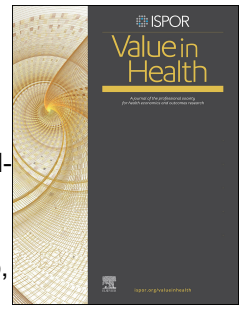
Journal Pre-proof



Flexible approaches based on multi-state models and microsimulation to perform real-world cost-effectiveness analyses: an application to pcsk9-inhibitors

Caterina Gregorio, PhD, Federico Rea, PhD, Francesca Ieva, PhD, Arjuna Scagnetto, MS, Carla Indennidate, MD, Chiara Cappelletto, MD, Andrea Di Lenarda, MD, Giulia Barbati, PhD.

PII: S1098-3015(24)00123-2

DOI: https://doi.org/10.1016/j.jval.2024.03.008

Reference: JVAL 3995

To appear in: *Value in Health*

Received Date: 12 September 2023

Revised Date: 18 March 2024

Accepted Date: 20 March 2024

Please cite this article as: Gregorio C, Rea F, Ieva F, Scagnetto A, Indennidate C, Cappelletto C, Di Lenarda A, Barbati G, Flexible approaches based on multi-state models and microsimulation to perform real-world cost-effectiveness analyses: an application to pcsk9-inhibitors, *Value in Health* (2024), doi: https://doi.org/10.1016/j.jval.2024.03.008.

Target Journal: VIH

Flexible approaches based on multi-state models and microsimulation to perform real-world cost-effectiveness analyses: an application to pcsk9-inhibitors

Caterina Gregorio[1,2,3] PhD, Federico Rea[4,5] PhD, Francesca Ieva[1,6] PhD, Arjuna Scagnetto[7] MS, Carla Indennidate[7] MD, Chiara Cappelletto[7] MD, Andrea Di Lenarda[7] MD, Giulia Barbati[2] ,PhD.

[1]MOX - Modelling and Scientific Computing Laboratory, Department of Mathematics, Politecnico di Milano, Milano, Italy

[2]Biostatistics Unit, Department of Medical Sciences, University of Trieste, Trieste, Italy

[3]Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden

[4]Department of Statistics and Quantitative Methods, Unit of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, Milan, Italy.

[5] National Centre for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy.

[6] HDS, Health Data Science center, Human Technopole, Milan, Italy

[7]Cardiovascular Center, Territorial Specialistic Department, University Hospital and Health Services of Trieste, Italy

Corresponding author:

Caterina Gregorio

Mailing Address: Via A.Valerio 4\1, 34127, Trieste, Italy

caterina.gregorio@units.it

+39 0405587334

Précis: Using flexible statistical methods and target trial emulation, PCSK9-I were found cost-effective to treat Hyperlipidemia using data from an electronic health database.

Word Count: 4026

Number of Pages: 26

Number of Figures: 3

Number of Tables: 3

Appendix:

Pages: 18

Figures: 6

Tables: 6

Author Contributions:

Concept and design: Caterina Gregorio, Andrea Di Lenarda and Giulia Barbati

Acquisition of data: Arjuna Scagnetto

Statistical Analysis: Caterina Gregorio, Federico Rea

Analysis and interpretation of data: Caterina Gregorio, Chiara Cappelletto, Giulia Barbati, Federico Rea

Provision of study materials/partecipants: Chiara Cappelletto, Carla Indennidate, Andrea Di Lenarda, Arjuna Scagnetto

Drafting of the manuscript: Caterina Gregorio, Francesca Ieva, Carla Indennidate, Federico Rea

Critical revision of the paper for important intellectual content: Francesca Ieva, Chiara Cappelletto

Obtaining funding: Andrea Di Lenarda

Administrative, technical, or logistic support: Arjuna Scagnetto

Supervision: Francesca Ieva and Giulia Barbati

Funding/Support: This work was supported by MA Provider S.r.l.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Highlights

• Cost-effectiveness analysis utilizing electronic health records databases can offer valuable real-world evidence for drugs recently approved. However, the statistical aspects of economic-health evaluations in chronic illnesses using decision models are often overlooked.

• We proposed the application of a flexible multi-state decision model based on microsimulation to replicate a target trial using observational data, enabling the study of PCSK9-I cost-effectiveness. Notably, these methods overcome the limitations of standard Markov decision models by considering the dependence of individuals' healthcare paths on their past.

• This study provides novel insights into the real-world cost-effectiveness of PCSK9-I in hyperlipidemia. Furthermore, the statistical approach employed here could also be useful for other diseases, treatments, or healthcare systems.

Abstract:

Objectives: This study aims to show the application of flexible statistical methods in real-world cost-effectiveness analyses applied in the cardiovascular field, focusing specifically on the use of PCSK9 inhibitors for hyperlipidaemia.

Methods: The proposed method allowed us to use an electronic health database to emulate a target trial for cost-effectiveness analysis using multi-state modelling and microsimulation. We formally established the study design and provided precise definitions of the causal measures of interest, while also outlining the assumptions necessary for accurately estimating these measures using the available data. Additionally, we thoroughly considered goodness-of-fit assessments and sensitivity analyses of the decision model, which are crucial to capture the complexity of individuals' healthcare pathway and to enhance the validity of this type of health economic models.

