Annals of Internal Medicine RESEARCH AND REPORTING METHODS

Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes

David E. Kloecker, MPhil; Melanie J. Davies, MD; Kamlesh Khunti, MD, PhD; and Francesco Zaccardi, MD, PhD

The restricted mean survival time (RMST) has been advocated as an alternative or a supplement to the hazard ratio for reporting the effect of an intervention in a randomized clinical trial. The RMST difference allows quantification of the postponement of an outcome during a specified (restricted) interval and corresponds to the difference between the areas under the 2 survival curves for the intervention and control groups. This article presents examples of the use of the RMST in a research and a clinical context. First, the authors demonstrate how the RMST difference can answer research questions about the efficacy of different treatments. Estimates are presented for the effects of pharmacologic or strategy-driven glucose-lowering interventions for adults with type 2 diabetes from 36 trials and 9 follow-up studies reporting cardiovascular outcomes and mortality. The authors show how these measures may be used to mitigate uncertainty about the efficacy of intensive glucose control. Second, the authors demonstrate how the RMST difference may be used in the setting of a clinical consultation to guide the decision to start or discontinue a treatment. They then discuss the advantages of the RMST over the absolute risk difference, the number needed to treat, and the median survival time difference. They argue that the RMST difference is both easy to interpret and flexible in its application to different settings. Finally, they highlight the major limitations of the RMST, including difficulties in comparing studies of heterogeneous designs and in inferring the long-term effects of treatments using trials of short duration, and summarize the available statistical software for calculating the RMST.

Ann Intern Med. doi:10.7326/M19-3286 **Annals.org**

For author affiliations, see end of text.

This article was published at Annals.org on 24 March 2020.

Delivery of high-quality, patient-centered care requires appropriate interpretation and translation of scientific evidence. Randomized controlled trials (RCTs) are the gold standard of scientific evidence for assessing treatment effects of interventions. In trials, the treatment effect is frequently reported as the hazard ratio (HR). The HR is a comparison between the rates at which events occur in an intervention and a comparison group and, in a Cox regression, is assumed to be constant over time. However, this measure may be difficult to explain to patients (1–3); consequently, interest is increasing in quantifying and reporting treatment effects with alternative, more interpretable metrics (4).

Among these metrics is the restricted mean survival time (RMST), which has been recommended particularly in cardiovascular medicine (5, 6) and oncology (7–9). The RMST difference compares the areas under the 2 survival curves for the intervention and control groups for a specified (restricted) interval. This contrast corresponds to the mean temporal postponement of the outcome in one group compared with the other, with each group-specific RMST quantifying the average delay in the event over the specified time horizon. In a hypothetical trial comparing the effect of a drug versus placebo on the risk for death, an RMST difference of 10 days over 5 years favoring the drug indicates that, on average, patients would survive 10 days longer over 5 years by taking the drug. A key issue for its use, however, is that estimates of RMST differences depend on, and should be interpreted with reference to, the event rate in both groups and the duration of follow-up or, rather, the specified time horizon.

