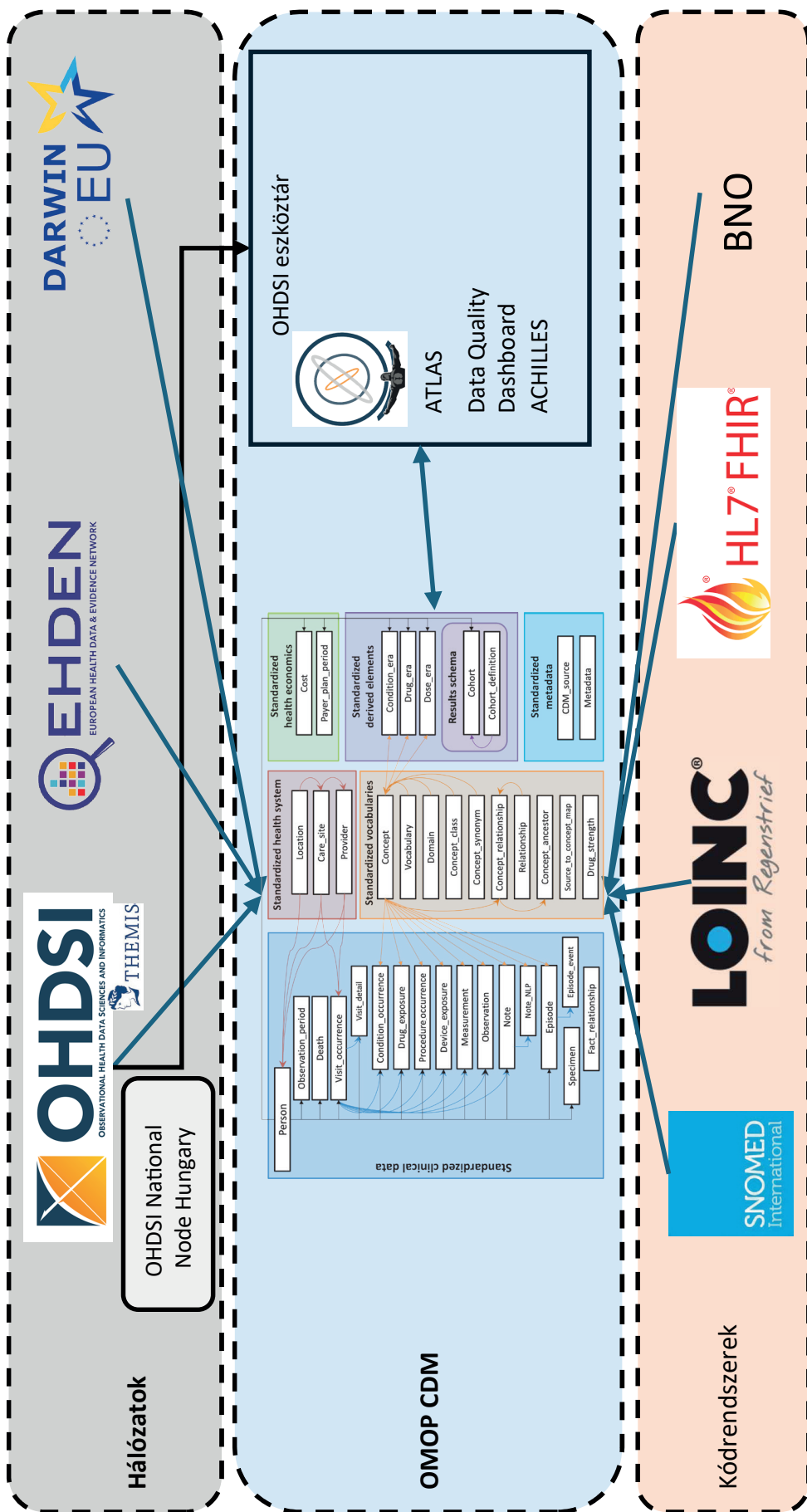
szabványosított formátumban állnak elő. Az ETL-folyamat többféle kontextusban értelmezhető, például üzleti intelligencia, „big data” vagy szemantikai szempontból [21]. A jelen megközelítésben a szemantikai kontextusra helyezük a hangsúlyt, mivel ennek szerepe az adatok heterogenitásának áthidalása.

Az OHDSI olyan négy fázisból álló megközelítést azonosított a legjobb gyakorlatként, amelyben a szereplők is meghatározásra kerültek: 1) ETL-folyamat megtervezése adatszaktörők és adatmodell-szaktörők által; 2) fogalomtérképek létrehozása orvosi ismeretekkel rendelkező informatikai munkatársak által; 3) ETL-folyamat megvalósítása műszaki szakemberek által; 4) minőségértékelés minden érintett bevonásával [5]. Így érvényre jut az interdiszciplináris megközelítés, és minden érintett (informatika, adminisztratív és finanszírozási munkatársak) bevonásra kerül a folyamat megvalósítása során.

A harmadik, technikai megvalósítási fázis kapcsán fontos megemlíteni, hogy az OHDSI átfogó támogatási csomagot kínál. Az eszközrendszer karbantartása iteratív folyamat, amelyre különösen az OMOP CDM új verzióinak vagy szótárfrissítéseknek a megjelenésekor van szükség.

Az OMOP CDM kulcsfontosságú alkotóelemei közé tartoznak a Standardizált Szótárak (Standardized Vocabularies), amelyek több doménre (területre, tartományra) terjednek ki, beleértve a gyógyszereket, diagnózisokat, eljárásokat, eszközöket, megfigyeléseket, méréseket stb. Ez lényegében a közös nyelv, az interoperabilitás alapja. A Standardizált Szótárakban összesen közel 10 millió egészségügyi fogalom található [22] számos kódrendszerben, mint például a SNOMED vagy a laboreredmények leírását szolgáló Logical Observation Identifiers Names and Codes (LOINC) [23]. A betegellátásban ezeket a standard terminológiákat különböző mértékben használják az egyes országok. Hazai szinten a standard kódrendszerek használata (például SNOMED és LOINC) még nem terjedt el, és emellett a betegdokumentációban igen hangsúlyosan megjelennek nem strukturált adatok is, mint például az anamnézis szöveges leírása. Az OMOP CDM és a használt terminológiák, illetve az OMOP CDM-et használó hálózatok viszonyrendszerét az 1. ábra ismerteti [11, 12, 15, 18, 23–27].