Results: In the disease model, the Markov assumption was found to be inadequate, and a "time-reset" timescale was implemented together with the use of a time-dependent variable to incorporate past hospitalization history. Furthermore, the microsimulation decision model demonstrated a satisfying goodness-of-fit, as evidenced by the consistent results obtained in the short-term horizon compared to a non-model-based approach. Notably, only in the long-term follow-up PCSK9 inhibitors revealed their favorable cost-effectiveness, with a minimum willingness-to-pay of 39,000 Euro/LY gained.

Conclusions: The approach demonstrated its significant utility in several ways. Unlike non-model based or alternative model-based methods, it enabled to 1) investigate long-term cost-effectiveness comprehensively, 2) employ an appropriate disease model that aligns with the specific problem under study, and 3) conduct subgroup-specific cost-effectiveness analyses to gain more targeted insights.

key-words

real-world data; cost-effectiveness; electronic health records; microsimulation; target trial emulation

Introduction

Economic evaluation to predict the cost-effectiveness (CE) profile and the financial consequences of adopting interventions for the healthcare system are of increasing importance as life expectancy, prevalence of chronic diseases, and costs for innovative treatments are rising. Such analyses are typically based on data gathered from randomized clinical trials (RCTs). However, RCTs have several limitations and strict enrollment criteria partially limit regulatory agencies in rules for real-world populations(1,2). Indeed, the efficacy observed in pre-marketing studies may be quite different from the effectiveness in clinical practice because of the following reasons: (i) frail patients are usually excluded from RCTs, (ii) trials are carried out in controlled environments whereas patient's low drug adherence and therapeutic inertia are common in real-practice, and (iii) the short-length of follow-up limits the assessment of long-term treatment benefits (and harms).

The attention of medical research for retrospective observational studies, especially those based on electronic health records or healthcare utilization databases (EHR), has progressively increased. Since all health services provided to the patients are included in these databases, the complete care pathway experienced by subjects can be identified, including clinical outcomes and healthcare costs. Therefore, EHR can be used to assess the impact of drugs introduced into the market in terms of the effectiveness of the treatment in reducing the progression of the disease for which they are prescribed and cost-effectiveness profile in specific areas and populations(3).Therefore, these data sources have the potential to enable more targeted and area-specific public health interventions.

However, statistical challenges in performing such analyses are the identification of appropriate methods to 1) consider the observational nature of data, 2) model health outcomes and cost in a complex time-to-event framework, 3) integrate methods for health-economic evaluations.

Decision models are common choices in economic evaluations to perform a comparison between competing decisions under uncertainty. Although these models are usually adopted to perform cost-effectiveness analyses based on data derived from RCTs, they are less applied in studies based on Real-World Data (RWD).

Methods for cost-effectiveness can be broadly categorized into non-model based and model-based approaches. Non-model-based approaches are useful for describing the current situation using available data. However, for formal comparisons and predictions of alternative treatment strategies, model-based methods are necessary, especially when generalizability to larger populations, lifetime scenarios or focus on specific subgroups is desired.

Model-based methods can be further divided into cohort models and individual-level (microsimulation) models (4,5). Cohort models are commonly used in health economics. However, they may not capture the complexities of real healthcare system mechanisms since they often assume Markovianity and time-homogeneity, which may not hold. On the other hand, microsimulation models generating individual life-course trajectories between health states are more flexible in taking into account in the subjects' temporal dynamics.

The application that motivated our work is the study of cost-effectiveness of Antibodies that inhibit proprotein convertase subtilisin–kexin type 9 (PCSK9-I). These are a new class of drugs that lower low-density lipoprotein (LDL) levels, preventing major cardiovascular events. The majority of the evidence in terms of the risk-benefit profile and economic-health assessments on the use of PCSK9-I is based on RCTs (6–8).

The main objectives of this study are to 1) present flexible statistical approaches to real-world cost-effectiveness analyses; 2) show how state-of-the-art multi-state methods can be combined to microsimulation to build a framework able to generate reliable and timely evidence of the sustainability of drug treatments; 3) combine target trial emulation techniques to limit the

danger of biases due to non-randomization; 4) perform a cost-effectiveness analysis for the addition of PCSK9-I to lipid-lowering therapy (LLT) in patients with hyperlipidemia from the payer perspective.

Methods

Study design and target trial emulation

Observational data from an Italian EHR is used to emulate a target trial for individuals eligible to the use of PCSK9-I according to the criteria established by the Italian Medicines Agency (AIFA). To emulate a target trial, it is necessary to develop a comprehensive protocol that outlines the fundamental design and analytical elements of the study (i.e., eligibility criteria, treatment strategy, assignment procedures, outcomes, follow-up period, causal contrast of interest and statistical analysis)(9–11). A summary of the components of the emulated trial's protocol for studying the cost-effectiveness of PCSK9-I is given in Table 1, together with the mitigation strategies used to address challenges in using EHR and potential sources of bias.

Data sources

The study is conducted using data from the Observatory of Cardiovascular Diseases of the Friuli-Venezia Giulia region(12) that systematically collects integrated administrative and cardiological clinical data that refer to the Trieste and Gorizia area (366.732. inhabitants). In Italy, all residents have equal access to health care by the National Health Service. The data sources interrogated for the present work are the Registry of Births and Deaths, Hospital Discharge data, Public Drug Distribution System, Exemption codes, cardiological e-chart (C@rdionet) and, examination results of public laboratories. According to the current Italian law, the study protocol was approved by the Unique Regional Ethics Committee Friuli-Venezia Giulia, with Protocol ID 185_2022.

