



Economic Evaluations of Pharmacological Treatments in Heart Failure Patients: A Methodological Review with a Focus on Key Model Drivers

Gian Luca Di Tanna¹ · Shuxian Chen² · Anna Bychenkova³ · Heidi S. Wirtz⁴ · Karen L. Burrows⁵ · Gary Globe⁴

Published online: 26 August 2019
© The Author(s) 2019

Abstract

Various decision analytic models exist for evaluating the cost-effectiveness of pharmacological interventions for heart failure (HF). Despite this, studies that explore drivers influencing these modeling approaches remain scarce. Through a systematic review of the literature, the present study sought to identify model drivers that emerge from economic evaluations of HF pharmacological interventions. Among the 72 cost effectiveness papers evaluating HF drug interventions, the most frequently identified, top 5 ranked model drivers impacting the incremental cost-effectiveness ratio (ICER) were cost of treatment and utility, identified in 10% of studies, respectively. Other drivers that emerged as top 5 ranked drivers in > 5% of studies included treatment effect on mortality (or cardiovascular mortality), duration of treatment, and baseline cardiovascular mortality. Model drivers reported at the top of tornado diagrams were treatment effect on mortality or on cardiovascular mortality. Collectively, these observations highlight the key importance of treatment effect in driving cost-effectiveness models for HF.

Key Points for Decision Makers

Treatment effect on mortality was the main model driver that emerged from models captured in this systematic review of cost-effectiveness studies of pharmacological interventions for heart failure (HF).

The most frequently identified, top 5 ranked model drivers impacting on the incremental cost-effectiveness ratio were cost of treatment and utility, identified in 10% of relevant studies.

While cost of treatment emerged as a main model driver, our findings demonstrate the higher relative importance of treatment effect on mortality in cost-effectiveness models of HF. This could be due to the simplicity of the model structures employed.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s41669-019-00173-y>) contains supplementary material, which is available to authorized users.

✉ Gary Globe
gglobe@amgen.com

¹ The George Institute for Global Health, Sydney, Australia

² The Comparative Health Outcomes, Policy and Economics (CHOICE) Institute, University of Washington, Seattle, WA, USA

1 Introduction

Heart failure (HF) is a well-recognized public health concern that imposes high economic and societal costs worldwide [1–3], and is predicted to rise as the population ages [1, 4]. Evidence-based treatment of HF can be effective for lowering both morbidity and mortality [5–8]. However, to guide clinical decision making, estimates of their cost-effectiveness profiles and related main drivers are needed. Gathering evidence from modeling studies of pharmacological treatments that report on cost-effectiveness in HF is of particular value. In addition to highlighting relevant core modeling specifications and structures, systematic analysis may allow the determination of key components that drive these economic models. Through a systematic review of the literature, we previously highlighted some similarities in modeling approaches when assessing the cost-effectiveness of interventions for HF [9]. However, our analysis recognized the need for modeling approaches that would allow patients to be reliably assessed throughout their journey,

³ Global Health Economics, Amgen Inc., Uxbridge, UK

⁴ Global Health Economics, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

⁵ Envision Pharma Group, Horsham, UK

e.g., through hospitalizations, re-hospitalizations, and ideally outpatient visits [9]. We also raised the need to further investigate what was driving these cost-effectiveness models, as studies that specifically explore model drivers remain scarce. As an extension to our previous review, this study sought to identify model drivers that emerge from a systematic review of cost-effectiveness models in HF treatment.

2 Methods

A systematic review of the literature was conducted using a combination of search terms (see Appendix I in the supplementary electronic material [ESM], and described in full elsewhere [9]). This follow-on analysis used the same systematic strategy within MEDLINE, EMBASE, the National Health Service Economic Evaluation Database, the Cost-Effectiveness Analysis Registry, the SchARR Health Utilities Database, and the Cochrane Library Database, from January 1, 1997 extended to December 1, 2018 (Appendix I, see the ESM). The search was confined to cost-effectiveness studies conducted in patients with HF who were aged ≥ 18 years that provided details related to a pharmacological treatment for HF, and documented any mathematical model of cost-effectiveness specific to patients. Studies that were experimental, preclinical, pharmacokinetic, pharmacodynamic or observational in nature were excluded. Case reports or a case series ($n < 10$ patients), letter(s) to an editor, or opinion pieces or review articles were not eligible for inclusion. Studies published prior to 1997 were not captured within the search parameters (Appendix I, see the ESM).