Amennyiben egy adatforrás a nemzetközi standardoktól eltérő terminológiákat, például országos érvényességű kódrendszert használ, vagy egyáltalán nem használ ilyet, akkor a fogalmak között új kapcsolatokat kell megteremtetni, vagyis illeszteni (mappelni) kell őket egy nemzetközi standardhoz. Az elsődleges cél az adattérképezési módszertanok alapján megtalálni egy 1 : 1 megfeleltetést, az azonos fogalmat a nemzetközi terminológiák között. Tekintsük példaként az Orvosi Eljárások Nemzetközi Osztályozása (OENO) 33974 „Coronaria stent beültetés” kódot, amelyet a SNOMED-hez szeretnénk illeszteni, így a következő azonos jelentésű megfelelőt kapjuk: SNOMED 4283892 „Placement of stent in



1. ábra

Az OMOP CDM, a kapcsolódó hálózatok, eszköztár és terminológiák kapcsolati rendszere (az ábra forrása: a szerzők saját szerkesztése, képek és információk forrásai: [11, 12, 15, 18, 23–27])

ACHILLES = egészségügyi információk automatizált jellemzése nagy méretű longitudinális bizonyítékrendszerekben; BNO = Betegségek Nemzetközi Osztályozása; CDM = közös adatmodell; DARWIN EU® = Adatelemzés és Valós Világbeli Megfigyelési Hálózat; EHDEN = Európai Egészségügyi Adat- és Bizonyítékhálózat; HL7® FHIR® = Egészségügyi Interoperabilitási Erőforrások; LOINC = Logikai Megfigyelési Azonosítók Nevei és Kódjai; OHDSI = Megfigyelésen Alapuló Egészségügyi Adattudományok és Informatika; OMOP = megfigyelésen alapuló orvosi eredmények társulása; SNOMED = a gyógyszer (klinikai terminológiák) rendszerezett némenklatúrája

coronary artery”. Ha az illesztés egyáltalán nem lehetséges, új standardként be kell vezetni az OMOP CDM egységes formátuma szerint. Annak ellenére, hogy a lehetőség adott, az új standardok létrehozása a jelentős karbantartási feladatok miatt az OHDSI által nem kifejezetten támogatott módszer [5]. Fontos továbbá figyelembe venni, hogy az illesztések úgy történjenek, hogy ne keletkezzen jelentős információvesztés, így a teljes információ megőrzése érdekében olyan esetek is előfordulnak, amikor egy fogalom egyszerre több standard terminológiához is hozzárendelésre kerül [5]. A kódrendszerek hierarchikus felépítését és a leképezési kapcsolatokat a 2. ábra szemlélteti [28]. A hozzárendelések, illesztések eredményét fogalomtérképnek nevezzük.

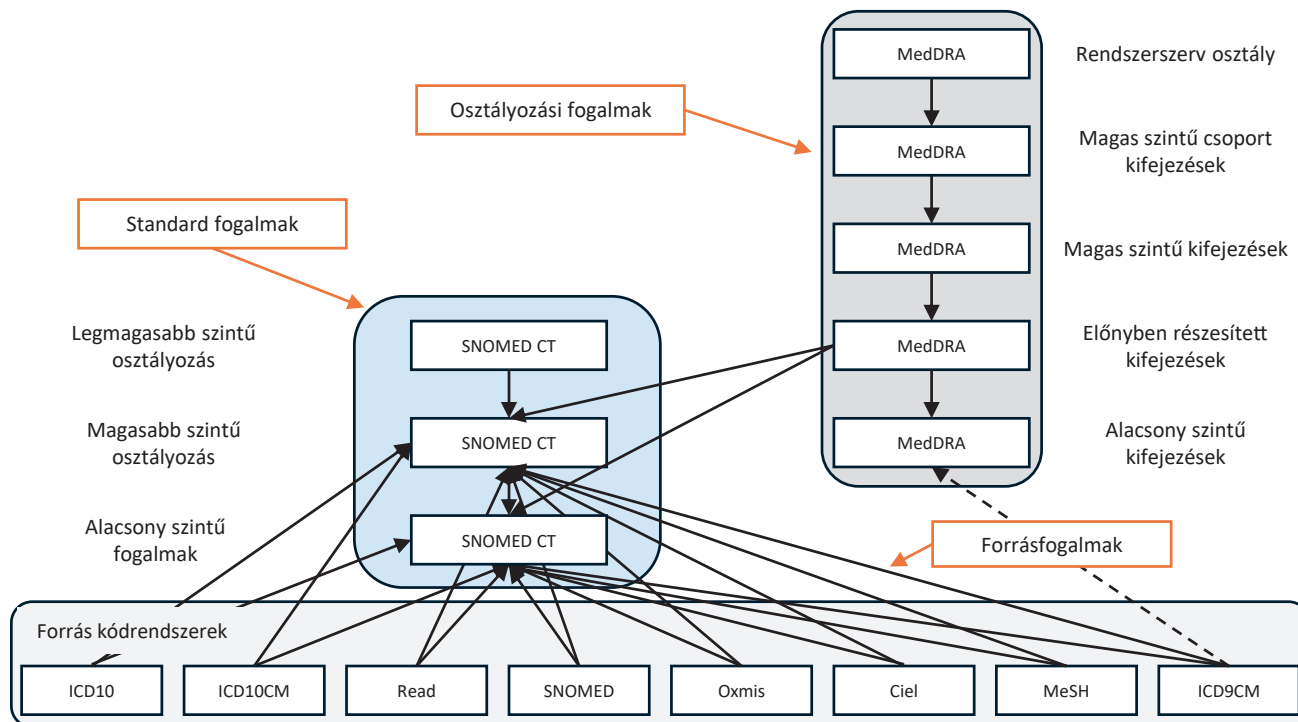
Ugyanakkor, az egyezmények használata teljes mértékben nem elegendő ahhoz, hogy elérjük a kívánt adatminőséget. Az OMOP CDM, a standardok használata mellett, jelentős hangsúlyt fektet a minőségbiztosításra is [29]. A vizsgálati bizonyítékok minősége négy komponens köré épül: 1) az adatok minőségére, 2) a klinikai érvényességre, 3) a szoftver hitelességére (az alkalmazott eszközök és elemzési folyamatok megbízhatósága) és 4) a módszer érvényességére is [5].