Notation and estimand of interest

Let P be the treatment strategy indicator for the use of PCSK9-I, D the time to death and C the time to administrative censoring. Censoring time is assumed to be non-informative. The observation time is $Y^D = \min(D, C)$ and $\delta^D = I(Y^D = D)$ is the event indicator. We also denote by $M(W)$ the total medical costs up to a time horizon W. Because of death and censoring, the observed values related to the cost accrued up to time W that can be observed are $Y^W = \min(M(W), M(D), M(C))$. We let $M^H(w) = \{M(u), u \leq w\}$ be the intermediate cost history where $M(u)$ is the observed accumulated cost up to time u.

We also define $D^{(P)}$as the potential timing for the terminal event under the binary treatment strategy P and $M(W)^{(P)}$ the potential medical cost accumulated up to time W under the treatment strategy P.

The quantity of interest is the Incremental Cost-Effectiveness Ratio (ICER) at a time-horizon w defined as:

$$ICER \quad (w) = \frac{E\left[M(w)^{(1)} - M(w)^{(0)}\right]}{E\left[\, f[min\left(D^{(1)}, w\right)] - f[min\left(D^{(0)}, w\right)]\right]} \quad (1)$$

Note that $E\left[\, min\left(D^{(P)}, w\right)\right]$ are the mean Life Years (LY) over a time horizon w and $f(\cdot)$ denotes a generic function of the LY to encompass measures of quality of life, such as Quality Adjusted Life Years (QALY).

Moreover, as a further objective, we are interested in estimating a subgroup specific ICER:

$$ICER_x \quad (w) = \frac{E[M(w)^{(1)} - M(w)^{(0)} | X = x]}{E\left[\, f[min\left(D^{(1)}, w\right)] - f[min\left(D^{(0)}, w\right)] | X = x\right]} \quad (2)$$

where x defines eligibility subgroups as explained in Supplementary Material.

Two time horizons *w* are considered: a short-term one corresponding to the median follow-up in the cohort under study and the *lifetime* one.

To identify the causal contrasts involving the potential outcomes in the definition of the ICER, the usual assumptions for causal inference must hold (10). In our context, *consistency* refers to the principle that the time to the terminal event and the medical costs in a world where we intervene with treatment strategy *P* are the same in the real world where we observe the use of PCSK9-I. *Conditional exchangeability* assumes that the potential outcomes are independent of the allocation of the treatment, conditionally on the vector of observed covariates Z . Methods to achieve conditional exchangeability are discussed in the next Section. Moreover, censoring times are assumed to be conditionally independent of all potential event times. Finally, to satisfy *positivity*, for each vector of covariates Z , the probability of being treated with PCSK9-I must be greater than zero.

Adjustment by Inverse Probability of Treatment Weighting

To achieve conditional exchangeability, we consider Inverse Probability of Treatment Weights (IPTW) (13,14). A multivariable gradient boosting classifier as implemented in the *twang* R package (15) is used to estimate the weights in terms of possible measured confounders: demographics, Charlson Comorbidity index, past atherosclerotic cardiovascular disease (ASCVD), diabetes with Target Organ Damage (TOD) or a risk factor (smoking, obesity and hypertension), the history of treatment with statins (duration and adherence, measured as the Proportion of Days Covered(16,17) by treatment) and the eligibility date. The IPTW weights used to obtain the identifiability of the quantity in Equation (1) include all the above cited confounders, while for the quantity in Equation (2) the covariates used to define the eligibility subgroups are excluded. When implementing the methods outlined in the subsequent sections, it is consistently assumed that the dataset utilized has been weighted using IPTW.

Non-model based approach

It involves estimating the ICER non-parametrically by estimating the LY for each treatment group using the area under the Kaplan-Meier survival curve, while considering the treatment group as a stratifying factor. Simultaneously, the mean medical costs are estimated using the Bang and Tsiatis estimator (18), which takes into account censoring. It accomplishes this by appropriately weighting the sample mean medical costs in the two treatment groups. Confidence intervals are obtained through non-parametric bootstrap. It is important to note that, although this method does not rely on any modeling assumptions, only the short-term horizon can be considered since non-parametric estimates tend to become unstable when the number of individuals being observed is small.

Model-based approach

An alternative approach consists of specifying a suitable statistical model to describe the risk of terminal events in the two treatments group and the medical cost-generating process. Individual-level health economic models are considered here for their flexibility among model-based methods. The steps involved in obtaining such model consists in a) specify and fit a suitable disease model; b) specify a suitable cost and health outcomes model; c) run the decision (economic) model through the microsimulation; d) perform the decision analysis by estimating the ICER.

Disease model

In the context of this application, the healthcare paths of individuals over time can be achieved by employing a multi-state model that depicts the potential multiple hospitalizations an individual may experience until their death. According to previous studies(19), the set of discrete mutually exclusive states considered are: “out-of-hospital (out-of-hosp)”, “in-hospital for acute coronary syndrome (in-ACS)”, “in-hospital for ischemic stroke (in-IS)”, “in-hospital

for periphery artery disease (in-PAD)" "in-hospital for other cardiovascular causes (in-others CV)", "in-hospital for non-cardiovascular causes (in-others no CV)", "death". The model is further defined by the transition intensities, $q_{rs}(t)$, which represent the instantaneous probability of moving from one generic state *r* to another generic state, *s*, conditionally on being still alive. The possible states and permitted transitions are illustrated in Figure 1.

