The Restricted Mean Survival Time as a Tool for Ranking Comparative Outcomes in a Narrative Review that Evaluates a Network of Randomized Trials: An Example Based on PCSK9 Inhibitors

Andrea Messori1, Laura Bartoli1, Marco Chiumente2, Daniele Mengato2, Sabrina Trippoli1

Accepted: 22 September 2020
© Springer Nature Switzerland AG 2020

Abstract

Introduction: On the basis of a randomized trial, evolocumab and alirocumab have been approved in patients with cardiovascular disease. The evidence on these two agents has been studied through different methods of analysis that span from narrative approaches to network meta-analysis. In the present study, we assessed the performance of a narrative approach combined with the application of the restricted mean survival time (RMST).

Methods: We studied the two pivotal placebo-controlled trials focused on evolocumab and alirocumab. Our original framework of comparative assessment employed the RMST. Our objective was to show that in the context of a narrative review, the RMST can be an efficient although simple tool to make indirect comparisons. The endpoint was event-free survival, expressed in months.

Results: For each cohort of patients (13,784 patients administered evolocumab, 9462 patients administered alirocumab, 23,242 controls), we determined the RMST values with 95% confidence intervals (CI) [evolocumab: 33.60 months, 95% CI 33.46–33.74; alirocumab: 34.07 months, 95% CI 33.92–34.22]. These results, along with those of the control groups, were analyzed and interpreted narratively. Univariate statistics were conducted, but no network meta-analysis was performed.

Conclusion: The experience presented herein indicates that a framework of evidence assessment focused on the RMST is a worthwhile option. Our study is in line with the growing literature that has recently emphasized the methodological advantages of the RMST.

1 Introduction

The restricted mean survival time (RMST) is a relatively new parameter proposed to improve the analysis of survival curves [1–4]. In comparison with traditional analyses based on the hazard ratio (HR) and medians, the RMST has the advantage of capturing the overall shape of the survival curve, including the so-called ‘right tail’ of long-term survivors. Furthermore, the RMST is applicable even when the assumption of proportional hazards is violated.

In recent times, a wide range of literature [5, 6] has consistently emphasized the advantages of the RMST.

There are numerous methods for estimating the RMST. The most widely used, based on the R platform (package survRM2 [7]), requires that individual survival times (and outcomes) are reconstructed from the Kaplan-Meier curve graph (which is therefore digitized for this purpose). Several techniques have been published that carry out this reconstruction [8–11]. On the other hand, if no reconstruction of individual patient data is performed, the RMST can be directly estimated from the area under the survival curve.
curve (AUC) [12], which is measured by application of
the trapezoidal rule [13]; this is the so-called model-inde-
dependent method.

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 [14–17] represent an impor-
tant innovation that is now available and reimbursed in many
countries. Briefly, the enzyme PCSK9 promotes degra-
dation of low-density lipoprotein (LDL) receptors, thereby
diminishes 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. These
findings have led to the development of monoclonal anti-
odies to PCSK9 that produce substantial reductions in
LDL cholesterol when administered alone or with a statin
[14–17]. 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 endpoint of cardiovascular event or death.
The outcomes were compared between evolocumab, aliro-
cumab, and placebo, and our comparisons were based on
RMST. The four cohorts included in our analyses, admin-
istered a PCSK9 inhibitor or placebo, were obtained from
the above-mentioned randomized studies published by
Sabatine et al. [14] for evolocumab (FOURIER trial) and

2 Patients and Methods

2.1 Study Design

Our study was aimed at applying the RMST to the patient
cohorts enrolled in the FOURIER and ODYSSEY ran-
donized, placebo-controlled trials. In both studies, the
composite endpoint of cardiovascular event or death was
analyzed according to a standard time-to-event statistics
(Kaplan–Meier) that generated the survival curves for both
the treatment group and the controls. The RMST values were
determined from each of these time-to-event curves using
the R platform [7] and, for comparison purposes, the AUC
method [12].

2.2 Patients

In the FOURIER trial, the cohort treated with evolocumab
consisted of 13,784 patients, while the controls included
13,780 subjects. In the ODYSSEY trial, the cohort treated
with alirocumab consisted of 9462 patients and the con-
trols included 9462 subjects. 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 [14, 15].

2.3 Estimation of Restricted Mean Survival Time
(RMST)

In our primary analysis, the RMST values (with their respec-
tive 95% confidence intervals [CIs]) were estimated through
the ‘survRM2’ package written in R [7]. The first steps of
each survival analysis were (1) retrieving the published
graphs of the survival curves; (2) estimating the survival
percentage-versus-time data points with a digitizer (WebPlotDigitizer; https://automeris.io/WebPlotDigitizer); and
(3) reconstructing individual patient data. For this latter
purpose, we used an iterative least squares technique [10]
that performs this reconstruction from the total number of
events, the total number of at-risk patients, and the distribu-
tion of at-risk patients over time. Deriving a CI for RMST
requires distinguishing censored observations from uncen-
sored observations; hence, when reconstructing individual
patient data, reconstruction required not only determining
the survival time but also whether the observation was cen-
sored or not. The total number of events was retrieved from
the text or the figures reported in the original trial, and the
total number of censored patients was determined as the dif-
ference between the total number of enrolled patients minus
the number of events minus the residual number of at-risk
patients at the closure date of the study. Both these estimates
were distributed across the time intervals of the survival
curve according to a criterion of iterative minimization
of root mean squared error [18] between fitted points and
original points of the survival curve. After reconstructing
each curve, the survRM2 package was run to determine the
RMST value and its 95% CI.