A bizonyítékok minősége általánosságban a kutatások tervezésével van szoros összefüggésben. A jó minőségű, megbízható megállapításokat tevő vizsgálatoknak megismételhetőnek (azonos eredmény ismételt megvalósítás esetén), reprodukálhatónak (más kutatók által való meg-

ismétlés esetén is azonos eredmény), lemásolhatónak (hasonló feltételek mellett hasonló eredmények), robusztusnak (alternatív elemzési módszerek alkalmazása esetén hasonló eredmények), hitelesíthetőnek és ellenőrizhetőnek (érvényességet befolyásoló negatív tényezők azonosíthatósága) kell lenniük [5].

A megbízható bizonyítékokat szolgáltató tanulmány egyik legfontosabb feltétele a jó adatminőség, amely azért érdemel különös figyelmet, mert a megfigyeléses kutatásoknál használt adatokat elsősorban nem kutatási célokkal gyűjtötték és rögzítették [30]. Számos probléma fakadhat az adatok kézi beviteléből, az adatok tárolásával kapcsolatos dokumentációs hiányosságokból, az adatok értelmezéséből és az adatfeldolgozásból. Ezek közül az OMOP CDM-be való adatbetöltés során a dokumentációs hiányosságok jól kontrollálhatók, a hibák pedig csökkenthetők. Az adatok értelmezésének kontrollálását a vizsgálati bizonyítékok fentebb ismertetett minőségi komponensei (2–4) segítik [5]. A minőség-ellenőrzés egységes, jól ellenőrzött szoftvercsomaggal valóítható meg.

Az adatminőségnek három típusát: 1) a megfelelőséget (conformance), 2) a teljességet (completeness), 3) a hihetőséget (plausibility), illetve két értékelési módját különböztethetjük meg: az ellenőrzést (verification) és a validációt (validation) *Kalm és mtsai* nyomán [30]. Az adatok megfelelőségének ellenőrzésére az OHDSI-közösség olyan eszközöket hozott létre, amelyek az el-



2. ábra

A kódrendszerek és standard koncepciók közötti hierarchia szemléltetése (az ábra forrása: a szerzők saját szerkesztése *Kent és mtsai* [28] alapján)

ICD = Betegségek Nemzetközi Osztályozása; ICD9CM = Betegségek Nemzetközi Osztályozása, 9. revízió (kiadás), klinikai módosítás; ICD10CM = Betegségek Nemzetközi Osztályozása, 10. revízió (kiadás), klinikai módosítás; MedDRA = Szabályozó Tevékenységek Orvosi Szótára; MeSH = Orvosi Tárgejszavak Terminológiájának Orvosi Szótára; SNOMED (CT) = a gyógyszerészet (klinikai terminológiák) rendszerezett nomenklatúrája

ellenőrzésekre automatizált módon nyújtanak lehetőséget, ez az Automated Characterization of Health Information at Large-scale Longitudinal Evidence Systems (ACHILLES) [31]. Ez a szoftver a CDM-nek megfelelően jellemzi és megjeleníti az adatbázist úgy, hogy az adatok értékelhetők legyenek [32]. Egy másik, az adatminőség javítását célzó eszköz a Data Quality Dashboard (DQD) [33], amely a vizuális ábrázolás helyett táblánként és mezőnként számszerűsítve értékeli az adatokat aszerint, hogy megfelelnek-e a CDM-specifikációknak, vagy sem. Összesen több mint 1500 ilyen ellenőrzést futtat a rendszer, és a kapott minőségi értékeket minden esetben összeveti az elfogadható küszöbértékkel, majd „megfelelt” vagy „nem megfelelt” minősítést határoz meg [5].

Az OHDSI ezen ellenőrzések elvégzését általában minden elemzést megelőzően javasolja, de szükség lehet olyan ellenőrzésekre is, amelyek az adott vizsgálat esetében specifikusak. Például egyetlen új, piacra vezetés előtt lévő gyógyszernek nem lenne szabad megjelennie az adatok között, csak majd a piacra vitelt követően. Amennyiben ilyen vagy hasonló, a várttal ellentétes mintázatokat figyelhetünk meg, ott vélhetően adatminőségi probléma merül fel. Az esetleges hibák kiküszöbölésére az elemzett adatok esetében javasolt az illesztések ismételt áttekintése. A CDM-ben nem csupán a standard fogalmak, hanem az eredetileg a forrásban rögzített kódok is eltávolíthatók, így informatikailag könnyedén megoldható az ismételt ellenőrzés [5].

Ezek az ellenőrzési pontok elősegítik, hogy egy vizsgálat során minél megbízhatóbb módszereket és eszközöket felhasználva, minél pontosabb és megbízhatóbb eredményeket érjünk el.

Kutatás az OMOP CDM adatokon

Az összevont hálózat lényege, hogy a nyers, elemi és személyes adatot nem szükséges átadni a kutatók részére, hanem a helyi informatikai rendszeren belül valósul meg a vizsgálat. A kutatás kezdeményezője átadja a kutatási protokollt és az elemzési kódot az adatpartnernek, aki lefuttatja azt a rendelkezésre álló adatokon, betartva a helyi adatvédelmi és kutatásengedélyezési szabályokat, majd az aggregált eredményeket, statisztikákat visszacsatolja a vizsgáló fél számára [34].

