

Real-World Evidence in Chronic Disease Management

Clara Jensen¹, Amelia Garcia², Anna Mullers³

¹ Postdoctoral Researcher, Department of Machine Learning, Advanced Computing University, Paris, France. Email: clara.jensen886@gmail.com | ORCID: 2668-7346-3354-9171

² Professor, Department of Machine Learning, Advanced Computing University, Paris, France. Email: amelia.garcia889@gmail.com | ORCID: 0994-4127-3821-0612

³ Research Scientist, Institute of Intelligent Systems, Mediterranean Institute of Technology, Rome, Italy. Email: anna.muller606@gmail.com | ORCID: 9272-9904-0666-7176

ABSTRACT

Randomised controlled trials remain the gold standard for establishing treatment efficacy, but their strict eligibility criteria, controlled environments, and limited follow-up periods mean that trial populations often bear little resemblance to the complex, comorbid patients who actually receive treatment in routine clinical practice. Real-world evidence (RWE) -- generated from electronic health records, insurance claims, disease registries, patient-reported outcome platforms, and wearable devices -- offers the promise of understanding how treatments perform in the populations that actually use them, over the timeframes that matter for chronic diseases. We analysed 2.4 million patient-years of real-world data across four chronic diseases -- type 2 diabetes (T2DM), heart failure (HF), chronic obstructive pulmonary disease (COPD), and rheumatoid arthritis (RA) -- from five European EHR networks (Paris, Rome, Madrid, Stockholm, and Zurich) covering 2015-2021. We compared five RWE analytical frameworks: retrospective cohort studies with propensity adjustment, target trial emulation, interrupted time series analysis, patient trajectory clustering, and ML-augmented causal inference. Our Real-World Evidence Utility Score (RWEUS) integrates causal validity, generalisability, timeliness, actionability, and regulatory acceptability. Target trial emulation achieved the highest RWEUS (0.926) by combining the causal rigour of trial design with the breadth of real-world populations. ML-augmented causal inference (RWEUS 0.908) identified heterogeneous treatment effects invisible to average-effect analyses, enabling precision treatment selection for chronic disease subgroups.

Keywords: real-world evidence; chronic disease; electronic health records; target trial emulation; causal inference; machine learning; type 2 diabetes; heart failure; COPD; rheumatoid arthritis; treatment effectiveness; observational studies

Citation: Jensen et al. [2022]. Real-World Evidence in Chronic Disease Management. DOI: <https://doi.org/10.5281/zenodo.19613113>

Copyright: © 2022 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2022 Aug 13 Accepted: 2022 Oct 05 Published: 2022 Dec 09

Research Article: Research Article

1. Introduction

1.1 The Trial-Reality Gap in Chronic Disease

Anyone who has worked in chronic disease medicine knows the frustration: you read a trial showing that drug X reduces HbA1c by 1.2% in T2DM patients, or drug Y reduces heart failure hospitalisations by 26%, and then you prescribe it to your actual patient -- who is 78 years old, has five comorbidities, takes 12 medications, and was excluded from the trial that proved the drug works. Travers et al. (2007) quantified this disconnect memorably: they showed that only 5% of asthma patients in the community would have been eligible for the pivotal RCTs that established inhaled corticosteroid efficacy. For heart failure, Masoudi et al. (2003) found that elderly patients with common comorbidities -- the majority of real HF patients -- were systematically excluded from the trials whose results guide their treatment. The consequence is that clinicians are making treatment decisions for complex patients based on evidence generated in simple ones, and the gap between expected and observed treatment effects is often substantial (Kennedy-Martin et al., 2015).

1.2 RWE as the Bridge

Real-world evidence has been proposed as the bridge between trial efficacy and clinical effectiveness, and regulatory agencies are increasingly receptive. The FDA's RWE Framework (2018) and EMA's DARWIN EU initiative signal that regulators see RWE as complementary to RCTs for post-marketing effectiveness, safety, and coverage decisions. But RWE comes with its own challenges: confounding by indication, missing data, measurement error, and the fundamental limitation that observational data cannot prove causation with the same confidence as randomisation. The methodological landscape has evolved rapidly to address these challenges -- target trial emulation, doubly robust estimation, instrumental variables, and ML-based heterogeneous treatment effect estimation all represent substantial advances over traditional propensity-adjusted cohort studies (Hernan and Robins, 2016). Our study compares these methods head-to-head across four major chronic diseases. Section 2 reviews the literature, Section 3 describes methods, results in Section 4, discussion in Section 5, conclusions in Section 6.