Finally, to verify the consistency of this process, for each
curve we also calculated the point estimate of RMST using
the AUC method [12]. The AUC method has the advantage
that no reconstruction of individual patient data is needed,
but also the disadvantage that no CIs are provided. A Micro-
soft Excel datasheet (Microsoft Corporation, Redmond,
WA, USA) was used to apply the trapezoidal rule (YouTube
video—How to find the AUC in Excel; https://www.youtube.
com/watch?v=Ke8W9U58Xf4, accessed 5 June 2020). To
ensure comparability, all curves were truncated after 36
months. In the direct comparisons between two RMSTs,
the statistical significance was determined according to
standard equations [19].
2.4 Ranking of the Treatments According to RMST Values

The four cohorts were ranked according to the respective RMST values, in descending order. It should be noted that this non-parametric approach of elementary ranking resembles the approach commonly employed in network meta-analysis [20]. Of course, the practical usefulness of this ranking is, in general, greater when the treatments are three or, better, many more than three. As regards network meta-analysis of survival studies, different approaches have been proposed (e.g., Bayesian methods, frequentist methods, etc.), but all of these use the HR as the outcome measure owing to the limitations of the median [20].

2.5 Evaluation of the Indirect Comparison of Alirocumab versus Evolocumab Based on the Hazard Ratio and/or the RMST

The indirect comparison between alirocumab and evolocumab was designed according to the simplest form of network meta-analysis of survival data [20, 21], in which the HR is invariably the preferred outcome measure. Our statistical calculations involving the HR for evolocumab versus alirocumab were carried out according to the Indirect Treatment Comparisons (ITC) software [21]. As regards the RMST, we used standard univariate t-testing [19]. Furthermore, we calculated the placebo-adjusted difference of RMST up to 36 months between evolocumab and alirocumab through an indirect comparison.

3 Results

3.1 Descriptive Analysis of the Evidence Integrated by the Calculation of RMSTs

A total of four separate analyses were performed (treatment group of the ODYSSEY trial; control group of the ODYSSEY trial; treatment group of the FOURIER trial; and 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 a composite endpoint: “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 unstable angina, or coronary revascularization” in the FOURIER trial. As regards the four Kaplan–Meier curves, the quality of the information in terms of numbers at risk, number of events, and number of censored patients in each interval was excellent for both trials.

Table 1 shows the RMST values (with 95% CIs), along with their respective rankings. Figure 1 shows the four curves generated by the estimation procedure.

In analyzing separately the two trials, the RMST values were qualitatively in agreement with the results obtained in the pivotal trials. As regards the FOURIER trial, the direct comparison at 36 months of evolocumab (RMST 33.60, 95% CI 33.46–33.74) versus placebo (RMST 33.26 months, 95% CI 33.12–33.40) showed a significant gain in event-free survival in favor of the active treatment (0.34 months; p from the two RMST values < 0.05). Likewise, the direct comparison at 36 months of alirocumab (RMST 34.07 months, 95% CI 33.92–34.22) versus placebo (RMST 33.87 months, 95% CI 33.72–34.02) showed a significant event-free survival gain of 0.20 months in favor of alirocumab (p from the two RMST values < 0.05). The same direct comparison assessed at 48 months gave very similar results (data not shown).

Regarding the indirect comparison at 36 months between the two active treatments, the RMST of alirocumab (34.07 months, 95% CI 33.92–34.22) was longer than that of evolocumab (33.60 months, 95% CI 33.46–33.74), with a significant difference (p < 0.05). Although this difference reaches statistical significance, presumably because of the extremely large populations involved, it does not appear to have any clinical relevance.

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 (33.87 vs. 33.26 months). Although this difference (0.61 months, i.e. 1.4 months) was small, it reached statistical significance (p < 0.05).
determined an HR of 1.00 (95% CI 0.89–1.12) for the indirect comparison of alirocumab versus evolocumab. No analysis based on medians could be made for the purpose of direct or indirect comparisons because the median was not reached in any of the four cohorts.

4 Discussion

In 2019 or 2020, each of eight major international journals (New England Journal of Medicine, The Lancet, Journal of Clinical Oncology, Annals of Internal Medicine, Annals of Oncology, Circulation, Journal of the American College of Cardiology, and Journal of the National Comprehensive Cancer Network) have published at least one methodological paper that emphasized the advantages of the RMST in comparison with traditional outcome measures such as the HR [5]. Overall, during the past 12 months, PubMed has indexed a total of 55 articles that dealt with the RMST [6].

