THE RESTRICTED MEAN SURVIVAL TIME AS A TOOL FOR RANKING COMPARATIVE OUTCOMES IN A NARRATIVE REVIEW THAT EVALUATES A NETWORK OF RANDOMISED TRIALS: AN EXAMPLE BASED ON PCSK9 INHIBITORS

Andrea Messori[1], Laura Bartoli[1], Marco Chiumente[2], Daniele Mengato[2], and Sabrina Trippoli[1]

[1]Tuscany Region
[2]Italian Society for Clinical Pharmacy and Therapeutics, Milan, Italy

June 8, 2020

Abstract

Evolocumab and alirocumab have been approved based on a randomised trial. We analysed the outcomes reported in the two trials to develop an original framework of comparative assessment that uses the restricted mean survival time (RMST). The objective was to show that, in the context of a narrative review, the RMST can be an efficient though simple tool to make indirect comparisons. For each cohort of patients (13,784 patients given evolocumab, 9,462 given alirocumab, 23,242 controls; results expressed in months of event-free survival), we determined the values of RMST with 95% confidence intervals [CI] (evolocumab: 33.48, 95%CI: 33.45 to 33.50; alirocumab: 33.89, 95%CI: 36.86 to 33.92). These results, along with those of the control groups, were analysed and interpreted narratively. A univariate statistics was conducted; no network meta-analysis was done. The experience presented herein indicates that a framework of evidence assessment focused on RMST is a worthwhile option.

THE RESTRICTED MEAN SURVIVAL TIME AS A TOOL FOR RANKING COMPARATIVE OUTCOMES IN A NARRATIVE REVIEW THAT EVALUATES A NETWORK OF RANDOMISED TRIALS: AN EXAMPLE BASED ON PCSK9 INHIBITORS

Andrea Messori PharmD,[1] Laura Bartoli PharmD,[1] Marco Chiumente PharmD,[2]

Daniele Mengato PharmD,[2] Sabrina Trippoli, PharmD[1]

[1]HTA Unit, Toscana Region Health Service, Toscana Region,

[2]Scientific Direction, Italian Society for Clinical Pharmacy and Therapeutics,

Corresponding author:

Andrea Messori, PharmD

HTA Unit, Regional Health Service

Regione Toscana

Via Alderotti 46

50135 Firenze

Italy

Email:*andrea.messori.it@gmail.com*, Phone: +39-338-9513583

Total word count of the manuscript (excluding the title page, abstract, references, tables, and figures legends): 1,999 (max. 2,000).

ABSTRACT (word count 148, max 150)

Evolocumab and alirocumab have been approved based on a randomised trial. We analysed the outcomes reported in the two trials to develop an original framework of comparative assessment that uses the restricted mean survival time (RMST). The objective was to show that, in the context of a narrative review, the RMST can be an efficient though simple tool to make indirect comparisons. For each cohort of patients (13,784 patients given evolocumab, 9,462 given alirocumab, 23,242 controls; results expressed in months of event-free survival), we determined the values of RMST with 95% confidence intervals [CI] (evolocumab: 33.48, 95%CI: 33.45 to 33.50; alirocumab: 33.89, 95%CI: 36.86 to 33.92). These results, along with those of the control groups, were analysed and interpreted narratively. A univariate statistics was conducted; no network meta-analysis was done. The experience presented herein indicates that a framework of evidence assessment focused on RMST is a worthwhile option.

KEYWORDS: evolocumab; alirocumab; restricted mean survival time; Kaplan-Meier; network meta-analysis.

What is already known about this subject : Both meta-analytical and descriptive approaches can be used to summarise the evidence that results from 2 or more clinical trials

What this study adds : expanding the role of narrative reviews integrated by the estimation of restricted mean survival times can be worthwhile as opposed to the use of network meta-analysis based on the hazard ratio.

INTRODUCTION

The restricted mean survival time (RMST) is a relatively new parameter proposed to improve the analysis of survival curves.[1-]comparison with traditional analyses based on hazard ratio (HR) and medians, the RMST has the advantage of capturing the overall shape of survival curve, including the so-called "right tail" of long-term survivors. Since RMST and area under the survival curve (AUC) are identical parameters, the RMST can be rather easily calculated taking advantage of a model-independent method of pharmacokinetic AUC calculation,[5] that allows for an extreme simplification of the estimation procedure. Recently, this original model-independent method of RMST calculation[6,7] has been employed in patients with different types of cancer.[7-9]