Under the simplest model, we assume that the transition intensities depend solely on the time since entry into the study and the treatment indicator. In such a model, there is no dependence of the transition intensities on the "history" of the process up to that specific time, i.e., the previous states visited by the individual and the time spent in each of them. Essentially, the process is considered Markov. Given the complex nature of the process involving subjects' interactions with the health-care system, in this study, we explore models capable of addressing potential violations of the Markov assumption. One approach considers the "clock-reset" time scale (see e.g., Putter et al.(20)) where time returns to zero at every transition. This enables us to model the hazard based on the time scale *u*, which represents the time since entry into the current state. To fit the model using the available data, the following cause-specific hazard models are employed for the transition intensities, conditional on the treatment indicator:

$$q_{rs}(u|P = p) = q_{rs}^{0}(u)\exp\{p\beta\}$$

Where $q_{rs}^{0}(u)$ is the baseline transition hazard, p is the covariate for the treatment indicator and β is its corresponding coefficient. $q_{rs}^{0}(u)$ is assumed to be parametric, but it is modelled via natural cubic splines to accommodate different shapes for hazard according to the Royston-Parmar flexible parametric model(21). This class of models was fitted using the R package *flexsurv* (22).

To introduce further dependence of the process on its history, both time-dependent covariates (e.g., the number of previous hospitalizations) and a frailty model are considered. The frailty

model considers the correlation between potential multiple transitions of the same type for the same individual by incorporating individual-specific random quantities known as frailties.

For model selection, e.g., selection of number of degrees of freedom of the baseline transition hazard, Akaike Information Criteria and Bayesian Information Criteria are used. Moreover, the overall goodness-of-fit of the models for the transition-hazards is verified by comparing the predicted values of the cumulative hazard obtained from the models to the non-parametric estimates.

Finally, for the subgroup analysis, the previous model was modified by adding X as additional covariate.

Cost model & Health Outcomes model

The cost model, using Euros as currency, is formulated based on the regulations of the Italian public healthcare reimbursement system. In line with previous studies (19,23) and, as all patients in the present analysis were above retirement age, only direct costs associated with medication and hospitalizations were considered. Regarding the hospitalizations, in Italy each Diagnosis Related Group (DRG) code has a predetermined cost if the hospitalization duration is below a certain threshold. If the hospitalization exceeds that threshold, a daily cost is applied. In our model, we assume that when an individual is admitted to the hospital with a specific DRG code, the fixed cost is assigned, and additional costs based on the length of the simulated hospital stay are attributed. It is worth noting that multiple DRG codes are possible for each in-hospital state, so we use a state-specific multinomial probability distribution to determine the probabilities of different DRG codes, which are estimated from the available data (See Supplementary Material and Supplementary Table 1). The average daily drug cost for the two treatment groups is derived from the Public Drug Distribution System and complete adherence is assumed for both groups for the entire time-horizon.

In the main analysis, no adjustment for quality of life is incorporated so there is no utility model to be defined. However, a sensitivity analysis using QALY instead of LY is performed. The utility model used is reported in Supplementary Table 2.

Finally, both costs and health outcomes are discounted at an annual rate of 3%(24).

Decision model through microsimulation and decision analysis

In essence, microsimulation involves simulating the life trajectories of individuals based on a specified health economic model using a random-number generator over the given time horizon w. Continuous microsimulation is used here because it does not require specification of model cycles and runs considerably faster(5). Conceptually, microsimulation can also be viewed as an instrument to replicate the target trial based on the specified protocol.

To obtain confidence intervals, Probability Sensitivity Analysis (PSA)(25) based on parametric bootstrap can be applied within the microsimulation framework. In standard microsimulation, individuals' paths are simulated based on the pointwise Maximum Likelihood estimate of the parameters that define the transition hazards. With PSA, the parameters are assumed to follow a multivariate normal distribution, according to the asymptotic behavior of the Maximum Likelihood Estimator. Therefore, in microsimulation with PSA, we first generate a random sample of B values for the parameter vector. Then, for each drawn parameter vector, we conduct a microsimulation with a sample of N individuals for each treatment strategy.

The microsimulation is performed using the *hesim*(5) R package with N=1000 individuals for each of the 500 PSA samples and each treatment strategy (1 000 000 in total). Convergence diagnostics are reported in Supplementary Figure 1.

The required estimates to calculate the ICER under each treatment strategy scenario, over the horizon w, are computed by averaging the total health outcomes and total costs across the simulated patients for each sample $b = 1, \dots, B$.

Finally, the marginal and conditional ICER (along with their corresponding 95% confidence intervals) are calculated as the mean (2.5% and 97.75% quantiles) of the corresponding distribution derived from the B bootstrap samples.

Sensitivity Analyses

Different sensitivity analyses are carried out to assess the robustness and generalizability of the results. The first regards the extrapolation beyond the maximum follow-up observed in the data (Sensitivity A). First, an independent historical cohort of subjects has been extracted *ad hoc* from the health electronic health records database with a follow-up compatible with the duration of the lifetime microsimulation. Transition hazards are estimated using this dataset and they were subsequently incorporated in the decision model as baseline transition hazards, after having them opportunely recalibrated on the study cohort. In a second sensitivity analyses, a different scenario considering that a portion of individuals are not adherent to the treatment is considered (Sensitivity B). In addition, we used the EValue methodology (26) (using the *R* package *Evalue* (27)*),* to quantify the degree to which the cost-effectiveness results may be affected by different unmeasured confounders scenarios when estimating the treatment effects using observational data (Sensitivity C). The E-value was selected because of its lack of dependence on specific assumptions and its flexibility (28), making it suitable for integration into the methodology used for the cost-effectiveness analysis. Specifically, we consider three microsimulations with a less protective treatment effect for all-cause death compared to the main analysis. Finally, Sensitivity Analysis D incorporates utilities (Supplementary Table 2) to estimate QALY.