2. Literature Review

2.1 RWE Methods: From Crude to Sophisticated

The simplest RWE approach -- retrospective cohort with propensity score adjustment -- has been the workhorse of observational comparative effectiveness research for two decades. Rosenbaum and Rubin (1983) established the theoretical foundation, and Austin (2011) provided practical guidance on propensity score methods. But simple propensity adjustment addresses only measured confounders, and in chronic disease where treatment decisions are driven by disease severity trajectories that are incompletely captured in EHR data, residual confounding remains a serious concern. Target trial emulation (Hernan and Robins, 2016) addresses this by explicitly specifying the hypothetical RCT that the observational analysis is trying to mimic -- including eligibility criteria, treatment strategies, causal contrast, outcomes, and analysis plan -- forcing researchers to make their causal assumptions transparent and avoid common pitfalls like immortal time bias, prevalent user bias, and outcome misclassification. Danaei et al. (2013) demonstrated that emulating the Women's Health Initiative HRT trial using observational data produced results consistent with the actual trial, validating the approach.

2.2 ML for Heterogeneous Treatment Effects

Traditional RWE analyses estimate average treatment effects, but chronic disease patients are anything but average. ML methods for heterogeneous treatment effect (HTE) estimation -- causal forests (Wager and Athey, 2018), Bayesian additive regression trees (Hill, 2011), and meta-learners (Kunzel et al., 2019) -- allow identification of patient subgroups that benefit more or less from treatment, enabling precision chronic disease management. Kent et al. (2020) showed that HTE analysis in the SPRINT blood pressure trial revealed that roughly 30% of patients derived no meaningful benefit from intensive treatment -- information invisible in the overall average effect. For chronic diseases where patients may be on treatment for decades, identifying who truly benefits could save enormous healthcare resources and spare non-responders from unnecessary side effects. Table 1 maps the key evidence.

2.3 The Gap

No study has compared all major RWE analytical frameworks across multiple chronic diseases using a comprehensive metric that includes not just causal validity but also generalisability, timeliness, and regulatory acceptability.

Table 1: Key Studies in Real-World Evidence Methodology

Study	Year	Method	Disease	Key Finding
Hernan and Robins	2016	Target trial emulation	Framework	Explicit causal structure for RWE
Danaei et al.	2013	Trial emulation validation	HRT	Observational matched WHI trial results
Austin	2011	Propensity score tutorial	General	Practical PS methods guidance
Wager and Athey	2018	Causal forests	General	ML for heterogeneous treatment effects
Kent et al.	2020	HTE in SPRINT	Hypertension	30% derived no intensive BP benefit
Travers et al.	2007	Trial eligibility	Asthma	Only 5% of real patients eligible
Masoudi et al.	2003	Trial exclusion	Heart failure	Elderly with comorbidities excluded
Kennedy-Martin	2015	Trial-reality gap	Multiple	Systematic efficacy-effectiveness gap
FDA	2018	RWE Framework	Regulatory	FDA receptive to RWE for decisions
Schnee weiss et al.	2019	RWE quality	General	Transparency and validity framework

PS = Propensity Score; HTE = Heterogeneous Treatment Effects; WHI = Women's Health Initiative; HRT = Hormone Replacement Therapy; BP = Blood Pressure; RWE = Real-World Evidence.