A technique for estimating the 95% confidence interval (CI) of RMST has also been made available extending the application of RMST in the analyses of survival curves.[10] Likewise, the appropriate method for determining the statistical significance in the comparison between the RMSTs of two different treatments has been clarified.[11]

Patients who have had an acute coronary syndrome are at high risk for recurrent ischemic cardiovascular events. In this context, proprotein convertase subtilisin–kexin type 9 (PCSK9) inhibitors[12,13] represent an important innovation that is now available and reimbursed in many countries. Briefly, the enzyme PCSK9 promotes degradation of LDL receptors, thereby diminishing the clearance of LDL from the circulation. Studies have shown that mutations conveying gain or loss of function of PCSK9 result in a higher or lower level of LDL cholesterol, respectively, which determines a higher or lower risk of incident coronary heart disease.[14,15]These findings have led to the development of monoclonal antibodies to PCSK9 that produce substantial reductions in LDL cholesterol when administered alone or with a statin.[12,13] Currently there are two approved drugs based on this mechanism of action: evolocumab and alirocumab.

In the present paper, we studied the outcomes observed with these two drugs by focusing our analyses on the composite end-point of cardiovascular event or death. The outcomes were compared between evolocumab, alirocumab, and placebo. Our comparisons were based on RMST. The four cohorts included in our analyses, given PCSK9 inhibitor or placebo, were obtained from the above-mentioned randomized studies published by Sabatine et al.[12] for evolocumab (FOURIER trial) and Schwartz et al.[13] for alirocumab (ODYSSEY trial).

PATIENTS AND METHODS

Study design. Our study was aimed at applying the RMST to the patient cohorts enrolled in the FOURIER and ODYSSEY randomized, placebo-controlled trials. In both studies, the composite end-point of cardiovascular event or death was analyzed according to a standard time-to-event statistics (Kaplan-Meier) that generated the respective survival curves for the treatment group and the controls. The values of RMST were determined from each of these time-to-event curves by model-independent methods.

Patients. In the FOURIER trial, the cohort treated with evolocumab consisted of 13,784 patients while the controls were 13,780. In the ODYSSEY trial, the cohort treated with alirocumab consisted of 9,462 patients and also the controls were 9,462. The follow-up lasted up to 36 months for FOURIER and 48 months for ODYSSEY. Further details on these cohorts can be found in the original studies.[12,13]

Statistical analysis aimed at estimating RMST. The model-independent values of RMST (units, months per patient) were determined according to the AUC calculation previously described.[6,7] Briefly, our procedure retrieves the published graphs of the survival curves, then estimates the survival percentage-vs-time data points with a digitizer (WebPlotDigitizer. https://automeris.io/WebPlotDigitizer), and finally calculates the model-independent values of RMST. An Excel datasheet was used to apply the trapezoidal rule (YouTube video. How to find the area under the curve in Excel, https://www.youtube.com/watch?v=Ke8W9U5SXf4, accessed 5 June 2020). To improve the phase of data input, the procedure was transferred into an executable file compatible with Windows (both 32 and 64-bit versions). Each survival curve was truncated ("restricted") at the last time point in the follow-up (the so-called "milestone" or t*). Under the assumption that RMST can appropriately describe the time-to-event survival pattern in these patients, our analysis examined the 4 Kaplan-Meier curves of the two trials. To ensure comparability, all curves were truncated at t*=36 months. The 95% confidence intervals [CIs] for RMST were calculated for each curve as previously described.[10.] In the comparison between two RMST, the statistical significance was determined according to the equations reported by Messori et al.[11]

Ranking of the treatments according to RMST values . The four cohorts were ranked according to the respective values of RMST in descending order. It should be noted that this non-parametric approach of elementary ranking resembles the one commonly employed in network meta-analysis.[16] Of course, the practical usefulness of this ranking is -in general- greater when the treatments are three or, better, many more than three.[17]

Evaluation of the indirect comparison of alirocumab vs evolucumab based on the HR.

The indirect comparison between alirocumab and evolocumab was designed according to the simplest form of network meta-analysis developed thus far.[16,18] The HR was the outcome measure. Statistical calculations were carried out according to the ITC software.[18]

RESULTS

Descriptive analysis of the evidence integrated by the calculation of RMSTs

A total of 4 separate analyses were performed (treatment group of the ODYSSEY trial; control group of the ODYSSEY trial; treatment group of the FOURIER trial; control group of the FOURIER trial). The clinical material was represented by the Kaplan Meier time-to-event curves published in the two pivotal trials. In both trials, the event was represented by a composite end-point: "a composite of death from coronary heart disease, nonfatal myocardial infarction, fatal or nonfatal ischemic stroke, or unstable angina requiring hospitalization" in the ODYSSEY trial; "a

composite of cardiovascular death, myocardial infarction, stroke, hospitalization for unstableangina, or coronary revascularization" in the FOURIER trial.