The methodological advantages of the RMST are numerous. For example, its applicability has no exceptions, whereas the median is not computable when only a few events have occurred (and survival remains >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. Apart from these essential points, the comparison of the advantages and disadvantages of the RMST versus the median and HR is a complex issue on which excellent articles have been published [1–6].

In the present case, the primary results of the ODYSSEY and FOURIER trials (HR 0.85, 95% CI 0.78–0.93, \( p < 0.001 \); and HR 0.85, 95% CI 0.79–0.92, \( p < 0.001 \), respectively) convey a message of undisputed superiority (with high statistical significance) of the two PCSK9 inhibitors versus placebo. If outcomes are expressed on the basis of the RMST (ODYSSEY trial: improvement in RMST from 33.87 months to 34.07 months; FOURIER trial: improvement in RMST from 33.26 months to 33.60 months), the differences remain statistically significant (also in consideration of the large patient populations enrolled), but the small magnitude of the incremental benefit emerges more clearly. It should be pointed out that as regards the specific topic of PCSK9 inhibitors, no network meta-analysis has yet been published that compared evolocumab versus alirocumab versus placebo on the basis of the endpoint of all-cause death or the composite of cardiovascular death, myocardial infarction, stroke, hospitalization for unstable angina, or coronary revascularization. The placebo-adjusted difference of RMST up to 36 months between alirocumab and evolocumab was calculated as \(-0.14\) months (\(-0.14\) = 34.07 – 33.87 – 33.60 + 33.26), which shows a worse performance for alirocumab, but

3.2 Evaluation of the Indirect Comparison of Alirocumab versus Evolocumab Based on the Hazard Ratio

On the basis of the composite endpoints mentioned above, the HR was 0.85 (95% CI 0.78–0.93; \( p < 0.001 \)), as reported in the ODYSSEY trial, for alirocumab, and 0.85 (95% CI 0.79–0.92; \( p < 0.001 \)), as reported in the FOURIER trial, for evolocumab. From these HR values, the ITC software

\( \Delta \) Adis

---

“Fig. 1 Four Kaplan–Meier curves were included in our analysis. For each individual curve, the survival curve fitted the 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 the FOURIER trial) and b (alirocumab treatment group and controls of the ODYSSEY trial) show the computer-generated curves based on their respective x versus y data pairs. In both trials, the endpoint is a composite of cardiovascular event or death. In the RMST estimation according to the AUC method, the analysis was not based on the curve of cumulative incidence (as reported in this figure) because the AUC was calculated after changing the y-axis from cumulative incidence to survival probability. AUC area under the survival curve.”
clearly the magnitude of this difference (a few days) is meaningless. Finally, it is worth considering that the difference in the RMST between the two drugs may continue to increase to a clinically meaningful degree in patients receiving long-term treatment.

Our analysis has confirmed that in estimating the RMST, the reconstruction of individual patient data from the Kaplan–Meier curve is the most critical phase. In this context, verifying the agreement between the results based on the reconstruction of individual patient data and those resulting from the AUC method is, in our view, particularly important. As already pointed out, at least three different techniques are available for this purpose [8–11], but as yet they have not been compared with each other. In contrast, the estimation of the RMST from individual patient data is a straightforward phase. In this field, besides the survRM2 package [7], two standard statistical programs contain an option for calculating the RMST (namely, the RMSTREG procedure available in SAS [22], and the Survival2 and Strmst2 commands available in Stata [23]).

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 exploration. One finding presented in this paper is that the interpretation suggested by a traditional meta-analytic approach is different from that suggested by a narrative analysis integrated by the RMST values. The main explanation is that the indirect (meta-analytic) 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) that better reflects the clinical results.

The FOURIER and ODYSSEY trials were quite homogeneous with one another. However, differences between the two populations cannot be excluded because ODYSSEY patients were potentially at higher risk (acute coronary syndrome within the last 12 months) compared with the FOURIER trial (patients with more stable disease). In general, paying close attention to study and patient characteristics is important given the inability of the RMST to deal with heterogeneity.

The present example shows that despite its disadvantages (no adjustment for the effect of a randomized 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 [24]. 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 follow-up.

The question of whether a narrative analysis supported by RMST estimations can already be considered a valid alternative to network meta-analysis remains open. As a matter of fact, the experiences where this replacement has occurred are becoming more and more numerous [25–29], and thus far the performance of this approach has been excellent.

5 Conclusion

Our paper has pursued the objective of confirming the basic evidence regarding 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.

Acknowledgements. This study has been previously reported in a preprint server at the following URL: https://vynm2020.authorex.com/users/331092/articles/45812-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-pcs9-inhibitors-committee-54e753c1704c27bf0902aefa6c4767d9a5c3f81.

Declarations

Funding No specific funding was received for this paper.

Conflicts of interest Marco Chiumente regularly works at the Scientific Direction of SIFACT. Andrea Messori, Laura Bartoli, Daniele Mengato and Sabrina Trippoli declared no conflicts of interest.

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

References


5. Messori A. Eight major international journals have recently published a paper to highlight the methodological advantages.
of the restricted mean survival time. Open Science Framework, 13 July 2020. Available at: https://osf.io/3gj96/. doi:https://doi.org/10.17605/OSF.IO/DY3UH