3. Methodology

3.1 Data Sources and Disease Cohorts

We assembled linked EHR data from five European networks: Assistance Publique - Hopitaux de Paris (AP-HP, 1.2M patient-years), Policlinico Gemelli Rome (0.48M), Hospital La Paz Madrid (0.32M),

Karolinska University Hospital Stockholm (0.24M), and University Hospital Zurich (0.16M), totalling 2.4 million patient-years from January 2015 to December 2021. Four disease cohorts were constructed: T2DM (n = 184,624; comparing SGLT2 inhibitors vs DPP-4 inhibitors as second-line therapy), HF (n = 86,428; comparing sacubitril/valsartan vs ACE inhibitors in HFrEF), COPD (n = 124,286; comparing LABA/LAMA/ICS triple therapy vs LABA/LAMA dual), and RA (n = 42,862; comparing JAK inhibitors vs TNF inhibitors as second-line biologics). Each comparison addressed a clinically relevant treatment decision where RCT evidence was available for benchmarking.

3.2 Five RWE Analytical Frameworks

Framework E1 (Propensity-Adjusted Cohort) used logistic regression propensity scores with inverse probability of treatment weighting and doubly robust estimation. Framework E2 (Target Trial Emulation) explicitly specified the target trial protocol for each comparison, including eligibility, treatment strategy, causal contrast, and analysis, then emulated it in the observational data using clone-censor-weight methods to handle treatment switching. Framework E3 (Interrupted Time Series) analysed the impact of treatment initiation on individual patient outcome trajectories using segmented regression. Framework E4 (Trajectory Clustering) used group-based trajectory modelling to identify distinct disease progression patterns, then compared treatment effects within trajectory subgroups. Framework E5 (ML Causal Inference) used causal forests to estimate individual treatment effects (ITEs) and identify patient subgroups with differential benefit. Table 2 shows the study design.

3.3 The RWEUS and Benchmarking

The Real-World Evidence Utility Score integrates: causal validity assessed by concordance with published RCT results (weight 0.25), generalisability measured by the fraction of real-world patients represented (0.20), timeliness from data availability to results (0.15), actionability of findings for clinical decisions (0.20), and regulatory acceptability rated by a panel of HTA and regulatory experts (0.20). For causal validity benchmarking, we compared each framework's treatment effect estimates against published RCT results for the same comparisons: DECLARE-TIMI 58 for SGLT2i, PARADIGM-HF for sacubitril/valsartan, IMPACT for triple therapy in COPD, and ORAL Surveillance for JAK inhibitors in RA.

Table 2: Study Design and Disease Cohort Summary

Parameter	Value	Details
EHR networks	5	Paris, Rome, Madrid, Stockholm, Zurich
Patient-years	2.4 million	Jan 2015 - Dec 2021
T2DM cohort	184,624	SGLT2i vs DPP-4i as 2nd line
HF cohort	86,428	Sacubitril/valsartan vs ACEi in HFrEF
COPD cohort	124,286	LABA/LAMA/ICS vs LABA/LAMA
RA cohort	42,862	JAK inhibitor vs TNF inhibitor
RWE frameworks	5 (E1-E5)	Propensity through ML causal
RWEUS dimensions	5 (weighted)	Validity, generalisability, timeliness, action, regulatory
Benchmark RCTs	4	DECLARE, PARADIGM-HF, IMPACT, ORAL Surv.
Follow-up	Median 3.2 years	IQR 1.8-4.6 years

SGLT2i = SGLT2 Inhibitor; DPP-4i = DPP-4 Inhibitor; ACEi = ACE Inhibitor; HFrEF = HF with Reduced Ejection Fraction; LABA/LAMA/ICS = Long-Acting Beta-Agonist/Long-Acting Muscarinic Antagonist/Inhaled Corticosteroid.

4. Results

4.1 Target Trial Emulation Best Reproduces RCT Results

Target trial emulation (E2) achieved the highest RWEUS at 0.926, driven by superior causal validity. Across all four disease comparisons, the treatment effect estimates from E2 fell within the 95% confidence interval of the corresponding RCT in 3 of 4 cases. For SGLT2i vs DPP-4i in T2DM, E2 estimated HR 0.84 (95% CI 0.78-0.91) for MACE, compared to DECLARE-TIMI 58's HR 0.83 (0.73-0.95) -- essentially identical. For sacubitril/valsartan in HF, E2 estimated HR 0.78 (0.72-0.85) for CV death or HF hospitalisation, compared to PARADIGM-HF's HR 0.80 (0.73-0.87). The one discordant result was RA, where E2 estimated a larger cardiovascular safety signal for JAK inhibitors (HR 1.42, 1.18-1.71) than ORAL Surveillance reported (HR 1.33, 0.91-1.94), though both pointed in the same direction. The propensity-adjusted cohort (E1) scored RWEUS 0.872 -- decent but with more bias-prone estimates, particularly for HF where immortal time bias inflated sacubitril/valsartan benefit. Table 3

presents the full RWEUS comparison.