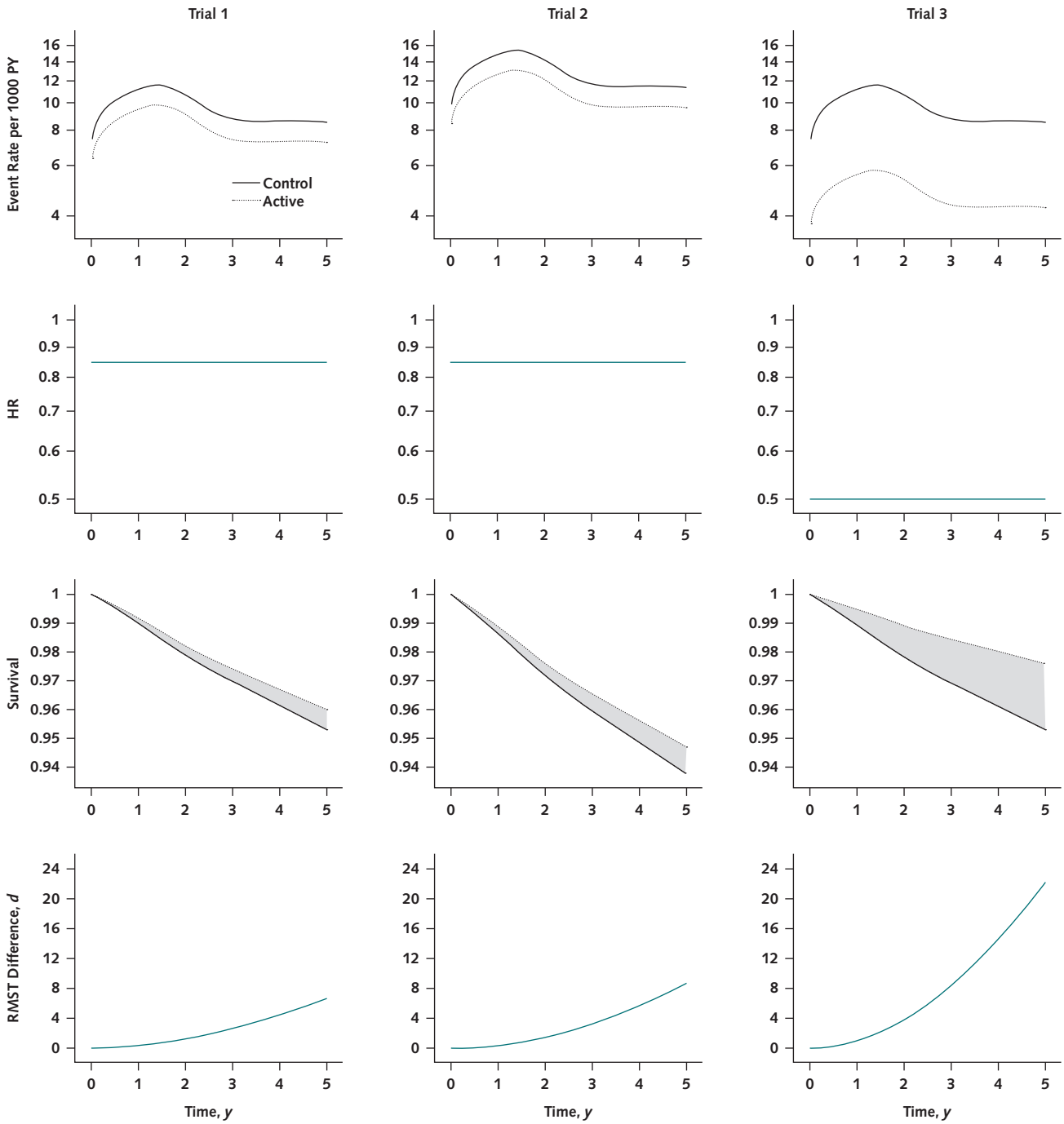
The **Figure** presents 3 simulated trials, each with 10 000 participants randomly assigned to a control or active treatment group. It shows the same data displayed in 4 different ways for each trial: group-specific event rates (that is, hazard rates; first row), the ratios of those hazards (HRs; second row), survival curves (third row), and the RMST difference (fourth row). Like the absolute difference in survival curves, the RMST difference is greater when the event rate in the control group is higher, keeping the HR and follow-up time constant. Moreover, for the same rate in the control group, the lower the event rate in the treatment group (that is, the smaller the HR, constituting a greater effect), the greater the RMST difference. The RMST difference also increases over time if the event rate is lower in the active than the control group (that is, HR <1).

To better understand the potential usefulness of the RMST metric, we present 2 examples related to patients with type 2 diabetes. The first illustrates its use in summarizing results of a body of trials examining the effects of glucose-lowering strategies on cardiovascular outcomes and death. The second is a case example showing how a clinician and patient might use an RMST difference reported from a trial of empagliflozin to enhance personalized decision making. We then review the advantages of the RMST difference over the HR and other measures of absolute risk difference, discuss limitations to the use of the RMST, and report available software for estimating the RMST.

See also:

Web-Only
Supplement

Figure. Relationships between RMST difference and other survival parameters.



Simulated trials with 10 000 participants randomly assigned to a control (n = 5076) or an active (n = 4924) group; each column represents a trial. The first row shows the hazard rates for the control (*solid lines*) and active (*dotted lines*) groups (y-axis on log scale); the distance between the 2 curves is constant (proportional hazards). The constant ratio between the curves, the HR of the active versus the control group, is shown in the second row (y-axis on log scale). The third row shows the survival curves, the cumulative effect of the hazard rates. The fourth row shows the RMST difference (active minus control), corresponding to the difference between the areas under the 2 survival curves. In the first and second columns, the shape of the hazard rate function for the control group (first row) and the HR comparing the active and control groups (second row) are the same (HR, 0.85), but the mean rate in the control group is different (9.6 per 1000 PY in the first column, 12.8 per 1000 PY in the second column). The first and third columns have different HRs (0.85 in the first column and 0.5 in the third). The hazard rate functions for the control groups are the same, both with a mean rate of 9.6 per 1000 PY. The graph shows that the RMST difference increases over time, corresponding to the difference between the survival curves, when the rates are lower in the active group (that is, HR <1). However, for the same HR, the RMST difference is greater if the rate in the control group is higher (second vs. first column), whereas for the same hazard function in the control group, the smaller the HR (that is, the greater the effect), the greater the RMST difference (third vs. first column). Of note, although the HR is 0.5, the RMST difference at 5 years is 22 days. HR = hazard ratio; PY = person-years; RMST = restricted mean survival time.