A szakmai közösség a kutatási, adathasznosítási tevékenységét három fő használati esethez (use case) kapcsolódóan végzi: 1) klinikai jellemzések készítése a betegség lefolyása, a terápia-igénybevétel és a minőségjavítás érdekében, ez elsősorban leíró statisztikát követel meg; 2) populációsintű hatásbecslés (population-level effect estimation) az ok-okozati tényezők feltárására; 3) páciensintű megközelítések, mesterségesintelligencia-alapú algoritmusok előrejelzései a precíziós orvoslás és a betegségmegelőzés érdekében [5]. Ezek a megközelítések eltérő módszertanokat igényelnek, amelyeket az OHDSI különböző nyílt forráskódú szoftveres megoldá-

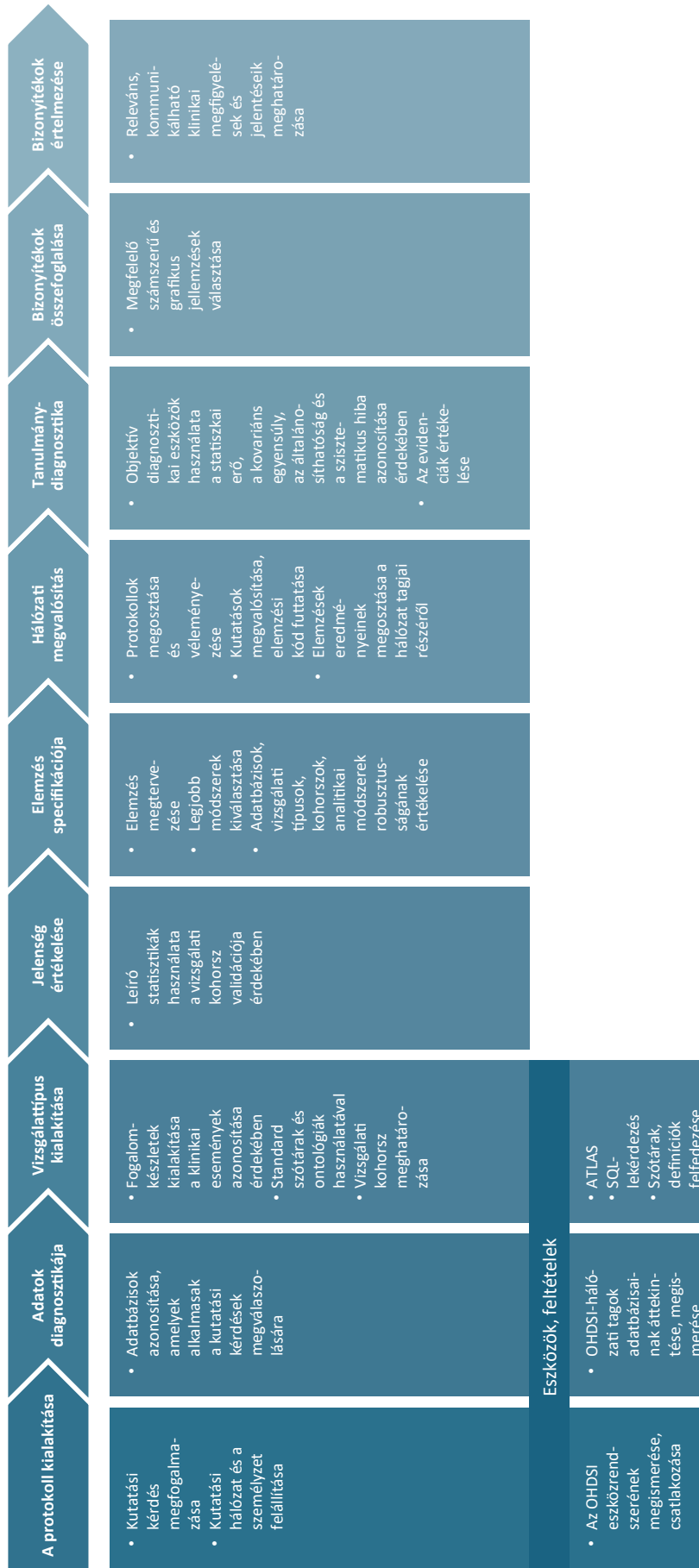
sokkal és analitikai eszközökkel támogat, mint például R statisztikai programcsomagok, amelyeket a nyilvános Github felületen tesznek elérhetővé [34, 35].

A betegadatokon való kutatás az OMOP CDM-ben tárolt formátumon többlépcsős folyamat, amelyet a 3. ábra részleteiben ismertet [36, 37]. A vizsgálati módszereket tekintve három különböző eljárás alkalmazható: 1) egyedi elemzési programkód megírása és futtatása; 2) az OHDSI Módszertani Könyvtárban elérhető nyílt forráskódú R-programcsomagok használata; 3) az OHDSI-közösség által fejlesztett ATLAS [26] interaktív analitikai platform használata. Ezek közül az első két módszer jelentős technikai, programozási ismereteket igényel, míg az utolsó felhasználóbarát megközelítést tesz lehetővé [5].

Az OHDSI-hálózaton belül az elmúlt években olyan RWD-alapú elemzések valósultak meg, mint például a krónikus betegségek kezelési lehetőségeinek [38], a rákos betegek korai kezelési módszereinek vizsgálata [39] vagy a 2-es típusú cukorbetegség hatékony klinikai kezelési módszereinek megismerése megfigyelési adatok alapján [40]. *Hripcsak és mtsai* [38] például a hálózatban részt vevő 11 adatforrásból 4 ország adatait elemezve összesen 250 millió beteg egészségügyi dokumentációját és finanszírozási adatát dolgozták fel. A vizsgálat célkitűzése a 2-es típusú diabetes, a magas vérnyomás és a depresszió kezelési útvonalainak felmérése volt. A kutatás rámutatott arra, hogy a világ az idő múlásával egyre következetesebbé vált a terápiák alkalmazásában, azaz összességében egységesebbé vált a terápiás gyakorlat, de az adatforrások közötti heterogenitás még mindig szignifikáns. A 2-es típusú diabetes esetében nagyobb volt a konszenzus az első terápia tekintetében, mint a magas vérnyomás vagy a depresszió esetében. Ugyanakkor az egyedi terápiás utat követő páciensek aránya a vizsgált kohorszon belül a diabetes és a depresszió esetében 10%, míg a magas vérnyomás esetében 25%. Az eredmények továbbá azt is alátámasztották, hogy az OHDSI-hálózat alkalmas nagy léptékű megfigyeléses kutatások megvalósítására [38].