The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) reporting guideline checklist (29) is reported in the Supplementary Material.

Results

Study cohort

We extracted data related to 96,886 subjects with at least one measure of LDL available in the enrollment period (from 1/7/2017 to 31/12/2020). Among these, at least one of the eligibility criteria occurred during the observation period for 1,976 subjects (2%). Among them, 161 (8%) were prescribed to PCSK9-I. The median follow-up time was 34 months.

Subjects on LLT+PCSK9-I were slightly younger, prevalently males, with more severe CV conditions and a higher rate of statin treatment with respect to the subjects belonging to the LLT group (Supplementary Table 3). Patients non-treated with PCSK-I showed higher prevalence of comorbidities such as diabetes, chronic obstructive pulmonary disease, and renal diseases.

Diagnostics of the IPTW procedure are shown in Supplementary Table 4. For all the covariates a satisfactory balance has been achieved. Individuals with estimated propensity scores indicating potential violations of the positivity assumption were excluded from the diagnostics and subsequent analysis (26% of the study cohort).

Disease-model, cost and utility model

The comparison between the "time-reset" scale and the "time-forward" (Markov model) is depicted in Figure 2, confirming the relevance of considering the former time scale in this context to accurately capture the transition intensities from in-hospital states to the out-of-hospital state (Figure 2, from panels f) to l)). Regarding the dependence on the past history of the process, a time-varying covariate that distinguishes the first hospitalization from subsequent ones exhibited the best goodness-of-fit based on the Akaike Information Criteria and it was therefore chosen among models with different definitions of the time-dependent variable and the frailty model.

The final models incorporated distinct baseline hazards for each transition, but the effect of PCSK9-I treatment was assumed to be consistent across different causes of hospitalization and between death occurring in and out of the hospital (Table 2). We made this assumption because considering a more complex specification of the model did not demonstrate a significant improvement in terms of goodness of fit.

A significant strong protective effect of PCSK9-I on all-cause death (HR=0.14, 95% CI 0.07-0.27) was observed and a significant protective effect of PCSK9-I was detected only for the transition towards the first hospitalization (HR=0.79, 95% CI 0.63-0.99).

All the disease models' parameters are reported in Supplementary Table 5 (main analysis) and in Supplementary Table 6 (subgroup analysis).

Cost-effectiveness results

The short-term ICER at 34 months obtained with the model-based and non-model based approaches were not statistically different and showed a minimum willingness to pay of Euro/LY >200 000 (Supplementary Figure 2). The results of the lifetime analysis model are reported in Table 3. According to these results**,** an ICER of 29 540 (95% CI: 23 773-38 949) Euro/LY was obtained (Figure 3, panel A).

The subgroup analysis showed that patients with diabetes with organ damage and/or a risk factor have a lower minimum willingness to pay (Figure 3, panel B and Supplementary Table 7).

Results of the sensitivity analyses

Sensitivity analysis A leads to results consistent with the ones obtained through the main lifetime decision model (Supplementary Figure 3a) and the long-term mortality rates of the

historical cohort are overlapping with the ones estimated from the disease model for the non-treated group (Supplementary Figure 3b) .

In Sensitivity Analysis B, considering an estimated fraction of 7% of non-adherent individuals to PCSK9-I, as reported by Arca et al.(30), an ICER of 29,905 (95% CI: 23 982-38 604) Euro/LY was observed (Supplementary Figure 4).

Assuming a treatment effect for all-cause death more similar to the one observed in RCTs, the ICER reached 58,000 Euro/LY (Supplementary Figure 5).

In Sensitivity Analysis D, an ICER of 29 292 (95% CI: 23 550-37 888) Euro/QALY was estimated (Supplementary Figure 6).

Discussion

Cost-effectiveness analysis using RWD is a promising yet challenging field. In our study, we combined target trial emulation with flexible statistical methods. In chronic illnesses such as cardiovascular conditions, assessing lifetime cost-effectiveness is crucial. Our application on PCSK9-I revealed substantial differences in cost-effectiveness between short and long-term perspectives.

To evaluate long-term cost-effectiveness, decision models based on RWD become essential since they encompass limitations common to RCTs and observational studies related to a limited follow-up period, when the focus is on drugs recently approved. Thus, selecting an appropriate decision model and assessing its goodness-of-fit using available observational data are crucial steps. While non-parametric methods have limitations in conducting comprehensive scenario analyses, they can still provide valuable insights. By keeping the same time horizon, the results obtained from the non-parametric approach should be consistent with those derived from the decision model, if the disease and cost models are correctly specified. Indeed, this was the case for our decision model. In this study, we were able to achieve this by employing a

flexible parametric multi-state model combined with a microsimulation model. While cohort models are suitable for obtaining marginal estimates of cost-effectiveness, microsimulation can simulate individual life-course trajectories between health states, allowing for more personalized analyses. Moreover, individual-level models naturally capture the accumulation of costs in real healthcare systems.