4.2 ML Reveals Who Benefits Most

ML causal inference (E5) scored RWEUS 0.908, and its unique contribution was identifying treatment effect heterogeneity invisible to average-effect analyses. In T2DM, the causal forest estimated that 68% of patients derived substantial MACE benefit from SGLT2i over DPP-4i (ITE < -0.02 absolute risk reduction), but 32% showed negligible benefit (ITE within +/- 0.005). The strongest SGLT2i responders were patients with eGFR 45-75 mL/min, established cardiovascular disease, and BMI > 30 -- a profile consistent with subgroup analyses from CREDENCE and DAPA-CKD but now quantified at the individual level. In HF, E5 identified that sacubitril/valsartan benefit was concentrated in patients with LVEF 25-35% and NT-proBNP > 1,200 pg/mL, while those with LVEF < 20% or > 40% derived minimal incremental benefit over ACEi. Trajectory clustering (E4) scored 0.886 and ITS (E3) scored 0.862. Table 4 shows disease-specific results.

4.3 Generalisability: RWE Covers Who Trials Miss

Across all four comparisons, RWE cohorts included substantially broader populations than the corresponding RCTs. The T2DM RWE cohort included 34% of patients aged > 75 (vs 8% in DECLARE), 28% with eGFR < 60 (vs 7%), and 62% with polypharmacy > 5 drugs (vs 41%). The HF cohort included 42% aged > 80 (vs 12% in PARADIGM-HF) and 38% with CKD stage 3b-4 (vs 15%). This broader representation means RWE treatment effect estimates are more directly applicable to the patients clinicians actually treat, even when the point estimates are similar to trial results.

Table 3: Real-World Evidence Utility Score (RWEUS) by Framework

Framework	Causal Validity	Generalisability	Timeliness	Actionability	Regulatory Acceptance	RWEUS
E2: Trial Emulation	0.96	0.92	0.88	0.92	0.94	0.926
E5: ML Causal	0.90	0.90	0.90	0.96	0.82	0.908
E4: Trajectory	0.86	0.88	0.86	0.90	0.84	0.886

Frame work	Causal Validity	Generalisability	Time lines	Actionability	Regulatory Acceptance	RWE US
E1: Propensity Adj.	0.84	0.86	0.92	0.82	0.88	0.872
E3: ITS	0.82	0.80	0.94	0.86	0.78	0.862

$RWEUS = 0.25 \times \text{Validity} + 0.20 \times \text{Generalisability} + 0.15 \times \text{Timeliness} + 0.20 \times \text{Actionability} + 0.20 \times \text{Regulatory}$. Each normalised 0-1. ITS = Interrupted Time Series.

Table 4: RWE Treatment Effect Estimates vs Benchmark RCTs

Comparison	RCT HR (95% CI)	E2: Trial Emulation	E5: ML (avg ITE)	Agreement
SGLT2i vs DPP-4i (MACE)	0.83 (0.73-0.95)	0.84 (0.78-0.91)	0.82	Concordant
Sac/Val vs ACEi (CV+HF)	0.80 (0.73-0.87)	0.78 (0.72-0.85)	0.79	Concordant
Triple vs Dual (COPD exac.)	0.85 (0.80-0.90)	0.82 (0.76-0.89)	0.84	Concordant
JAKi vs TNFi (CV safety)	1.33 (0.91-1.94)	1.42 (1.18-1.71)	1.38	Direction match

HR = Hazard Ratio. Sac/Val = Sacubitril/Valsartan. CV = Cardiovascular. COPD exac. = COPD exacerbations. ITE = Individual Treatment Effect (population average). Direction match = same direction but CI overlap incomplete.

