RWD197

Literature Review on the Use of Synthetic Data and AI Advances for Patient-Centered, Sustainable HEOR

Viktor Chirikov, PhD[1]; Sonja Kroep, PhD[2]

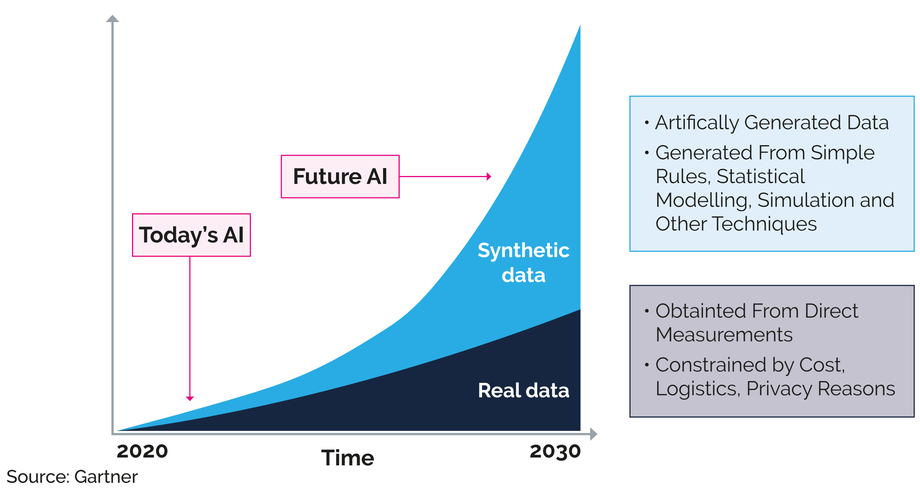
[1]OPEN Health HEOR & Market Access, New York, NY, USA Contact: viktorchirikov@openhealthgroup.com
[2]OPEN Health HEOR & Market Access, Rotterdam, the Netherlands

OPEN HEALTH

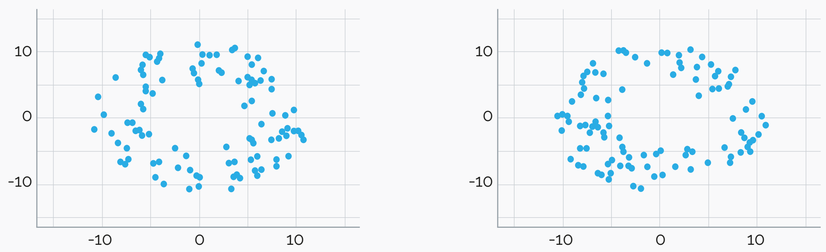
INTRODUCTION

- Long-term sustainability of healthcare systems, affordability, and patient-centricity are three of the main global healthcare drivers addressed in ISPOR Strategic Plan 2030.[1]
- Real-world data plays an instrumental part in achieving the strategic goal of sustainability, as it is increasingly being used to augment randomized controlled trials and accelerate drug approvals.[1]
- Recent advances in AI algorithms have increased the availability of real-world data as well as options for synthetic data, an artificial version of real data created algorithmically without compromising privacy.
- Synthetic data can be used as proxy for real patient data;[2,3] its use is expected to overshadow that of real data by 2030 (**Figure 1**).

Figure 1. Real data vs synthetic data and their forecasted use by 2030.



Source: Gartner



The synthetic data remains the structure of the original data but is not the same

OBJECTIVES

The primary objective of this study was to summarize how advances in AI and its applications to real-world data and synthetic data could provide avenues for patient-centered, sustainable HEOR aligned with the goals of the ISPOR Strategic Plan 2030.