Using this approach, we could overcome the limitations of Markov assumptions and incorporate history of hospitalizations. This significantly improved the goodness-of-fit. The inadequacy of Markov models for modeling healthcare paths is a well-established topic in biostatistics literature (31). Nevertheless, decision models based on the Markov assumption are standard methods in cost-effectiveness analyses for chronic illnesses.

In addition to the disease-model's goodness-of-fit, it is also essential to assess the convergence of the microsimulation and perform different sensitivity analyses. Lifetime decision models involve extrapolation. In this paper, we have tested the robustness of such extrapolation using data from a historical cohort extracted from our EHR database. We also considered a scenario in which not all individuals prescribed to the treatment adhere to it, according to the observed non-adherence rate in Italy for PCSK9-I(30). Finally, the treatment effect of the drug on the risk of death estimated in our study was much higher than the one observed in the RCTs(6–8). This may be partly due to the higher cardiovascular risk of our cohort of subject. As it has been shown in a RCTs subgroup analysis(6), patients at higher risk seem to benefit the most from PCSK9-I. However, unmeasured confounding could also not be ruled out given the observational nature of the study. Therefore, in another sensitivity analysis, we assessed how the estimate for the ICER changed according to different scenarios of unmeasured confounding resulting in a treatment effect of PCSK9-I on death closer to the one reported in a meta-analysis(32).

To the best of our knowledge, this study provides the first cost-effectiveness analysis on PCSK9-I for the Italian healthcare system using RWD. According to our results, the ICER about 30 000 Euro per health outcome gained, both considering the LY and QALY. However, in case of a much less protective treatment-effect on death, the ICER reaches 58 000 Euro per LY gained. Results from other investigations on the cost-effectiveness of PCSK9-I are heterogeneous, as are the health economic models employed. However, some results are in line with the one obtained in this study(19,23).

Conclusions

In conclusion, this work provides evidence on the cost-effectiveness of PCSK9-I using RWD. Furthermore, this study demonstrates the potential of individual-level decision models for cost-effectiveness analysis using RWD. The framework of disease and cost models presented here can be extended to other applications or healthcare systems. Moreover, it could be possible to consider scenarios where it is of interest to examine cost-effectiveness based on more detailed subject profiles, allowing for personalized analyses.

References

1. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ [Internet]. 1996 May 11 [cited 2024 Jan 7];312(7040):1215–8. Available from: https://pubmed.ncbi.nlm.nih.gov/8634569/

2. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products | FDA [Internet]. [cited 2024 Jan 7]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory

3. Bowrin K, Briere J-B, Levy P, Millier A, Clay E, Toumi M. Cost-effectiveness analyses using real-world data: an overview of the literature. J Med Econ. 2019 Jun;22(6):545–53.

4. Wu O. Microsimulation Model for Health Economic Evaluation of Public Health Policies: An Imperfect but Useful Tool. Vol. 144, Circulation. United States; 2021. p. 1377–9.

5. Incerti D, Jansen JP. hesim: Health Economic Simulation Modeling and Decision Analysis. 2021.

6. Ridker PM, Revkin J, Amarenco P, Brunell R, Curto M, Civeira F, et al. Cardiovascular Efficacy and Safety of Bococizumab in High-Risk Patients. N Engl J Med. 2017 Apr;376(16):1527–39.

7. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. N Engl J Med. 2017 May;376(18):1713–22.

8. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, et al. Alirocumab and Cardiovascular Outcomes after Acute Coronary Syndrome. N Engl J Med. 2018 Nov;379(22):2097–107.

9. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol [Internet]. 2016 Nov 1;79:70–5. Available from: https://doi.org/10.1016/j.jclinepi.2016.04.014

10. Hernan MA, Robins JM. Causal inference: What if. CRC Press. 2021.

11. Hernán MA. Methods of Public Health Research — Strengthening Causal Inference from Observational Data. N Engl J Med [Internet]. 2021 Oct 2;385(15):1345–8. Available from: https://doi.org/10.1056/NEJMp2113319

12. Iorio A, Sinagra G, Di Lenarda A. Administrative database, observational research and the Tower of Babel [Internet]. Vol. 284, International Journal of Cardiology. Elsevier Ireland Ltd; 2019 [cited 2021 Jul 2]. p. 118–9. Available from: https://pubmed.ncbi.nlm.nih.gov/30558991/

13. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika [Internet]. 1983 Jun 27;70(1):41–55. Available from: http://www.jstor.org/stable/2335942

14. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol. 1974;66:688–701.

15. Cefalu M, Ridgeway G, McCaffrey D, Morral A, Griffin BA, Burgette L. twang: Toolkit for Weighting and Analysis of Nonequivalent Groups [Internet]. 2021. Available from: https://cran.r-project.org/package=twang

16. Andrade SE, Kahler KH, Frech F, Chan KA. Methods for evaluation of medication adherence and persistence using automated databases. Pharmacoepidemiol Drug Saf. 2006;15(8):565–74.