Titles and abstracts of all retrieved records were screened and duplicated publications removed. Relevant full-text records were examined in detail for information relating to study design, participants and population, intervention(s), treatment(s), and outcome(s). Manuscripts were also assessed to identify model type and software used, model drivers, and information related to probabilistic sensitivity analysis (PSA) and value of information (VOI).

The top 5 model drivers were extracted from tables and tornado diagrams for each study, where available. We ranked the parameters in each model by their scale of impact on the incremental cost-effectiveness ratio (ICER), and documented the first five with the largest impact (e.g., the first five parameters listed on the tornado diagram). When extracted, model drivers were grouped into categories and ranked according to frequency of identification. Model drivers listed on the top of tornado diagrams or first reported within tables (where available) or reported in the study conclusions text regarding sensitivity analyses were also captured. We considered these “first-reported” model drivers as the main driver in the economic model. First-reported model drivers were also

grouped according to category and ranked according to frequency of inclusion in the models. A “key model driver” was considered for descriptive purposes in this study as a model driver identified in $> 5\%$ of studies, but all model drivers were captured a priori.

Both the search and screening phases were independently conducted by two trained investigators. Any disagreements were resolved by consensus with a senior investigator. For models retrieved from the study selection process, we reviewed good modeling practice and reporting by using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist [10, 11]. One study by Polistena et al. [12] was excluded from the CHEERS assessment as the paper was published in Italian; however, information was easily extractable for the qualitative analysis.

3 Results

3.1 Study Selection

Of the 4276 citations identified, 69 were duplicates. A further 4086 articles failed to meet inclusion criteria, based on title and abstract screening. Forty-nine publications were excluded following full-text review, due to reasons outlined in Fig. S1 (Appendix II, see the ESM). As a result, 72 articles were included. One study by Ollendorf et al. [13] was not included in the model driver assessment as the paper only reported a summary of results from an earlier cost-effectiveness analysis.

3.2 Description of Studies

Studies meeting eligibility criteria are summarized in Table S1 (Appendix III, see the ESM). Most of the 72 studies were carried out in the USA (26%) or UK (24%). Sacubitril/valsartan (19%), eplerenone (17%), or ramipril (14%) were the most commonly studied interventions. The majority of study participants were male, and the mean age ranged from 57 to 76 years. Study samples ranged from 49 to 11,055 patients. The majority of studies (72%) used the New York Heart Association (NYHA) classification for categorizing HF; however, 28% did not report HF classification.

For the 72 studies, Markov modeling (53%) and a trial or registry-based analysis (28%) were the most common approaches used. Discrete-event simulation (8%) and survival analysis (7%) were the only other approaches used by more than one study.

3.3 Sensitivity Analyses

Sensitivity analyses were carried out by the majority of studies. Fifty-seven studies (79%) used deterministic sensitivity

analysis (DSA); 53 of these (93%) employed a one-way sensitivity analysis. Variables considered in studies using one-way sensitivity analyses are given in Table S1 (Appendix III, see the ESM). Among the studies that conducted one-way sensitivity analysis, 14 of these (26%) used 95% confidence intervals where applicable. Ten studies (19%) reported that they varied the parameters over plausible ranges, while 12 studies (23%) specified very large variations (from $\pm 20\%$ up to $\pm 50\%$) from the point estimates. Eleven studies overall (15%) carried out a two-way analysis. A PSA was carried out by 33 studies (46%). Of the studies that carried out a PSA, 27 (82%) reported the number of simulations (most commonly 1000). Twelve studies (17%) reported use of bootstrapping techniques with the aim to capture uncertainty. The majority of studies that used bootstrapping were conducted in a trial-registry-based dataset (83%). Twenty-four studies (33%) included a tornado diagram.

In general, we found that PSA reported adequate parametric distributions around the various parameters evaluated. The most common distributions used for baseline risks were the beta ($n = 10/18$ [56%]) and Gompertz approach ($n = 4/18$ [22%]). For treatment effect and relative risk, the log-normal approach was used by the majority of relevant studies ($n = 9/16$ [56%]), followed by a normal distribution (19%). For utilities, a beta distribution was most commonly used ($n = 11/15$ [73%]). For costs, a gamma distribution was most commonly employed ($n = 7/17$ [41%]), followed by a triangular or log-normal approach ($n = 3$ each).