Az Unióban egy további hasonló kezdeményezésként hozta létre az Európai Gyógyszerügynökség (European Medicines Agency – EMA) és az Európai Gyógyszerészeti Szabályozási Hálózat a Data Analysis and Real World Interrogation Network-ot (DARWIN EU®), amelynek célja a valós idejű és megbízható tényeken alapuló információ szolgáltatása az emberi felhasználású gyógyszerek használatáról, biztonságáról, hatékonyságáról [25]. A DARWIN EU® Adathálózatban jelenleg 16 ország vesz részt 30 adatpartnerrel, 39 adatforrással és több mint 181 millió beteg adataival úgy, hogy az adatok, például a Norvég Rákregiszter esetében, akár 1960-ig visszanyúlhatnak [41].

A DARWIN EU® Adathálózatához a Semmelweis Egyetem is csatlakozott, és ún. adatpartnerként vesz részt a határon átnyúló kutatási együttműködésekben. Jelenleg is több kutatás zajlik a DARWIN EU® kereté-



3. ábra | Az OHDSI-hálózaton megvalósuló kutatások tervezésének és kivitelezésének lépései (az ábra forrása: a szerzők saját szerkesztése az OHDSI tájékoztató alapján [36, 37])

OHDSI = megfigyelésen alapuló egészségügyi adattudományok és informatika, SQL = strukturált lekérdezési nyelv

ben, mint például a cystás fibrosisban szenvedő európai egyének jellemzése, amelynél a kutatás célja epidemiológiai bizonyítékok létrehozása az érintett páciensek klinikai jellemzőiről és monitorozásáról 2015 és 2024 között, 5 európai tagállam bevonásával [42]. További hasonló kezdeményezések például bizonyos ráktípusok előfordulásával, valamint gyermekkori magas vérnyomással kapcsolatos vizsgálatok [43]. Egy további kutatási példa a vényköteles opioidtartalmú gyógyszerek felhasználásának vizsgálatát célozta az amerikai egyesült államokbeli „opioidkrízis” kapcsán [44, 45]. Összességében a kutatás rámutatott arra, hogy az opioidtartalmú gyógyszerek használatának incidenciája vagy enyhén csökkent, vagy nem változott a vizsgált adatbázisokban, illetve az is megfigyelhető volt, hogy a COVID-19-járvány időszakában a nem daganatos elváltozás kezeléséhez kapcsolódó opioidhasználat jelentősen csökkent, 2022 után viszont az opioidhasználat mértéke visszatért legalább a járvány előtti időszak szintjéhez.

Megbeszélés

További adatmodellek

Az OMOP CDM csak egy azok közül a kezdeményezések közül, amelyek közös adatmodellben kezelik a különböző forrásokból származó egészségügyi adatokat. Egyes szerzők mélyrehatóan foglalkoztak az eltérő adatmodellek összehasonlításával, különösen az azok között fellelhető különbségekkel és hasonlóságokkal [46–48], míg mások, a kihívásokra válaszolva, a CDM-ek jövőbeli fejlődési irányait határozták meg, hangsúlyozva a mesterseges intelligencia térnyerését [49].

A különböző CDM-ek eltérő célok mentén jöttek létre. Míg az OMOP CDM – a legelső létrehozott CDM-ek egyike – egységes és összefüggő formátumot kívánt kialakítani, különösen a klinikai ellátásban hasz-

nált adatokra, addig a Sentinel CDM sokkal inkább az adminisztratív és kiegészítő adatokra helyezte a hangsúlyt, és célja a forgalomba hozatal utáni gyógyszerfelügyelet. Emellett a páciens által jelentett adatokat is tartalmazza. Végül, a The National Patient-Centered Clinical Research Network (PCORnet) célkitűzése az OMOP CDM-hez hasonlóan a klinikai együttműködések támogatása, ennek elérését viszont sokkal egyszerűbb modellben, szűk felhasználói kör számára valósítja meg [46]. A terminológiák számát tekintve az OMOP CDM rendelkezik a legszélesebb lefedettséggel, mivel számos nemzetközi kódszótárt magában foglal [48] (1. táblázat) [12, 46, 50, 51].

A CDM-ek összehasonlításával foglalkozó források és az 1. táblázatban bemutatott adatok alapján látható, hogy az OMOP CDM a legelterjedtebb, rendkívüli lehetőségeket nyújtva összevont kutatások megvalósítására.

További lehetőségek

Mindezt összevetve látható, hogy az OMOP CDM adatmodell képes a megfigyeléses kutatások támogatására úgy, hogy szabványos terminológiák használatával átjárhatóságot teremt a különböző adatbázisok között, illetve teljes tervezési és megvalósítási eszköztárat is biztosít a kutatók számára. A nyílt adatmodell előnye, hogy igény szerint bárki használhatja, a további nemzetközi kezdeményezések, mint például az OHDSI, az EHDEN vagy a DARWIN EU® által átfogó, több populációt lefedő kutatások is elvégezhetők, nagy erejű bizonyítékokat teremtve az orvosi szakma számára. A jövőbeli kutatási irányokat tekintve, a nagy mennyiségű adatokon alapuló vizsgálatok élveznek előnyt. Ebbe az irányba mutat az Európai Unió törekvése is az EHDS létrehozásával [8].

A hálózaton keresztül saját kezdeményezésű kutatási projektek is indíthatók és átfogó vizsgálatok végezhetők a részt vevő tagállamok által biztosított adatokkal. A hazai érdeklődők számára a Magyarországi Nemzeti OHDSI Csomópont biztosít elsősorban kapcsolódási lehetőséget. Szükség szerint iránymutatás kapható az OHDSI-hálózathoz való csatlakozással kapcsolatban, és informatikai támogatás is igénybe vehető.

Az OMOP-adatmodell használatán alapuló vizsgálatok előnyeit az alábbiak szerint összegezzük:

- nagy esetszám,
- valós környezetben gyűjtött, klinikai állapotra vonatkozó adatok,
- valós fejlesztések, együttműködések jönnek létre,
- nemzetközi együttműködésekben alapul,
- az eredményeket könnyebb publikálni,
- az adatok jobb minőségben állnak rendelkezésre,
- könnyen használható betegbiztonsági fejlesztések céljából.