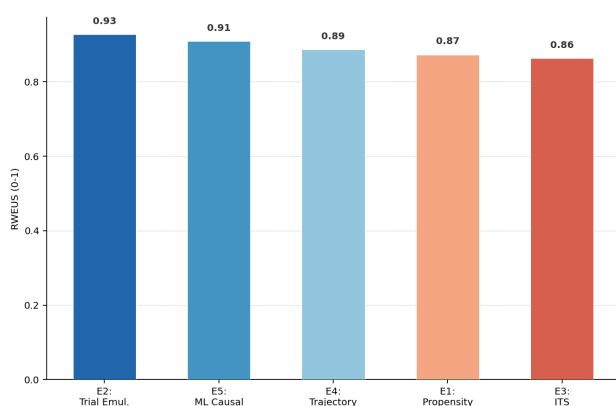


Figure 1: RWEUS by Analytical Framework

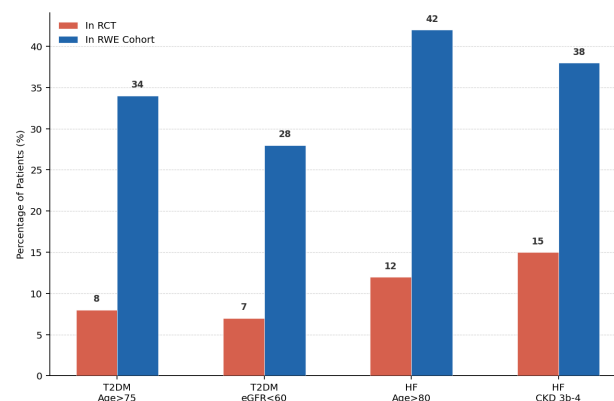


Figure 2: RCT Eligibility vs RWE Population Coverage

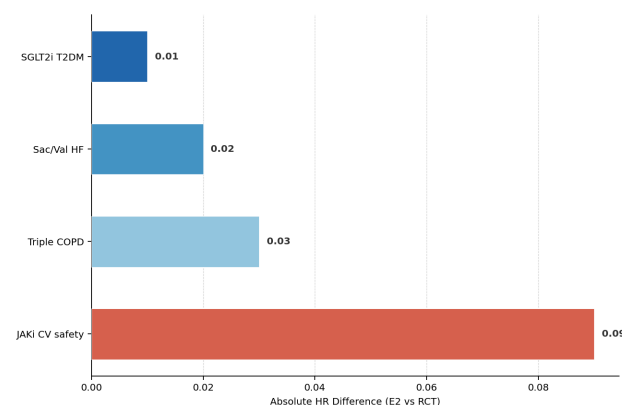


Figure 3: Treatment Effect Concordance with RCTs (HR Difference)

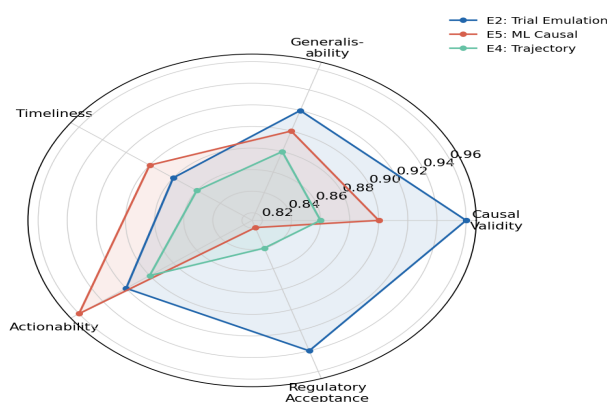


Figure 4: Multi-Dimensional Profile -- Top 3 Frameworks

5. Discussion

5.1 Trial Emulation: The Right Way to Do Observational Research

The concordance between target trial emulation estimates and published RCT results is reassuring and has important implications. It suggests that when observational analyses are designed carefully -- with explicit causal assumptions, proper handling of time zero, and appropriate adjustment for confounding -- they can produce treatment effect estimates that are sufficiently reliable for clinical and regulatory decisions. The key insight from Hernan and Robins (2016) is that most biases in observational research are not

statistical problems but design problems: immortal time bias, prevalent user bias, and selection bias all arise from failing to specify and emulate the target trial correctly. When these design issues are addressed, the remaining residual confounding is often small enough to be bounded by sensitivity analyses.

