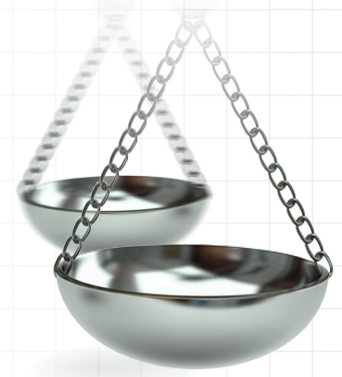




R WE ready for reimbursement? A round-up of developments in real-world evidence relating to health technology assessment: part 22

Paul Arora^{1,2} & Sreeram V Ramagopalan^{*,3}

¹Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada

²Inka Health, 100 King St W. Suite 1600. Toronto, ON M5X 1G5, Canada

³Centre for Pharmaceutical Medicine Research, King's College London, London, SE19NH, UK

*Author for correspondence: sreeram@ramagopalan.net

In this update, we examine the growing potential of synthetic data in addressing data access challenges for health technology assessment. We explore studies evaluating synthetic data as external control arms in clinical trials and the application of synthetic data in health economic evaluation. Additionally, we review real-world evidence on the clinical impact of formulary restrictions in the US.

First draft submitted: 5 September 2025; Accepted for publication: 17 September 2025; Published online: 16 October 2025

Keywords: clinical trials • external control arms • formulary restrictions • health technology assessment • real-world data • real-world evidence • simulation study • synthetic data

The challenge of data access continues to be a significant barrier in the generation of real-world evidence (RWE). Synthetic data has emerged as a promising solution to address privacy concerns while maintaining analytical value for research and regulatory decision-making [1]. We have previously discussed a demonstration study using synthetic data for comparative effectiveness research [1]. Recent publications have evaluated the potential for synthetic data use in additional contexts.

Elvatun *et al.* investigated the feasibility of using synthetic data as external control arms, addressing the often-lengthy application and approval processes required to access observational healthcare data, especially from the Nordic registers [2]. To create a synthetic external control arm, artificial patient records are generated by training algorithms on original observational data. These algorithms learn the statistical patterns and relationships within the dataset, then produce new synthetic records that maintain the same distributional properties without directly copying any real patient information. This approach creates datasets that preserve the clinical and statistical characteristics necessary for research while protecting individual privacy. Algorithms may vary greatly in terms of their modeling approach, making some better suited for particular data types and requirements. Elvatun and colleagues therefore evaluated six different synthetic data generation algorithms: Conditional Tabular Generative Adversarial Network (CTGAN), Survival CTGAN (specialized for survival data), Differentially Private GAN (DP-GAN), Private Aggregation of Teacher Ensembles GAN (PATE-GAN), PrivBayes (a Bayesian network-based approach with differential privacy) and Uniform (a baseline generator using random sampling). A key methodological contribution of Elvatun *et al.* was the introduction of a 'reversible data generalization' procedure specifically designed to address the challenges of generating synthetic data from small datasets with high-dimensional characteristics, making individual records more susceptible to re-identification. The generalization process reduces input data cardinality before synthesis by binning continuous variables like age and survival time, then reverses this process post generation of synthetic data to restore the original data structure while maintaining privacy protection. The original real-world data (RWD) used was advanced non-small-cell lung cancer data from Norwegian health registries, comparing 645 patients who received either chemotherapy (control arm; $n = 397$) or pembrolizumab monotherapy (active arm; $n = 248$). The approach was evaluated for three dimensions: data resemblance, utility and privacy. The results demonstrated that the reversible data generalization mechanism provided substantial

benefits across all evaluated algorithms, appearing to enhance their ability to capture genuine statistical patterns while reducing the risk of memorizing specific patient information. In terms of performance, the overall ranking identified PrivBayes, CTGAN and PATE-GAN as the most promising algorithms.

The external control arm analysis showed that the top-performing algorithms (PrivBayes and CTGAN) produced survival curves that closely aligned with original data for approximately 25 months, with only minor deviations in later time periods. Hazard ratio estimates remained consistent with the original data, demonstrating that the same clinical insights could be drawn from studies using synthetic external control arms. For manufacturers, these findings suggest that synthetic external control arms could serve as viable alternatives to observational data in drug development.