EXAMPLE 1: DEMONSTRATION OF THE USE OF RMST TO EXAMINE AN EVIDENCE BASE OF GLUCOSE-LOWERING INTERVENTIONS

The following example is a systematic review that shows how the RMST difference may be used to examine results from trials evaluating the potential benefits of intensive glucose reduction. See the **Supplement** (available at Annals.org) for the protocol (**Supplement Table 1**, available at Annals.org), the search strategy (**Supplement Figure 1**, available at Annals.org), and a list of excluded trials (**Supplement Table 2**, available at Annals.org).

In brief, we searched PubMed and the Cochrane Central Register of Controlled Trials from inception until 28 June 2019 for randomized trials of glucose-lowering interventions published in English. Two authors screened search results to identify trials of any duration in adult patients with type 2 diabetes mellitus that randomly assigned participants to a specific treatment or strategy and that reported the Kaplan–Meier plot for the primary outcome of interest and, when available, all-cause mortality. Two authors independently extracted trial data and assessed the risk of bias of the trials (10). First, we extracted data on time and survival probability coordinates from the Kaplan–Meier plots by using Engauge Digitizer (version 10.11) software and, where available, on the total number of events and the number of patients at risk. We used Stata, version 15.0 (StataCorp), for data manipulation, analyses, and graphs, specifically the *ipdfc* command to reconstruct individual-level time-to-event data (11), the *strmst2* command to estimate the RMST difference between the intervention and control group survival curves (positive values indicate postponement of the outcomes in the intervention group) (12), and the *strate* command to determine the outcome rate in the control group. We restricted the time horizon (t*, defined as tau in the *strmst2* command) to the longest possible value, equal to the minimum of the largest observed event time in the trial groups. To assess the accuracy of our estimations, we also calculated the HR with Cox regression for each outcome and compared it with the HR reported in the study.

We identified 45 studies reporting data on 36 RCTs (**Supplement Tables 3** to **5**, available at Annals.org) (13–57). Of note, UKPDS 80 (United Kingdom Prospective Diabetes Study 80) was a follow-up to both UKPDS 33 and UKPDS 34 (52–54). Although the original DCGP (Diabetes Care in General Practice) study was excluded because of a lack of usable figures, we were able to include its follow-up study (DCGP FU) and used the original trial population characteristics to describe the study (23, 58). We rated the overall risk of bias of trials as low (**Supplement Table 6**, available at Annals.org). Studies were published between 1998 and 2019 and included a total of 226 991 participants. The 36 trial cohorts reported a total of 28 744 events for the primary outcome, which in most cases was a composite of nonfatal myocardial infarction, nonfatal stroke, and death from cardiovascular causes (**Supplement Table 7**, available at Annals.org); 35 of these cohorts reported data on 17 704 deaths. The 9 follow-up studies accounted for 15 553 primary outcome events and 11 646 deaths.

Estimation of the RMST was based on studies reporting Kaplan–Meier curves for both the primary outcome and all-cause mortality ($n = 25$), the primary outcome only ($n = 18$), and all-cause mortality only ($n = 3$) (**Supplement Table 3**). The RMST differences and the time horizons (t*) are graphically presented in **Supplement Figure 2** (available at Annals.org) and detailed in **Table 1** for the primary outcomes and in **Supplement Figure 3** (available at Annals.org) and **Table 2** for all-cause mortality. Differences in RMST for the primary outcomes were statistically significant ($P < 0.050$) in 14 of 43 trials. The largest differences, in days, were 694 in the 13-year follow-up of Steno-2 (Intensified Multifactorial Intervention in Patients With Type 2 Diabetes and Microalbuminuria), 523 in the follow-up of UKPDS 34, 452 in UKPDS 34, 362 in DCGP FU, 132 in the 12-year follow-up of VADT (Veterans Affairs Diabetes Trial), and 131 in UKPDS 33. All these studies had a follow-up longer than 10 years. Smaller, statistically significant differences also were found for some of the more recent trials with follow-ups shorter than 10 years. These ranged from 9 days in HARMONY Outcomes (Effect of Albiglutide, When Added to Standard Blood Glucose-Lowering Therapies, on Major Cardiovascular Events in Subjects with Type 2 Diabetes Mellitus) to 32 days in the CANVAS (Canagliflozin Cardiovascular Assessment Study) Program.