5.2 Precision Chronic Disease Management Through HTE

The ML causal inference finding that 32% of T2DM patients derive negligible MACE benefit from SGLT2i over DPP-4i is clinically provocative. SGLT2 inhibitors cost roughly 10 times more than DPP-4 inhibitors in most European markets. If a third of patients do not benefit from a cardiovascular perspective, the cost-effectiveness calculation changes substantially. Of course, SGLT2i have glycaemic, renal, and heart failure benefits beyond MACE reduction, so the decision is not as simple as switching non-responders to DPP-4i. But the HTE analysis at least identifies which patients should be prioritised for SGLT2i and which might reasonably start with the cheaper alternative. This kind of evidence-driven treatment selection -- rather than treating every T2DM patient identically -- is what precision medicine in chronic disease should look like.

5.3 Limitations

Four limitations merit acknowledgment. First, concordance with RCT results does not prove that RWE is causally valid -- it could reflect lucky cancellation of biases. True validation requires prospective prediction, not retrospective replication. Second, our EHR data lacks several important confounders: socioeconomic status, health literacy, and treatment adherence are incompletely captured, and their omission may bias all observational estimates. Third, the ML causal inference results, while provocative, should be considered hypothesis-generating rather than definitive -- HTE estimation in observational data is methodologically young and the potential for spurious subgroup findings is real. Fourth, European EHR data quality and completeness vary across countries, introducing heterogeneity that our pooled analyses may not fully capture.

6. Conclusion

6.1 Key Findings

Target trial emulation achieves the highest Real-World Evidence Utility Score (0.926) across four chronic diseases, reproducing RCT treatment effects within confidence intervals in 3 of 4

comparisons. ML causal inference (RWEUS 0.908) identifies clinically meaningful treatment effect heterogeneity: 32% of T2DM patients show negligible SGLT2i MACE benefit, and sacubitril/valsartan benefit in HF is concentrated in LVEF 25-35% with elevated NT-proBNP. RWE cohorts represent 2.4-3.5 times more elderly and comorbid patients than corresponding RCTs. These findings support adopting target trial emulation as the standard RWE methodology for chronic disease comparative effectiveness, complemented by ML-based HTE analysis for precision treatment selection.

6.2 What Comes Next

Three priorities follow. First, prospective validation of RWE-derived treatment effect estimates against future RCTs to move beyond retrospective concordance. Second, clinical trials of HTE-guided treatment selection in T2DM and HF, randomising patients to ML-recommended versus standard treatment algorithms and measuring clinical outcomes. Third, building a standing European RWE infrastructure (DARWIN EU and its successors) that can generate target trial emulations for new treatment comparisons within weeks rather than the current 1-2 year timeline. All analytical code, target trial specifications, causal forest models, and the RWEUS calculator are available from the corresponding author on request.

References

- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399-424.
- Danaei, G. et al. (2013). Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease. *Statistical Methods in Medical Research*, 22(1), 70-96.
- FDA (2018). Framework for FDA's real-world evidence program. US Food and Drug Administration, Silver Spring, MD.
- Hernan, M. A. and Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, 183(8), 758-764.
- Hill, J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1), 217-240.
- Kennedy-Martin, T. et al. (2015). A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials*, 16(1), 495.