Az OMOP CDM hazai adaptáció széles körű megvalósulása többek közt azért is lenne előnyös, mert így megteremthetővé válhat az egészségügyi informatikai rendszerek közötti átjárhatóság. Számos esetben talál-

1. táblázat | A három legnépszerűbb [46] CDM összehasonlítása (2025. október 8-i adatokkal). Források: [12, 50, 51]

Adtmodell	OMOP CDM	Sentinel CDM	PCORnet
Elterjedtség	Világszintű, 83 országban	Amerikai Egyesült Államok	Amerikai Egyesült Államok
Adatpartnerek száma	4294 együttműködő	13 adatpartner	78 adatpartner
A közösség tagjai	<ul style="list-style-type: none"> • Adatbirtokosok • Kutatók • Egészségügyi szolgáltatók • Páciensek és fogyasztók 	<ul style="list-style-type: none"> • Tudományos orvosi központok • Egészségügyi rendszerek • Egészségügyi biztosítók 	<ul style="list-style-type: none"> • Páciensek és gondozók • Klinikusok, klinikai szakemberek • Biztosítók • Döntéshozók • Mások

CDM = közös adatmodell; OMOP = Megfigyelésen Alapuló Orvosi Eredmények Társulása; PCORnet = Betegközpontú Klinikai Kutató-hálózat

kozhatunk azzal a jelenséggel, hogy két különböző informatikai rendszerből, azonos mutató ellenőrzése esetén, eltérő értékeket kapunk, vagy a rendszeres adatgyűjtések eredményeit nem tudjuk ágazati szinten, együttesen értékelni. Az egészségügyi rendszer adatainak elemzése olyan értéket jelent, amely nemzetgazdasági szinten is hasznosítható. A kutatás-fejlesztési tevékenység előmozdításával a bizonyítékalapú orvoslás is fejlődik, és új eljárásrendek, protokollok jöhetnek létre a megbízhatóság növeléséért. Az RWE átfogó célja, hogy betegszintű eredményeket és pozitív irányú változást érjünk el, de mindennek az alapköve a jó adatminőség, a megbízható vizsgálati módszerek és kivitelezés, így a szabványkövetésnek nemcsak az ellátás szintjén, hanem a legegyszerűbb szinten, az adatgyűjtés és -tárolás szintjén is meg kell jelennie.

Korlátok

Bármennyire törekszik is az OHDSI az adatbázisok és a modell teljességére, továbbra is előfordulnak olyan kutatási célkitűzések, amelyek vizsgálatára a bemutatott módszerek nem elegendők. Ezek közé tartozik a beavatkozások ok-okozati tényezőinek feltárása a placebohatással szemben, a vény nélkül kapható gyógyszerekre vonatkozó kutatások, azok a megfigyeléses vizsgálatok, amelyeknél az adatok elérhetősége a rögzítési gyakorlatból kifolyólag korlátozott, illetve a kezelések pontos eredményeinek nyomon követése [5]. Sok esetben az orvosszakmai igények mélyebbek, komplexebbek, mint amit a kódrendszerek használata a rutinszerű adatgyűjtés során lehetővé tesz, ezért folyamatos fejlesztési szemlélet szükséges.

A kutatási környezet versenyképességének fenntartása érdekében ugyanakkor figyelmet kell fordítani az OMOP CDM-et övező jövőbeli kihívásokra is. *Li és mtsai* [46] a génszekvenálás, a viselhető eszközök és a mesterséges intelligencia alkalmazásának elterjedésével kapcsolatban hangsúlyozzák a terület fejlődési szükségleteit [46]. Álláspontjuk szerint igény lesz arra, hogy további adatforrások is beépítésre kerüljenek a CDM-be, illetve a mesterséges intelligencia nyújtotta lehetőségek is az eszköztár részét képezzék [46].

Végül, fontos figyelembe venni, hogy az OMOP CDM adaptálása jelentős erőforrásokat igényel, amely mind az időráfordítás szükségességét, mind a szakmai kompetenciákat magában foglalja. Az adaptáció során megvalósítandó ETL-folyamat interdiszciplináris együttműködésekre támaszkodik, ehhez több szakterület bevonása szükséges, és jelentős koordinációt igényel. Az orvosi szakértelem nem választható el teljesen az informatikai folyamatoktól, hanem ellenkezőleg, előnyös, ha mind a bevont orvosi és egészségügyi szakembereknek van informatikai érdeklődésük, mind az informatikus kollégák nyitottak az egészségügyi terület sajátosságaira.

Következtetés

A standard terminológiák használata látható előnyöket jelent a megfigyelésalapú vizsgálatok megvalósítása során. Az OMOP-adatmodell a standard terminológiákat egységes keretbe szervezi, és további szabványokat határoz meg az adatok minőségére, illetve azok felhasználására. A standard adathasználat lehetővé teszi a nagy megbízhatóságú klinikai bizonyítékok, azaz RWE előállítását, ezzel segítve az egészségügyi ellátás fejlesztését, különös tekintettel a megbízhatóságra.

Az OMOP CDM hosszú távú fenntarthatósága azonban nemcsak közösségi érdek, hanem közös cél és feladat is kell hogy legyen. Az OHDSI-hálózat bővítése az adatmodell adaptációjával, további kutatási adatbázisok bevonásával tovább szélesíti az RWD-alapú kutatási lehetőségeket, nemcsak hazai, hanem nemzetközi szinten is jelentős értéket teremtve.

Ebben a cikkben az OMOP CDM megvalósításának informatikai vetületeit részleteiben nem tárgyaltuk, tekintettel a célközönség érdeklődésére és ismereteire. Ennek ellenére az ismertett információk elegendőek lehetnek ahhoz, hogy egy egészségügyi ellátó az érdeklődése szerint további lépéseket tegyen az OHDSI-hálózathoz való csatlakozás érdekében, például a Magyarországi Nemzeti OHDSI Csomóponttal [20] való együttműködés kialakítása által. Emellett az OHDSI nyílt forrásai további lehetőségeket biztosítanak arra, hogy az informatikai irányultságú szakemberek elmerüljenek a terület rejtelmeiben.