17. Dima AL, Dediu D. Computation of adherence to medication and visualization of medication histories in R with AdhereR: Towards transparent and reproducible use of electronic healthcare data. PLoS One [Internet]. 2017 Apr 26;12(4):e0174426. Available from: https://doi.org/10.1371/journal.pone.0174426

18. Bang H, Tsiatis AA. Estimating medical costs with censored data. Biometrika [Internet]. 2000 Jun 1;87(2):329–43. Available from: https://doi.org/10.1093/biomet/87.2.329

19. Landmesser U, Lindgren P, Hagström E, Van Hout B, Villa G, Pemberton-Ross P, et al. Cost-effectiveness of proprotein convertase subtilisin/kexin type 9 inhibition with evolocumab in patients with a history of myocardial infarction in Sweden. Eur Hear J - Qual Care Clin Outcomes. 2022;8(1):31–8.

20. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007 May;26(11):2389–430.

21. Royston P, Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med [Internet]. 2002 Aug 15 [cited 2022 Nov 1];21(15):2175–97. Available from: https://pubmed.ncbi.nlm.nih.gov/12210632/

22. Jackson CH. flexsurv: A Platform for Parametric Survival Modeling in R. J Stat Softw. 2016 May;70.

23. Azari S, Rezapour A, Omidi N, Alipour V, Behzadifar M, Safari H, et al. Cost-

effectiveness analysis of PCSK9 inhibitors in cardiovascular diseases: a systematic review. Heart Fail Rev. 2020;25(6):1077–88.

24. Organization WH, Baltussen RMPM, Adam T, Tan-Torres Edejer T, Hutubessy RCW, Acharya A, et al. Making choices in health : WHO guide to cost-effectiveness analysis / edited by T. Tan-Torres Edejer ... [et al] [Internet]. Geneva PP - Geneva: World Health Organization; Available from: https://iris.who.int/handle/10665/42699

25. O'Hagan A, Stevenson M, Madan J. Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA. Health Econ [Internet]. 2007 Oct [cited 2022 Nov 1];16(10):1009–23. Available from: https://pubmed.ncbi.nlm.nih.gov/17173339/

26. Haneuse S, Vanderweele TJ, Arterburn D. Using the E-Value to Assess the Potential Effect of Unmeasured Confounding in Observational Studies. JAMA [Internet]. 2019 Feb 12 [cited 2023 Dec 18];321(6):602–3. Available from: https://jamanetwork.com/journals/jama/fullarticle/2723079

27. Van Der Weele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med [Internet]. 2017 Aug 15 [cited 2023 Dec 18];167(4):268–74. Available from: https://pubmed.ncbi.nlm.nih.gov/28693043/

28. Chung WT, Chung KC. The use of the E-value for sensitivity analysis. J Clin Epidemiol. 2023 Nov 1;163:92–4.

29. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. Value Heal. 2022 Jan 1;25(1):3–9.

30. Arca M, Celant S, Olimpieri PP, Colatrella A, Tomassini L, D'Erasmo L, et al. Real-World Effectiveness of PCSK9 Inhibitors in Reducing LDL-C in Patients With Familial Hypercholesterolemia in Italy: A Retrospective Cohort Study Based on the AIFA Monitoring Registries. J Am Heart Assoc. 2023 Nov;12(21):e026550.

31. Ieva F, Jackson CH, Sharples LD. Multi-State modelling of repeated hospitalisation and death in patients with Heart Failure: the use of large administrative databases in clinical epidemiology. Stat Methods Med Res [Internet]. 2017 Jun 1 [cited 2022 Mar 24];26(3):1350. Available from: /pmc/articles/PMC4964935/

32. Imran TF, Khan AA, Has P, Jacobson A, Bogin S, Khalid M, et al. Proprotein convertase subtilisn/kexin type 9 inhibitors and small interfering RNA therapy for cardiovascular risk reduction: A systematic review and meta-analysis. de Carvalho LSF, editor. PLoS One [Internet]. 2023 Dec 6 [cited 2024 Jan 7];18(12):e0295359. Available from: https://pubmed.ncbi.nlm.nih.gov/38055686/

Table 1 Summary of the protocol components of a target trial to study the cost-effectiveness of PCSK9-I in Hypercholesterolemia.

Protocol Component	Description of the target trial	How was the protocol element emulated using observational EHR data?	Challenges and potential sources of bias	Mitigation strategies to overcome challenges and potential bias
Eligibility criteria ***Who will be included in the study?***	Individuals eligible for the use of PCSK9-I to treat Hypercholesterolemia and living in the Trieste-Gorizia area of Italy according to the related reimbursement criteria established by AIFA. See the Supplementary Material for the specific criteria.	Same as for the target trial with the exception that a single measurement of LDL above the threshold was considered valid. *Required data for each person: age, LDL measurements, anamnesis and family history of ASCVD and pharmacological treatment history*	Data might be insufficient to emulate the trial's eligibility criteria Leading to selection bias/confounding	The EHR data source used has complete coverage of the population of interest and contains all the available information necessary to define the eligibility criteria for PCSK9i (e.g. laboratory values and treatment history). Expert opinion was used to translate the target trial criteria (e.g. definition of ICD9 from inpatient data and their coupling with clinical diagnoses made by the cardiologists during specialist visits)
			The population selected for the study might include patients for whom there is no equipoise between treatment strategies leading to confounding bias	According to the expert opinion, there were no reason to assume no equipoise for specific subgroups. During the IPTW diagnostics, we assessed the presence of patients for which there was no

				equipoise according to the distribution of the estimated propensity score, excluding them from the subsequent analyses.
			The population selected for the study might fail to include subgroups of interest for the cost-effectiveness analysis leading to non-generalizable results or omission of relevant subgroup analyses	The EHR data source is representative of the target population and subgroups of interest were identified using expert opinion.
Treatment strategies ***What interventions will eligible persons receive?***	Either standard LLT or LLT+ PCSK9-I	Same as for the target trial *Required data for each person: date of first prescription of* PCSK9-I	The definition of the intervention might differ from the intervention of interest.	It was possible to define precisely LLT and PCSK9-I using ATC codes and they reflect the ones routinely used in clinical practice, according to expert opinion.
			The comparator strategy might not be defined with a sufficient level of detail.	It was possible to define precisely LLT using ATC codes and they reflect the ones routinely used inclinical practice, according to expert opinion