3.4 Key Model Drivers and Value of Information

Of the studies identified by the search strategy (Appendix II, see the ESM), 37 studies (51%) explicitly discussed model drivers in detail, within the text of the manuscript (Table S1, Appendix III, see the ESM). Model drivers that emerged from the analysis are given in Fig. 1. Cost of treatment and utility were the most frequently identified top 5 ranked model drivers, reported in 10.1% of models. Other key drivers (i.e., identified in more than 5% of studies) included treatment effect on mortality, duration of treatment, treatment effect on cardiovascular (CV) mortality, and baseline CV mortality (Fig. 1a). Treatment effect on mortality was the main (first-reported) model driver in 24.3% of studies. Treatment effect on CV mortality and model time horizon emerged as first-reported model drivers, in 13.5% and 8.1% of studies, respectively. Other first-reported model drivers in $> 5\%$ of studies were baseline all-cause mortality risk, cost of hospitalization, duration of effectiveness, and baseline CV mortality (5.4% of studies, respectively; Fig. 1b). Beyond these, a list of heterogeneous model drivers was identified, with many drivers reported by only one model.

Few studies reported on VOI, as just two studies (3%) included an analysis of expected value of perfect information (EVPI).

3.5 CHEERS Assessment

The majority of CHEERS items [10, 11] were documented, and most articles demonstrated reasonable methodological reporting quality (Fig. S2; Appendix II, see the ESM). Domains frequently not reported on included “characterizing heterogeneity” (by 18 studies [25%]) and “choice of model” (by seven studies [10%]).

4 Discussion

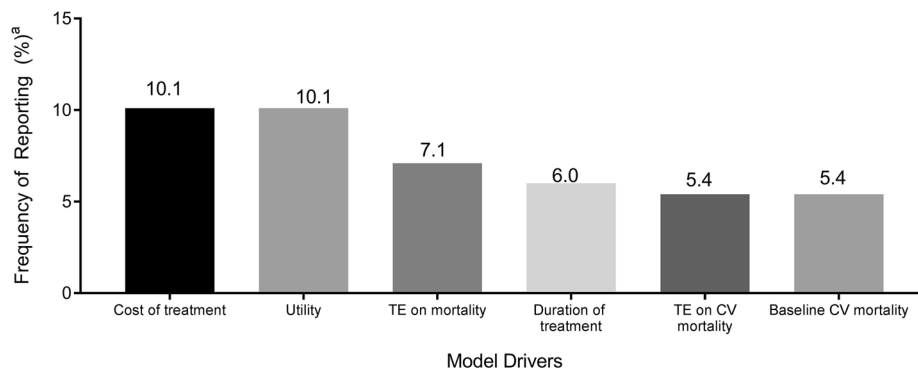
Through a systematic review of the literature, this study highlights the key model drivers that emerge from an assessment of cost-effectiveness studies for HF management. From the top 5 ranked model drivers, cost of treatment and utility most commonly emerged, and the model drivers ranked first were treatment effect on mortality or on and CV mortality. Interestingly, treatment effect on mortality was more frequently the main model driver than treatment time horizon or hospitalization costs. These observations highlight the critical importance of treatment effect in driving modeling outcomes in these HF studies, and the importance of carefully assessing treatment effects when building cost-effectiveness models for HF management.

Beyond drivers relating to treatment, within the top 5 ranked model drivers, baseline CV mortality risk more often emerged than factors related to hospitalizations, reinforcing a need to monitor at-risk patients along their trajectory of decline [14]. Although baseline mortality risk (all-cause or CV mortality) and cost of hospitalization were among the first-reported, main model drivers, no study modeled rehospitalization change explicitly. Given the simplicity of the model structures employed, it is not unexpected that these were among the main model drivers. Indeed, because most of the models employed in HF are simply alive $>$ (hospitalization) $>$ death, it is easy to understand why treatment effects on mortality and baseline CV mortality were among the most commonly identified drivers. We therefore believe there is an unmet need to develop more complex models of cost-effectiveness and risk around rehospitalization, to educate healthcare providers and payers. This is particularly relevant as in recent years the USA has increased efforts to reduce 30-day readmission rates, by imposing financial penalties if hospitals have a high readmissions ratio [15].