Anyagi támogatás: A közlemény publikálását az RRF-2.3.1-21-2022-00003 projektazonosítójú Nemzeti Kardiovaszkuláris Laboratórium tette lehetővé.

Szerzői munkamegosztás: Sz. O., M. Á., K. L. és B. Zs. részt vett a kutatás megtervezésében. Az irodalomkutatásban Sz. O., M. Á. és K. L. vett részt. Ezek alapján Sz. O. írta meg a dolgozat vázlatát, és állította elő a végleges kéziratot. A szakmai tartalom helytállóságához továbbá különösen hozzájárult B. O., K. E. és A. A. Valamennyi szerző elolvasta a kéziratot, és szakmai javaslatokkal egészítette ki azt. A közlemény végleges változatát minden szerző elolvasta és jóváhagyta.

Érdekltségek: A szerzőknek nincsenek érdekltségeik.

Irodalom

- [1] Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet* 2017; 390(10092): 415–423.
- [2] Kelemen É, Bella Zs, Erdélyi E, et al. Acute rhinosinusitis in children requiring hospitalization between 2016 and 2022 – retrospective analysis. [Hospitalizációt igénylő gyermekkori akut rhinosinuszitis 2016 és 2022 között klinikánkon – retrospektív elemzés.] *Orv Hetil.* 2024; 165: 747–753. [Hungarian]
- [3] Jánosi A, Ofner P, Andréka P. Clinical and scientific value of the Hungarian Myocardial Infarction Registry. [A Nemzeti Szív-

- farktusz Regiszter klinikai és tudományos értéke.] *Orv Hetil.* 2024; 165(24–25): 944–949. [Hungarian]
- [4] Ministry of Health. Decree No. 23/2002. (V. 9.) of the Ministry of Health on medical research involving humans, Section 20/A. [23/2002. (V. 9.) EüM rendelet az embereken végzett orvostudományi kutatásokról, 20/A §.] Available from: <https://net.jogtar.hu/jogszabaly?docid=a0200023.eum> [accessed: February 27, 2026]. [Hungarian]
- [5] Observational Health Data Sciences and Informatics. The book of OHDSI. Available from: <http://book.ohdsi.org/> [accessed: December 14, 2025].
- [6] Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence – what is it and what can it tell us? *N Engl J Med.* 2016; 375: 2293–2297.
- [7] Dang A. Real-world evidence: A primer. *Pharmaceut Med.* 2023; 37: 25–36.
- [8] European Parliament and the Council of the European Union. Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847 (Text with EEA relevance). *Official Journal of the European Union*; 2025. Article 51. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202500327 [accessed: December 14, 2025].
- [9] Reisinger SJ, Ryan PB, O'Hara DJ, et al. Development and evaluation of a common data model enabling active drug safety surveillance using disparate healthcare databases. *J Am Med Inform Assoc.* 2010; 17: 652–662.
- [10] Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med.* 2010; 153: 600–606.
- [11] Observational Health Data Science and Informatics. OMOP common data model. Available from: <https://ohdsi.github.io/CommonDataModel/> [accessed: December 14, 2025].
- [12] Observational Health Data Sciences and Informatics. Available from: <https://www.ohdsi.org> [accessed: December 14, 2025].
- [13] Observational Health Data Sciences and Informatics. Standardized data: The OMOP common data model. Available from: <https://www.ohdsi.org/data-standardization/> [accessed: December 14, 2025].
- [14] Observational Health Data Sciences and Informatics. OMOP CDM refresh process. Available from: <https://ohdsi.github.io/CommonDataModel/cdmRefreshProcess.html> [accessed: December 14, 2025].
- [15] European Health Data & Evidence Network. Available from: <https://www.ehden.eu> [accessed: December 14, 2025].
- [16] European Health Data & Evidence Network. Vision, mission & objectives. Available from: <https://www.ehden.eu/vision-and-mission/> [accessed: December 14, 2025].
- [17] European Health Data & Evidence Network. A federated network of data partners. Available from: <https://www.ehden.eu/datapartners/> [accessed: December 14, 2025].
- [18] SNOMED International. What is SNOMED CT? Available from: <https://www.snomed.org/what-is-snomed-ct> [accessed: December 14, 2025].
- [19] Mészáros Á, Kovács S, Héja T, et al. Mapping Hungarian procedure codes to SNOMED CT. *BMC Med Res Methodol.* 2023; 23: 240.
- [20] Observational Health Data Science and Informatics. OHDSI National Node Hungary. Available from: <https://ohdsi-europe.org/index.php/national-nodes/hungary> [accessed: December 14, 2025].
- [21] Boulahia C, Behja H, Boulahia Z. The multi-criteria evaluation of research efforts based on ETL software: from business intelligence approach to big data and semantic approaches. *Evol Intell.* 2024; 17: 2099–2124.
- [22] Observational Health Data Sciences and Informatics. Athena. Available from: <https://athena.ohdsi.org/search-terms/terms?query> [accessed: December 14, 2025].
- [23] Logical Observation Identifiers Names and Codes. The international standard for identifying health measurements, observations, and documents. Available from: <https://loinc.org> [accessed: December 14, 2024].
- [24] Observational Health Data Sciences and Informatics. Themis working group. Available from: <https://www.ohdsi.org/web/wiki/doku.php?id=projects:workgroups:themis> [accessed: December 14, 2025].
- [25] Data Analysis and Real World Interrogation Network (DARWIN EU®). Available from: <https://www.darwin-eu.org/index.php> [accessed: December 14, 2025].
- [26] Observational Health Data Sciences and Informatics. ATLAS. Available from: <https://atlas-demo.ohdsi.org/#/search> [accessed: December 14, 2024].
- [27] HL7 International. The standard. Available from: <https://blog.hl7.org/hl7-publishes-flir-release-4> [accessed: December 14, 2025].
- [28] Kent S, Burn E, Dawoud D, et al. Common problems, common data model solutions: evidence generation for health technology assessment. *PharmacoEconomics* 2021; 39: 275–285.
- [29] Observational Health Data Sciences and Informatics. How to determine the drug_exposure_end_date when it is not given explicitly in the data. Available from: https://ohdsi.github.io/Themis/drug_end_date_not_in_data.html [accessed: December 14, 2025].
- [30] Kahn MG, Callahan TJ, Barnard J, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. *EGEMS (Wash DC)* 2016; 4: 1244.
- [31] Observational Health Data Sciences and Informatics. Achilles. Available from: <https://github.com/ohdsi/achilles> [accessed: December 14, 2025].
- [32] Huser V, DeFalco FJ, Schuemie M, et al. Multisite evaluation of a data quality tool for patient-level clinical datasets. *EGEMS (Wash DC)* 2016; 4: 1239.
- [33] Observational Health Data Sciences and Informatics. DataQualityDashboard. Available from: <https://github.com/OHDSI/DataQualityDashboard> [accessed: December 14, 2025].
- [34] Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform.* 2015; 216: 574–578.
- [35] Observational Health Data Science and Informatics. Github repository. Available from: <https://github.com/OHDSI/> [accessed: December 14, 2025].
- [36] Falconer T, Chen R, Hripcsak G. Running an OHDSI network study. Available from: https://www.ohdsi.org/web/wiki/lib/exe/fetch.php?media=research:nci-ohdsi_instruction_manual2.pdf [accessed: December 14, 2025].
- [37] Martin B, Cai C, Golozar A, et al. OHDSI network study execution framework and templating. Available from: https://www.ohdsi.org/wp-content/uploads/2023/10/martin-ben_ohdsi-network-study-framework_2023symposium-Benjamin-Martin.pdf [accessed: December 14, 2025].
- [38] Hripcsak G, Ryan PB, Duke JD, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci USA* 2016; 113: 7329–7336.
- [39] Chen R, Ryan P, Natarajan K, et al. Treatment patterns for chronic comorbid conditions in patients with cancer using a large-scale observational data network. *JCO Clin Cancer Inform.* 2020; 4: 171–183.
- [40] Vashisht R, Jung K, Shah N. Learning effective treatment pathways for type-2 diabetes from a clinical data warehouse. *AMIA Annu Symp Proc.* 2017; 2016: 2036–2042.