While synthetic external control arms demonstrate promise for comparative effectiveness research, manufacturers must also consider their utility in broader health economic evaluation that form the cornerstone of health technology assessment (HTA) submissions. Health economic evaluations typically rely on individual patient data (IPD) from clinical trials and patient registries to assess treatment effectiveness, costs and quality of life outcomes. However, accessing such data may similarly face barriers. When individual patient data are unavailable, researchers often resort to aggregate data or digitized summary statistics, which lack the granularity to capture patient-level variation, correlations between parameters and the influence of patient characteristics on outcomes. Synthetic data generation offers a potential solution. Van der Linden *et al.* designed a simulation study using a hypothetical disease called 'shame' to test synthetic data performance in economic evaluations [3]. They created an original dataset (Dorg) containing 1000 patients in the base case, representing results from a fictitious clinical trial comparing 'shamectomy' versus usual care for preventing shame. This dataset included clinical outcomes (time to disease progression and death), costs, utilities and patient characteristics generated using realistic statistical distributions and correlations. From Dorg, they generated 500 synthetic datasets (Dsyn) using the *synthpop* package in R, which creates variables by drawing from conditional distributions fitted to the original data. Identical model-based economic evaluations using both original and synthetic data was performed, comparing incremental cost-effectiveness ratios (ICERs), quality-adjusted life years (QALYs) and reimbursement decisions based on a willingness-to-pay threshold of €20,000/QALY. The analysis encompassed multiple scenarios testing different conditions: varying sizes of original datasets (50, 500, 1000 and 10,000 patients), inflated synthetic datasets, different synthesis orders, proper versus simple synthesis methods, and missing data handling approaches. In the base case with 1000 patients, the original data yielded an ICER of €25,848/QALY, while synthetic data produced a mean ICER of €25,857/QALY with a 95% confidence interval of €16,776-60,021. Critically, while the original data led to a negative reimbursement decision (ICER above threshold), 15% of synthetic datasets resulted in positive reimbursement decisions. Dataset size emerged as the most important factor determining synthetic data reliability. With smaller original datasets (50 and 500 patients), ICER ranges expanded dramatically to -€151,542 and -€669,717, respectively. For the smallest original dataset ($n = 50$), 17% of synthetic datasets could not be analyzed due to lack of progression or death events. Larger synthetic datasets ($n = 10,000$) showed reduced variability but potentially misleading precision if the original dataset was small. Different synthesis approaches showed minimal impact on outcomes, while missing data scenarios suggested that imputation during synthesis outperformed pre-synthesis imputation approaches. Based on their findings, the authors propose a stepwise approach for using synthetic data in health economic evaluations. They recommend considering synthetic data only when original datasets contain more than 1000 records, as smaller datasets may lead to substantial deviations. For data owners generating synthetic data, they strongly recommend publishing multiple synthetic datasets (at least 100) with the same size as the original data, allowing analysts to understand uncertainty introduced by the synthesis process. This research demonstrates that while synthetic data shows promise in economic evaluation, caution is warranted, particularly with smaller datasets.

Synthetic data could address RWD access challenges, particularly in rare diseases or where data has strict privacy regulations. Both studies discussed underscore the need for rigorous methodological standards and transparent reporting frameworks when using synthetic data. The field currently lacks specific guidelines for synthetic data use in HTA submissions. The studies mentioned above provide case studies, but HTA bodies should consider developing explicit criteria for synthetic data use, including minimum dataset requirements, sensitivity analyses and transparency reporting on synthesis methodologies. The guidance proposed by van der Linden *et al.* provides a practical framework, but HTA bodies need to signal their willingness to accept synthetic data to encourage appropriate investment in synthetic data methodologies. Ultimately, while synthetic data represents a promising

solution to RWD access barriers, its integration into HTA decision-making requires collaborative development of standards that ultimately enables patient access to medicines.