The 4 Kaplan-Meier curves allowed for the estimation of the values RMST reported in Table 1 along with their respective rankings. Figure 1 shows the 4 curves generated by the estimation procedure.

Of course, the analysis based on the RMSTs confirmed the results obtained in the pivotal trials. As regards the FOURIER trial, the direct comparison at 36 mos of evolocumab (RMST, 33.48 mos; 95%CI: 33.45 to 33.50) vs placebo (RMST, 33.10 mos; 95%CI: 33.07 to 33.13) showed a significant gain in event-free survival in favor of the active treatment (0.38 months; p from the two RMST values <0.001). Likewise, the direct comparison at 36 mos of alirocumab (RMST, 33.89; 95%CI: 33.86 to 33.92) vs placebo (RMST. 33.47 mos; 95%CI: 33.44 to 33.50) showed a significant event-survival gain of 0.42 mos in favour of alirocumab (p from the two RMST values <0.001). The same direct comparison assessed at 48 mos gave very similar results (RMST of 44.62 mos; 95%CI: 44.58 to 44.65 for alirocumab vs 44.14 mos; 95%CI: 44.10 to 44.18 for placebo; gain of 0.48 mos in favour of alirocumab; p from the two RMST values <0.001).

As regards the indirect comparison at 36 mos of evolocumab vs alirocumab, the RMST of alirocumab (33.89 mos; 95%CI: 33.86 to 33.92) was longer than that of evolocumab (33.48 mos; 95%CI: 33.45 to 33.50) with a significant difference (p<0.001).

One finding in Table 1 raises some controversy. The controls of the ODYSSEY trial showed a slightly better RMST than the treatment group of the FOURIER trial (33.55 vs 33.48 mos). Although this difference (0.07 mos, i.e. 2.1 days) reaches statistical significance (p<0.05) presumably because of the extremely large populations involved, it does not appear to have any clinical relevance.

Evaluation of the indirect comparison of alirocumab vs evolocumab based on the hazard ratio.

On the basis of the composite end-points mentioned above, the HR was 0.85 (95%CI, 0.78 to 0.93; p<0.001) as reported in the ODYSSEY trial for alirocumab and 0.85 (95%CI, 0.79 to 0.92; p<0.001) as reported in the FOURIER trial for evolocumab. From these values of HR, the ITC software determined a HR of 1.00 (95%CI, 0.89 to 1.12) for the indirect comparison of alirocumab vs evolocumab.

DISCUSSION

The model-independent RMST has an important advantage because its applicability has no exceptions in analysing time-to-event curves. In contrast, the median is not computable when only a few events have occurred and survival remains above 50%. Another advantage of the RMST is that it is an absolute parameter and is determined on a scale of time; the HR is instead a relative parameter represented by a dimensionless number. In terms of ease of applicability, the RMST (in its model independent version) does not require any access to the data base of individual patients since the starting material is the graph of the Kaplan-Meier curve; furthermore, no advanced statistical software is necessary. Apart from these essential points, the comparison of the advantages and disadvantages of RMST vs median and HR ratio is a complex issue on which excellent articles have been published.[1-4,19]

The present paper has essentially a provocative purpose because, while many standard approaches commonly used in evidence-based medicine are unquestionable, some specific points related to the meta-analytic or descriptive nature of a comparative analysis still require to be explored.

One finding presented in this paper is that the interpretation suggested by a standard indirect-treatment-comparison (ITC) analysis (i.e. a meta-analytic approach) is qualitatively different from that of a descriptive analysis integrated by the values of RMST (i.e. a narrative approach). The main explanation is that the indirect (meta-analyitic) comparison is typically based on a relative outcome measure (namely, the HR) while the descriptive analysis, as presented herein, is based on an absolute outcome measure (the RMST). Hence, making comparisons is difficult.

This is not, however, the only difference. In general, no evidence-based analysis should be used to directly compare the efficacy of two treatments across heterogeneous studies; in other words, in order to compare two

drugs (in this case, evolocumab vs. alirocumab), it is necessary to prove that the two trials are comparable, and information should be provided about the trial design, patients' inclusion/exclusion, etc. The FOURIER and ODYSSEY trials, however, were extremely homogeneous with one another. Hence, applying the ITC approach to these two trials is likely to be an appropriate choice, and consequently the HR for the indirect comparison of alirocumab vs evolocumab (1.00; 95%CI, 0.89 to 1.12) is unquestionable.