Time zero and Follow-up period ***During which period will eligible persons be followed in the study?***	Time zero is the moment in which the subject starts being eligible for PCSK9-I. The recruitment period was from 1/7/2017 (entry of PCSK9-i in the Italian market) to 31/12/2020. The follow-up ends at the earliest of death loss at follow-up, or administrative end of the study (31/12/2021).	Time zero was assumed for individuals treated with PCSK9-I as the date of the first prescription and for the comparator group as the date at which the patients satisfied all the eligibility criteria. The follow-up was defined as in the target trial since complete follow-up data was available. *Required data for each person: date of first prescription of PCSK9-I, date of eligibility to PCSK9-I, date of death, date of censoring.*	The start of follow-up might predate the assessment of the eligibility criteria leading to selection bias.	Time zero was chosen so that the start of the follow-up started when the assessment of the eligibility criteria had been made.
			The time of treatment assignment might not be aligned with that of eligibility assessment and the start of follow-up leading to immortal time bias.	The time zero chosen ensures that it minimizes time to treatment initiation since for PCSK9-I the date of treatment initiation should very closely follow the date of the first prescription, according to clinical guidelines and routine clinical practice.
Assignment procedures ***How will eligible persons be assigned to the interventions?***	Eligible participants will be randomly assigned to the two strategies and will be aware of the strategy to which they have been assigned.	Eligible persons will be assigned to the strategies with which their data are compatible.		
Outcomes ***What outcomes in eligible persons will be compared among intervention groups?***	Medical costs and LY on a) a short time horizon (34 months) and b) life-time horizon taking into account possible repeated hospitalizations over time**.**	Same as for target trial (cost-effectiveness outcome) *Required data for each person: dates of entry/exits from hospital with corresponding*		

		ICD9-CM and DRG code and date of death.		
Causal contrasts of interest ***Which counterfactual contrasts will be estimated using the above data?***	Intention-to-treat effect (effect of being assigned to treatment).	Observational analogue of the Intention-to-treat effect.		
Statistical analysis ***How will the counterfactual contrasts be estimated?***	Intention-to-treat analysis via estimation of the ICER through multi-state models and microsimulation.	Same as intention-to-treat analysis	Confounding might exist after emulating the main components of the target trial, from both measured and unmeasured prognostic factors.	Inverse Probability of Treatment Weighting is used (together with diagnostics to assess the achievement of balance between the treatment groups) to eliminate confounding due to measured confounders. Sensitivity analyses based on the E-Value method are used to address the impact of possible unmeasured residual confounding in the cost-effectiveness results.

AIFA= Italian Medicines Agency; DRG: Diagnosis Related Group; EHR: Electronic Health Records; ICER= Incremental CERatio; LDL: Low-Density Lipoprotein; LLT= Lipid-Lowering Therapy; LY: Life Years; PCSK9-I= Proprotein Convertase Subtilisin–Kexin type 9- Inhibitors;

Table 2. Effect of PCSK9-I+LLT estimated on transition hazards through the disease model.

Transition type		HR	95% CI
Towards Hospital	PCSK9-I+LLT vs LLT first hospitalization	0.79	0.63 ; 0.99
	PCSK9-I+LLT vs LLT 2+ vs 1 hospitalization	0.91	0.62 ; 1.34
Towards Out-of-Hospital	PCSK9-I+LLT vs LLT 2+ vs 1 hospitalization	1.21	0.93 ; 1.56
	PCSK9-I+LLT vs LLT 2+ vs 1 hospitalization	0.87	0.58 ; 1.31
Towards Death	PCSK9-I+LLT vs LLT	0.14	0.07 ; 0.27

CI: Confidence Interval; HR: Hazard Ratio; LLT= Lipid-Lowering Therapy; PCSK9-I= Proprotein Convertase Subtilisin–Kexin type 9- Inhibitors;

Table 3 Results of the microsimulation economic model.

	LLT (95% CI)	**PCSK9-I+LLT (95% CI)**
***Mean Utility* (years)**	17.13 (16.45, 17.75)	20.02 (19.56, 20.32)
***Mean Costs*: Drugs (Euros)**	1,850 (1,776, 1,917)	88,655 (86,611, 89,986)
***Mean Costs*: Hospitalizations length of stay below threshold (Euros)**	6,228 (3,806, 11,577)	5,510 (3,241, 10,217)
***Mean Costs:* Hospitalizations extra days (Euros)**	2,704 (1,920, 3,965)	1,954 (1,229, 2,906)

CI: Confidence Interval; HR: Hazard Ratio; LLT=Lipid-Lowering Therapy; PCSK9-I= Proprotein Convertase Subtilisin–Kexin type 9- Inhibitors.