- [41] Data Analysis and Real World Interrogation. Network (DARWIN EU®). The DARWIN EU® data network. Available from: <https://www.darwin-eu.org/index.php/data/data-network> [accessed: December 14, 2025].
- [42] European Medicines Agency. DARWIN EU® – RR Characterisation of individuals with cystic fibrosis in Europe. Available from: <https://catalogues.ema.europa.eu/node/4613/administrative-details> [accessed: December 14, 2025].
- [43] Data Analysis and Real World Interrogation Network (DARWIN EU®). Studies. Available from: <https://www.darwin-eu.org/index.php/studies> [accessed: December 14, 2025].
- [44] European Medicines Agency. DARWIN EU® – RR1 Drug utilisation study of prescription opioids. Available from: <https://catalogues.ema.europa.eu/node/4380/administrative-details> [accessed: December 14, 2025].
- [45] European Medicines Agency. DARWIN EU® – RR2 Drug utilisation study of prescription opioids. Available from: <https://catalogues.ema.europa.eu/node/4516/administrative-details> [accessed: December 14, 2025].
- [46] Li R, Romano JD, Chen Y, et al. Centralized and federated models for the analysis of clinical data. *Annu Rev Biomed Data Sci.* 2024; 7: 179–199.
- [47] Rosenbloom ST, Carroll RJ, Warner JL, et al. Representing knowledge consistently across health systems. *Yearb Med Inform.* 2017; 26: 139–147.
- [48] Wolfien M, Ahmadi N, Fitzer K, et al. Ten topics to get started in medical informatics research. *J Med Internet Res.* 2023; 25: e45948.
- [49] Painter JL, Ramcharran D, Bate A. Perspective review: will generative AI make common data models obsolete in future analyses of distributed data networks? *Ther Adv Drug Saf.* 2025; 16: 20420986251332743.
- [50] Sentinel. Who is involved. Available from: <https://www.sentinelinitiative.org/about/who-involved> [accessed: December 14, 2025].
- [51] The National Patient-Centered Clinical Research Network. PCORnet®: An introduction. 2025. Available from: https://pcornet.org/wp-content/uploads/2025/07/PCORnet_An_Introduction_11july2025_widescreen.pdf [accessed: December 14, 2025].

(Székely Orsolya,
Budapest, Baross u. 52., 1085
e-mail: szekely.orsolya99@phd.semmelweis.hu)

MEGHÍVÓ

Az Észak-budai Szent János Centrumkórház Tudományos és Oktatási Bizottsága

tisztelettel meghívja az érdeklődőket a következő tudományos ülésére.

Időpont: **2026. április 30. (csütörtök) 14 óra**

Helyszín: **Szent János Kórház Auditórium – 1125 Budapest, Diós árok 1–3.**

Üléselnök: *Prof. Dr. Hirschberg Andor*

Program:

<i>Szepessy Zsuzsanna dr.</i> (Szemészeti Osztály): Uveitisek etológiája új megközelítésből	15 perc
<i>Hargitai János dr.</i> (Szemészeti Osztály): Szaruhártya műtétek, a donor kiválasztás szempontjai	15 perc
<i>Székely Péter Pál dr.</i> (Idegsebészeti Osztály): Cervico-brachygalgia	15 perc
<i>Klenk Gusztáv dr.</i> (Fül-, Orr-, Gége-, Fej-Nyak és Szájsebészeti Osztály): Mivel foglalkozik az arc-állcsont és szájsebészet? Jelen, múlt, jövő	15 perc

Minden érdeklődőt szeretettel várunk!

A cikk a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>) feltételei szerint publikált Open Access közlemény, melynek szellemében a cikk bármilyen médiumban szabadon felhasználható, megosztható és újraközölhető, feltéve, hogy az eredeti szerző és a közlés helye, illetve a CC License linkje és az esetlegesen végrehajtott módosítások feltüntetésre kerülnek. (SID_1)