On the other hand, from the descriptive analysis reported in Table 1 one unexpected result emerged. In ranking the four cohorts according to the RMST, the controls of the ODYSSEY trial showed a slightly better RMST than the treatment group of the FOURIER trial (33.55 vs 33.48 mos). The difference was clinically negligible but, curiously enough, reached statistical significance (probably owing to the very large size of the patient populations involved). In this case, the descriptive analysis, that has a disadvantage in that it loses the link that randomisation determines within a single trial, could be thought to be less sensitive to very specific aspects of the patient populations involved, but -unexpectedly- captured a finding that the ITC, with its approach based on an overall pooling of control groups, had missed. Of course, no generalization can be made on the basis of this single example.

The advantages and disadvantages of the two approaches are subtle. However, the present example shows that, despite its disadvantages (no adjustment for the effect of a randomised design), a well-designed narrative analysis integrated with the estimation of RMSTs can represent an alternative to network meta-analysis, particularly when the treatments under comparison are numerous and the network meta-analysis can be confounded by the excess of pairwise comparisons.[20] Furthermore, one point of controversy is that, unlike the case of RMSTs, all survival meta-analyses disregard the length of follow-up because they focus on the HR that is independent on the follow-up.

On the side of the expected confirmation of the performance of these methods, it should be emphasized that the within-trial analyses comparing the treatment group vs the control group gave systematically the same results irrespective of the method of analysis adopted.

In conclusion, our paper has pursued the objective to confirm the basic evidence concerning PCSK9 inhibitors, but, more importantly, has identified an original framework for studying comparative effectiveness based on the RMST that might deserve to be further explored.

Sources of Funding

There was no specific funding for the present paper.

Disclosures

MC regularly works at the Scientific Direction of SIFACT. The other authors declared no conflict of interest.

Data Availability Statement

All data included in the analyses presented herein are available from the authors upon request.

.

REFERENCES

1. Trinquart L, Jacot J, Conner SC, Porcher R. Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials. J Clin Oncol. 2016 May 20;34(15):1813-9. doi: 10.1200/JCO.2015.64.2488.
2. Royston P, Parmar MK. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. Stat Med. 2011 Aug 30;30(19):2409-21. doi:10.1002/sim.4274.
3. McCaw ZR, Yin G, Wei LJ. Using the Restricted Mean Survival Time Difference as an Alternative to the Hazard Ratio for Analyzing Clinical Cardiovascular Studies. Circulation. 2019 Oct 22;140(17):1366-1368. doi:10.1161/CIRCULATIONAHA.119.040680.