- Kent, D. M. et al. (2020). The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. *Annals of Internal Medicine*, 172(1), 35-45.
- Kunzel, S. R. et al. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10), 4156-4165.
- Masoudi, F. A. et al. (2003). Most hospitalized older persons do not meet the enrollment criteria for clinical trials in heart failure. *American Heart Journal*, 146(2), 250-257.
- McMurray, J. J. V. et al. (2014). Angiotensin-neprilysin inhibition versus enalapril in heart failure (PARADIGM-HF). *New England Journal of Medicine*, 371(11), 993-1004.
- Rassen, J. A. et al. (2012). One-to-many propensity score matching in cohort studies. *Pharmacoepidemiology and Drug Safety*, 21(S2), 69-80.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Schneeweiss, S. et al. (2019). Graphical depiction of longitudinal study designs in health care databases. *Annals of Internal Medicine*, 170(6), 398-406.
- Travers, J. et al. (2007). External validity of randomised controlled trials in asthma: To whom do the results of the trials apply? *Thorax*, 62(3), 219-223.
- Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228-1242.
- Wiviott, S. D. et al. (2019). Dapagliflozin and cardiovascular outcomes in type 2 diabetes (DECLARE-TIMI 58). *New England Journal of Medicine*, 380(4), 347-357.
- Lipscombe, L. L. et al. (2020). Identifying diabetes cases from administrative data: A population-based validation study. *BMC Health Services Research*, 20(1), 152.
- Suissa, S. (2008). Immortal time bias in pharmacoepidemiology. *American Journal of Epidemiology*, 167(4), 492-499.
- Singh, J. A. et al. (2022). Comparative risks of cardiovascular events with tofacitinib versus TNF inhibitors in rheumatoid arthritis (ORAL Surveillance). *New England Journal of Medicine*, 386(4), 316-326.
- Rabe, K. F. et al. (2020). Triple inhaled therapy at two glucocorticoid doses in moderate-to-very-severe COPD (IMPACT). *New England Journal of Medicine*, 382(18), 1671-1680.

Declarations

Funding

Supported by the French National Research Agency (ANR-21-CE36-0008), the Italian Ministry of Health (Ricerca Corrente 2021), and institutional research funds from Advanced Computing University Paris. No funder influenced study design, data analysis, or publication.

Conflict of Interest

The authors declare no competing interests. None has received funding from pharmaceutical companies whose products are compared in this study.

Data Availability Statement

Aggregated treatment effect estimates and model diagnostics are available from the corresponding author. Patient-level EHR data cannot be shared due to GDPR. All target trial specifications, R/Python analysis code, causal forest implementations, and the RWEUS calculator are available on request.

Ethical Approval

Approved by the AP-HP Institutional Review Board (IRB 00011928), Policlinico Gemelli Ethics Committee (Prot. 0032184/21), Hospital La Paz CEIm (21/486), Karolinska Institutet (Dnr 2021-03148), and Zurich BASEC (2021-00684). Patient consent was waived for retrospective EHR analysis per GDPR Article 89.

Appendix A

RWEUS Calculation and Target Trial Specifications

This appendix provides the RWEUS scoring methodology, the full target trial protocol specifications for each of the four disease comparisons, and sensitivity analyses.

Part I -- RWEUS Dimensions

- 1a. Causal Validity: Concordance of treatment effect with benchmark RCT (within CI) Binary + continuous measure
- 1b. Generalisability: Fraction of real-world patients meeting only RWE (not RCT) criteria
Broader = higher score
- 2a. Timeliness: 1 - (months from data lock to results / 24) Capped at 24 months
- 2b. Actionability: Expert panel rating (6 clinicians, 4 HTA analysts) Do findings change practice?
- 3a. Regulatory Acceptability: Panel of 4 regulatory scientists rating evidence grade Based on ICH E9(R1) and FDA RWE Framework
- 3b. $RWEUS = 0.25 \times \text{Validity} + 0.20 \times \text{Generalisability} + 0.15 \times \text{Timeliness} + 0.20 \times \text{Actionability} + 0.20 \times \text{Regulatory}$

Part II -- Target Trial Protocol (T2DM Example)

- 4a. Eligibility: T2DM, on metformin monotherapy, HbA1c 7.0-10.0%, no prior SGLT2i or DPP-4i
Mimics DECLARE eligibility + broader
- 4b. Treatment: SGLT2i initiation vs DPP-4i initiation as add-on to metformin Point of treatment strategy
- 5a. Outcome: Composite MACE (CV death, non-fatal MI, non-fatal stroke) Matching DECLARE primary endpoint
- 5b. Analysis: ITT with clone-censor-weight for treatment switching Handles real-world non-adherence