METHODS

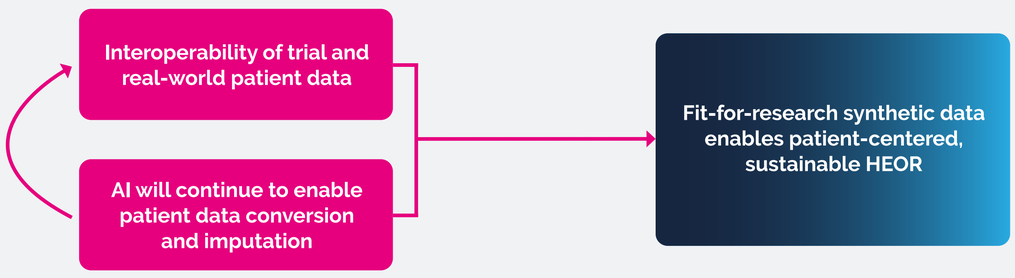
Data sources

- Structured query of Embase/Medline (January 1, 2022 through October 30, 2024) using the terms [("*synthetic data*" *OR* "*artificial intelligence*") *AND* ("*health economics*" *OR* "*outcomes research*")] was conducted.
- The literature query above was supplemented by unstructured review of the program and presentations content of ISPOR annual conferences 2022-2024 as well as healthcare technology conferences *HLTH 2022-2024 and Healthcare Information and Management Systems Society* 2022-2024.
- Review of the resulting publications and program content focused on four main topics:

i. AI advancements in generating patient data;
ii. AI methodological advancements in manipulating and analyzing patient data;
iii. The anticipated effect of the approaching patent cliff of many blockbuster therapeutics;
iv. The introduction of the Inflation Reduction Act in the US and it's influence on improving access and affordability of innovative health technologies.

CONCEPTUAL INTEGRATION OF RESULTS

Accelerated market access necessitates the use of rapid value evidence generation



Increased demand for external data information during rapid value evidence generation

RESULTS AND DISCUSSION

- The search terms resulted in 304 publications from 2022-2024 (N=78 in 2022, N=113 in 2023, and N=135 in 2024), which accounted for the majority of research work published on this topic to date (N=457).
- Elicitation of the topics presented across these publications, supplemented by review of the program content at the health conferences of interest, revealed the following surfacing themes:

Accelerated market access necessitates the use of rapid value evidence generation

- Market competition is anticipated to increase following the patent cliff of many blockbuster drug therapies between 2024-2030 as well as due to the stipulations of the Inflation Reduction Act, allowing Medicare to negotiate pricing with drug manufacturers in the US.
- Increased competition and market pressures will necessitate accelerated clinical development and market access underpinned by rapid value evidence generation.

Increased demand for external data sources during evidence generation

- The use of digital twins technology[4] and dynamic Bayesian borrowing partial information from external data sources will be more prominent in order to conduct well-powered, yet more efficient, clinical trials requiring reduced sample size.

AI will continue to enable patient data conversion and imputation

- AI will continue to enable the conversion of semi-structured or unstructured data into structured data as well as impute missing data fields.
- This will increase the availability of data fit for research.

Interoperability of trial and real-world patient data

- The increase in fit-for-research data will allow for the pooling and interoperability of multiple disparate sources of data, from clinical trial data to real-world patient data and vice versa.

Fit-for-research synthetic data minimizes privacy concerns and enables patient-centered, sustainable HEOR

- Interoperability of patient data will enable the creation of fit-for-research synthetic structured data that could be used, alongside real data, to augment the efficiency and speed of clinical trials as well as observational research.
- Additionally, the use of synthetic data will enable economic evaluations when available comparative effectiveness evidence is otherwise scarce.
- Synthetic data minimizes privacy concerns of using patient-identifiable data and, when used as part of rapid value evidence generation, helps pave the way for patient-centered, sustainable health economics and outcomes research.

REFERENCES

1. ISPOR—The Professional Society for Health Economics and Outcomes Research. ISPOR Strategic Plan 2030. An ISPOR Strategic Report. July 2024

2. Azizi Z, Zheng C, Mosquera L, Pilote L, El Emam K. Can synthetic data be a proxy for real clinical trial data? A validation study. BMJ open. 2021 Apr 1;11(4):e043497.

3. El Emam K, Mosquera L, Fang X, El-Hussuna A. An evaluation of the replicability of analyses using synthetic health data. Scientific Reports. 2024 Mar 24;14(1):6978.

4. Vidovszky AA, Fisher CK, Loukianov AD, Smith AM, Tramel EW, Walsh JR, Ross JL. Increasing acceptance of AI generated digital twins through clinical trial applications. Clinical and Translational Science. 2024 Jul;17(7):e13897.

Presented at: ISPOR Europe 2024