4. Kloecker DE, Davies MJ, Khunti K, Zaccardi F. Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes. Ann Intern Med. 2020 Mar 24. doi:10.7326/M19-3286. [Epub ahead of print]
5. Gibaldi M, Perrier D Pharmacokinetics (2nd Edition). Marcel Dekker, New York, 1982.
6. Damuzzo V, Agnoletto L, Leonardi L, Chiumente M, Mengato D, Messori A. Analysis of Survival Curves: Statistical Methods Accounting for the Presence of Long-Term Survivors. Front Oncol. 2019 Jun 4;9:453. doi: 10.3389/fonc.2019.00453.
7. Messori A, Damuzzo V, Agnoletto L, Leonardi L, Chiumente M, Mengato D. A model-independent method to determine restricted mean survival time in the analysis of survival curves. SN Comprehensive Clinical Medicine 2019 (in press), https://doi.org/10.1007/s42399-019-00199-7,
8. Chiumente M, Mengato D, Messori A. Tisagenlecleucel in non-Hodgkin lymphoma: the restricted mean survival time as a tool for estimating progression-free life expectancy better than the median Acta Haematologica 2020 (in press), preprint available at http://www.osservatorioinnovazione.net/papers/aha-accepted-paper.pdf
9. Messori A, Damuzzo V, Leonardi L, Agnoletto L, Chiumente M, Mengato D. CAR-T treatment: determining the survival gain in patients with relapsed or refractory diffuse large B-cell lymphoma. Clinical Lymphoma, Myeloma and Leukemia, 2020 (in press), preprint at http://www.osservatorioinnovazione.net/papers/clml-accepted-paper.pdf
10. Messori A, Bartoli B, Damuzzo V, Chiumente M, Mengato D. Restricted mean survival time estimated according to model-independent methods: procedure for calculating its 95% confidence interval. Abstract, Sifact Meeting, 21 January 2020, url https://osf.io/qemhc/? accessed 25 March 2020.
11. Messori A, Caccese E, D'Avella MC. Estimating the 95% confidence interval for survival gain between an experimental anti-cancer treatment and a control. Ther Adv Med Oncol. 2017 Nov;9(11):721-723. doi: 10.1177/1758834017731084.
12. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease (FOURIER trial). N Engl J Med 2017; 376: 1713-22.
13. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, R. Diaz R et al. Alirocumab and Cardiovascular Outcomes after Acute Coronary Syndrome (ODISSEY trial). N Engl J Med 2018;379:2097-107. DOI: 10.1056/NEJMoa1801174.
14. Cohen JC, Boerwinkle E, Mosley TH Jr,Hobbs HH. Sequence variations in*PCSK9* , low LDL, and protection against coronary heart disease. N Engl J Med 2006; 354:1264-72.
15. Abifadel M, Varret M, Rabes JP, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 2003; 34: 154-6.
16. Dias S, Caldwell DM. Network meta-analysis explained. Arch Dis Child Fetal Neonatal Ed. 2019 Jan;104(1):F8-F12. doi: 10.1136/archdischild-2018-315224.
17. ESTAR (evidenze sui farmaci). Numero 1, 29 Maggio 2020 (https://osf.io/gebnh/); Numero 2, 5 Giugno 2020 (https://osf.io/dp92f/) in Italian language, both available at http://portale-mav.estav-centro.toscana.it/noteestar.htm
18. Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.
19. Kloecker DE, Davies MJ, Khunti K, Zaccardi F. Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes. Ann Intern Med. 2020 Mar 24. doi:10.7326/M19-3286. [Epub ahead of print]
20. Messori A. Generalised anxiety disorder: a network meta-analysis with 462 binary comparisons. Open Science Framework, https://osf.io/78zq5/? , DOI 10.17605/OSF.IO/WC4K3

Table 1. Characteristics of the four cohorts and values of RMST at t*=36 months estimated from the time-to-event curves	**Table 1. Characteristics of the four cohorts and values of RMST at t*=36 months estimated from the time-to-event curves**	**Table 1. Characteristics of the four cohorts and values of RMST at t*=36 months estimated from the time-to-event curves**	**Table 1. Characteristics of the four cohorts and values of RMST at t*=36 months estimated from the time-to-event curves**
Data-set	**No. of patients**	**RMST (mos) with 95% confidence interval**	**Rank**
Alirocumab	9,462	33.89 (33.86 to 33.92)§	1
Controls (ODYSSEY trial)	9,462	33.55 (33.52 to 33.58)§§	2
Evolocumab	13,784	33.48 (33.45 to 33.50)*	3
Controls (FOURIER)	13,780	33.10 (33.07 to 33.137)**	4
§SEM = 0.0153 §§SEM =0.0159 *SEM = 0.0128 **SEM=0.0139 Abbreviations: RMST, restricted mean survival time; mos, months; t*, milestone; SEM, standard error of the mean.	§SEM = 0.0153 §§SEM =0.0159 *SEM = 0.0128 **SEM=0.0139 Abbreviations: RMST, restricted mean survival time; mos, months; t*, milestone; SEM, standard error of the mean.	§SEM = 0.0153 §§SEM =0.0159 *SEM = 0.0128 **SEM=0.0139 Abbreviations: RMST, restricted mean survival time; mos, months; t*, milestone; SEM, standard error of the mean.	§SEM = 0.0153 §§SEM =0.0159 *SEM = 0.0128 **SEM=0.0139 Abbreviations: RMST, restricted mean survival time; mos, months; t*, milestone; SEM, standard error of the mean.

Legends for figures

Figure 1. Four Kaplan-Meier curves were included in our analysis. For each individual curve, the survival curve fitting procedure generated the small red circles that, in the two panels, are superimposed to the original curves as published in the respective articles. According to these red circles, Panels A (evolocumab treatment group and controls of FOURIER trial) and B (alirocumab treatment group and controls of ODYSSEY trial) show the computer-generated curves based on their respective x-vs-y data pairs. For each curve, the values of cumulative incidence were firstly converted into event-free values; the AUCs were then estimated from these x-vs-y data pairs by application of the trapezoidal rule. In the model-independent approach, RMST is known to be identical to AUC.