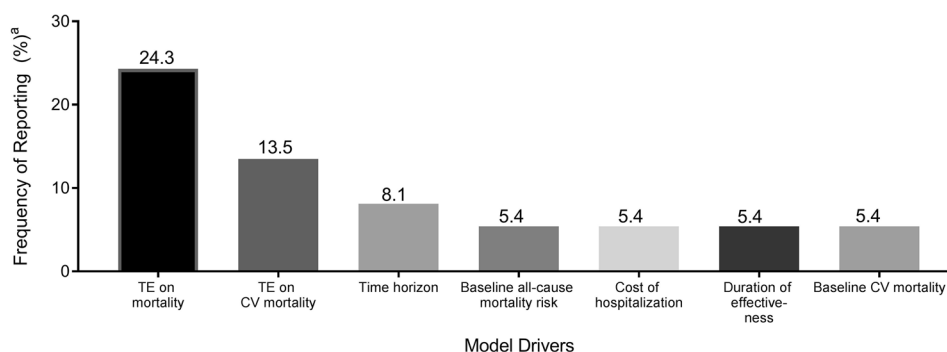
By highlighting the key model drivers, we hope to inform future efforts in building cost-effectiveness models in HF management. However, a key failing within the cost-effectiveness models was the general lack of studies carrying out

Fig. 1 Key model drivers: **a** top 5 ranked model drivers and **b** first-reported model drivers. *CV* cardiovascular, *TE* treatment effect

(a) In Top 5 Ranked Model Drivers



(b) First-Reported Model Driver



^a based on reporting from tornado diagrams or tables, or study conclusions regarding sensitivity analyses (where available), in > 5% of relevant studies

any type of VOI, with just two studies reporting EVPI and none reporting an expected value of partially perfect information (EVPPI) or expected value of sample information (EVSII). VOI is an important tool for studying uncertainty associated with a coverage decision and its implications, and can help to estimate the value of collecting additional data to reduce decision uncertainty [16, 17]. Authors have noted that application of VOI analyses remains limited, and that a better awareness is needed to align VOI assessments with the needs of decision makers [17]. Our study reflects this conclusion with regard to VOI analysis within recent cost-effectiveness models for HF.

The studies identified through our systematic approach typically employed a Markov or trial-based analytic approach, with most models displaying sufficient detail with regard to model perspectives and discount rates. However, more than half of the retrieved studies lacked transparency in important features of their analyses, including aspects of model structure, software used, simulations carried out, and completeness of results reported. Reassuringly, when more recent studies were assessed, we noted that transparency was generally increased. Collectively, we feel improved adherence with CHEERS standards and other similar reporting

guidelines [10, 18] needs to be upheld, alongside a wider open-source modeling initiative for publication of cost-effectiveness models, to allow for full understanding of the methodologies employed.

The current study should be interpreted in light of some limitations. As noted above, there was wide variation in the reporting of key information within studies, and our analysis is therefore limited to the information provided in the published manuscripts. Half of the studies were carried out in the UK or USA; therefore we cannot assume generalizability in other cohorts. Of the 23 studies published in the past 3 years, 74% involved a sacubitril/valsartan comparison with enalapril, assessing cost-effectiveness based on the PARADIGM-HF study. As such, a bias towards certain treatment interventions cannot be overlooked, particularly in most recent years. We have only been able to draw conclusions on model drivers based on the information available in 51% of retrieved studies, which explicitly discussed model drivers within the text of their manuscript with sufficient detail. It was beyond the scope of this brief communication to provide an in depth assessment of the mechanisms behind the relevance/identification of each model driver. Furthermore, we could only assess model drivers as factors

determining the robustness/uncertainty of ICERs within the models described. We therefore believe there are missed opportunities as a result of this lack of user-evaluable information in the design and utility of risk prediction models for the management of HF. In addition, the majority of studies that reported DSA did not specify the type of ranges adopted or used a large variation (i.e., $\geq 20\%$) from the point estimates. As a result, bias could have been introduced in the key drivers identified. In addition, structural uncertainty was not assessed. Further research is needed to conclude on how these model drivers can be used to build better cost-effectiveness models.

In conclusion, treatment effect on mortality was the most frequently identified first-reported model driver across the identified decision analytic models designed for evaluating the cost-effectiveness of HF. Among top 5 ranked model drivers, cost of treatment and utility were most frequently identified. While cost of treatment was a frequently identified model driver, our findings suggest a higher relative importance of treatment effect on baseline mortality outcomes in cost-effectiveness models of HF.