Beyond synthetic data applications, RWE continues to reveal important insights about how payer policies directly impact patient outcomes, as demonstrated by recent research examining formulary restrictions in multiple sclerosis (MS). US pharmacy benefit managers employ formulary restrictions and utilization management to secure manufacturer rebates and control drug spending. While health plans and pharmacy benefit managers argue these tools promote cost-effective prescribing, evidence indicates they may be overused, creating administrative burdens and interfering with clinical decision-making [4]. Blaylock and colleagues sought to test the appropriateness of formulary restrictions for MS, where treatment selection often depends on individual patient factors and clinical judgment [5,6]. While formulary exclusions may be suitable for drug classes with similar treatment effects, MS disease modifying therapies (DMTs) vary significantly in mechanism of action, route of administration, tolerability and efficacy profiles. Blaylock *et al.* analyzed Medicare administrative data from 2018 to 2022, examining both stand-alone prescription drug plans (PDPs) and Medicare Advantage Prescription Drug plans (MA-PDs). The study included 84,870 unique beneficiaries (50,162 in PDPs and 34,708 in MA-PDs) with relapsing-remitting MS who had used MS DMTs during a baseline period and remained in the same plan for at least five quarters. The cohort was followed using a rolling quarters design where the first four quarters served as baseline and the fifth quarter as follow-up. Formulary breadth was categorized as low or high coverage based on whether the four-quarter moving average of MS DMT coverage was below or above the median for each plan type and quarter. Coverage of 15 oral and injectable MS DMTs was examined. The primary outcome was MS relapse episodes during follow-up, identified through inpatient or outpatient treatment claims using validated algorithms adapted for administrative data. The study revealed extensive formulary restrictions across both plan types, with coverage becoming more restrictive over time. In 2022, more than 50% of plans excluded 11 of 15 DMTs in PDPs and nine of 15 DMTs in MA-PDs. After adjusting for patient and plan characteristics in multivariable analyses, high-coverage plans were associated with 6–7% lower odds of MS relapse in PDPs and 8–12% lower odds in MA-PDs. The findings challenge the assumption that formulary restrictions for MS DMTs serve only to promote cost-effective prescribing without harming patient outcomes. The 6–12% increase in relapse odds associated with restrictive formularies represents a clinically meaningful difference, as MS relapses are associated with disability progression, reduced quality of life and increased healthcare utilization [6]. There were several limitations to the analysis, but nevertheless the study provides compelling evidence that formulary restrictions on MS DMTs are associated with worse clinical outcomes in Medicare beneficiaries. The findings suggest that the current approach of using broad exclusions to manage costs may be inappropriate for heterogeneous conditions like MS, where treatment individualization is crucial for optimal outcomes. The study supports calls for more nuanced formulary management that considers both cost containment and patient benefit. For manufacturers, these findings highlight the strategic value of RWE in demonstrating the clinical consequences of payer policies. By generating robust RWE that quantify how formulary restrictions, prior authorization requirements, or step therapy protocols affect patient outcomes manufacturers can build compelling cases for broader access. Such evidence becomes particularly powerful when it demonstrates not only clinical harm from restrictions but also downstream economic consequences, including increased healthcare utilization and disability progression costs that ultimately burden both payers and patients.

Financial disclosure

Author SV Ramagopalan has received an honorarium from Becaris Publishing for the contribution of this work. The authors have received no other financial and/or material support for this research or the creation of this work apart from that disclosed.

Competing interests disclosure

The authors have no competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Writing disclosure

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. Arora P, Ramagopalan SV. R WE ready for reimbursement? A round up of developments in real-world evidence relating to health technology assessment: part 18. *J. Comp. Eff. Res.* 14(4), e250014 (2025).
2. Elvatun S, Knoors D, Brant S, Jonasson C, Nygård JF. Synthetic data as external control arms in scarce single-arm clinical trials. *PLOS Digit. Health* 4(1), e0000581 (2025).
3. van der Linden N, Pouwels XGLV, Jahn B, Siebert U, Koffijberg H. Who needs real data anyway? Exploring the use of synthetic data in economic evaluations of health interventions. *Value Health* S1098–3015(25) 02413-1 (2025).
4. Busis NA, Khokhar B, Callaghan BC. Streamlining prior authorization to improve care. *JAMA Neurol.* 81(1), 5–6 (2024).
5. Blaylock B, Van Nuys K, Joyce G. Formulary restrictions and relapse episodes in persons with relapsing-remitting multiple sclerosis. *JAMA Netw. Open.* 8(8), e2525155 (2025).
6. Jakimovski D, Bittner S, Zivadinov R *et al.* Multiple sclerosis. *Lancet* 403(10422), 183–202 (2024).