Acknowledgements We would like to thank Mitesh Nakum, MSc. (formerly of Curo Payer Evidence, part of the Envision Pharma Group) for his assistance with aspects of the manuscript's development.

Author contributions All authors fulfill International Committee of Medical Journal Editors (ICMJE) authorship criteria. GLDT and GG planned and designed this study. GLDT and SC conducted the review, screened the studies, and extracted the information. AB and HW were involved in the data analysis. All authors were involved in aspects of data interpretation. All authors were involved in either drafting of the manuscript or critical appraisal of the manuscript. All authors approved the final draft of the manuscript for submission and take responsibility for the manuscript content.

Compliance with Ethical Standards

Data availability Data sharing is not applicable to this article as no datasets were generated or analyzed during the current review.

Informed consent For this type of study, formal consent is not required.

Funding This study was funded by Amgen, Thousand Oaks, USA.

Conflict of interest GLDT was an employee of Amgen Inc when the study was conducted. SC was an intern student at Amgen Inc when the study was conducted. HW, AB, and GG are employees of Amgen. KLB is an employee of Envision Pharma Group, who were contracted by Amgen in relation to this project.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. *Int J Cardiol.* 2014;171:368–76.
2. Savarese G, Lund LH. Global public health burden of heart failure. *Card Fail Rev.* 2017;3:7–11.
3. Lesyuk W, Kriza C, Kolominsky-Rabas P. Cost-of-illness studies in heart failure: a systematic review 2004–2016. *BMC Cardiovasc Dis.* 2018;18:74.
4. Roth GA, Forouzanfar MH, Moran AE, Barber R, Nguyen G, Feigin VL, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. *N Engl J Med.* 2015;372:1333–41.
5. Burnett H, Earley A, Voors AA, Senni M, McMurray JJ, Deschaseaux C, et al. Thirty years of evidence on the efficacy of drug treatments for chronic heart failure with reduced ejection fraction: a network meta-analysis. *Circ Heart Fail.* 2017;10:e003529.
6. Deedwania PC, Carbajal E. Evidence-based therapy for heart failure. *Med Clin N Am.* 2012;96:915–31.
7. Raj L, Adhyaru B. An evidence-based review of recent advances in therapy for heart failure with reduced ejection fraction (HFREF). *Postgrad Med J.* 2016;92:726–34.
8. Reed BN, Sueta CA. A practical guide for the treatment of symptomatic heart failure with reduced ejection fraction (HFREF). *Curr Cardiol Rev.* 2015;11:23–32.
9. Di Tanna GL, Bychenkova A, O'Neill F, Wirtz HS, Miller P, Ó Hartaigh B, et al. Evaluating cost-effectiveness models for pharmacologic interventions in adults with heart failure: a systematic literature review. *Pharmacoeconomics.* 2019;37:359–89.
10. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)-explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health.* 2013;16:231–50.
11. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Value Health.* 2012;15:843–50.
12. Polistena B, Maggioni AP, Oliva F, Spandonaro F. Cost-effectiveness and budget impact of ivabradine in heart failure therapy. *G Ital Cardiol (Rome).* 2014;15:626–33.
13. Ollendorf DA, Sandhu AT, Pearson SD. Sacubitril-valsartan for the treatment of heart failure: effectiveness and value. *JAMA Intern Med.* 2016;176:249–50.
14. Lee DS, Alba AC. Risks and benefits of risk prediction in acute heart failure. *JACC Heart Fail.* 2015;3:748–50.
15. McIlvennan CK, Eapen ZJ, Allen LA. Hospital readmissions reduction program. *Circulation.* 2015;131:1796–803.
16. Bindels J, Ramaekers B, Ramos IC, Mohseninejad L, Knies S, Grutters J, et al. Use of value of information in healthcare decision making: exploring multiple perspectives. *Pharmacoeconomics.* 2016;34:315–22.
17. Tuffaha HW, Gordon LG, Scuffham PA. Value of information analysis in healthcare: a review of principles and applications. *J Med Econ.* 2014;17:377–83.
18. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine recommendations from the second panel on cost-effectiveness in health and medicine. *JAMA.* 2016;316:1093–103.