The RMST differences for all-cause mortality were statistically significant in 8 of 28 trials (**Table 2**). The largest differences were observed in studies longer than 10 years: 643 days in the 21-year follow-up of Steno-2, 510 days in the follow-up of UKPDS 34, and 203 in the follow-up of UKPDS 33. In studies with a follow-up shorter than 10 years, death occurred earlier in participants randomly assigned to the intensive treatment in ACCORD (Action to Control Cardiovascular Risk in Diabetes) (−16 days). Differences in RMST also were found for recent studies: 12 days in EXSCEL (Exenatide Study of Cardiovascular Event Lowering), 17 days in REWIND (Researching Cardiovascular Events With a Weekly Incretin in Diabetes), 21 days in EMPA-REG OUTCOME (BI 10773 [Empagliflozin] Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients), and 23 days in the CANVAS Program. Broad agreement was found between measures of statistical significance of the estimated RMST differences and the corresponding HRs (**Tables 1** and **2**).

This example demonstrates the usefulness of the RMST to illuminate an area of clinical uncertainty: It gives a different, complementary insight into the importance of long-term glucose-lowering strategies. Although convincing epidemiologic evidence exists of a linear relationship between glucose levels and risk for cardiovascular complications and death (59), uncertainty persists over the need for intensive glucose control in patients with type 2 diabetes (60). In some studies investigating an intensive glucose-lowering strategy, we showed a moderate to large postponement of the primary outcome and death.

Table 1. RMST Difference for Primary Outcomes†

RCT (Reference)‡	RMST, *d*			Follow-up (t*)		Estimated Event Rate in Control Group per 1000 PY (95% CI)	Estimated HR (95% CI)
	Control	Treatment	Difference (95% CI)	Days	Years		
Specific drug treatments							
Dipeptidyl peptidase-4 inhibitors							
CARMELINA (21)	1135	1133	−1.9 (−16.4 to 12.5)	1248	3.4	56.7 (51.6 to 62.3)	1.02 (0.89 to 1.17)
EXAMINE (28)	795	798	3.2 (−9.2 to 15.7)	878	2.4	79.1 (70.8 to 88.4)	0.95 (0.81 to 1.12)
OMNEON (36)	901	901	−0.6 (−9.5 to 8.3)	935	2.6	29.2 (24.4 to 35.1)	1.00 (0.77 to 1.29)
SAVOR-TIMI 53 (45)	844	844	0.2 (−4.3 to 4.7)	882	2.4	37.4 (34.5 to 40.5)	1.00 (0.89 to 1.12)
TECOS (50)	1351	1351	−0.1 (−10.8 to 10.6)	1464	4.0	40.6 (37.9 to 43.5)	0.98 (0.89 to 1.08)
Glucagon-like peptide-1–receptor agonists							
ELIXA (26)	1061	1061	−0.3 (−16.0 to 15.4)	1178	3.2	62.7 (56.7 to 69.2)	1.02 (0.88 to 1.17)
EXSCEL (29)	1656	1668	12.5 (−2.0 to 27.0)	1821	5.0	39.1 (36.6 to 41.8)	0.92 (0.84 to 1.02)
HARMONY Outcomes (30)	762	771	**8.9 (2.6 to 15.3)**	812	2.2	58.3 (53.1 to 64.1)	**0.78 (0.68 to 0.90)**
LEADER (34)	1436	1455	**18.7 (5.1 to 32.3)**	1567	4.3	41.2 (38.2 to 44.4)	**0.86 (0.78 to 0.96)**
PIONEER 6 (39)	530	534	4.3 (−0.5 to 9.1)	545	1.5	36.4 (29.1 to 45.6)	0.79 (0.57 to 1.11)
REWIND (44)	2025	2047	**22.0 (5.1 to 39.0)**	2180	6.0	26.5 (24.5 to 28.6)	**0.88 (0.79 to 0.99)**
SUSTAIN-6 (49)	732	741	**9.7 (1.4 to 18.0)**	767	2.1	43.9 (37.3 to 51.7)	**0.74 (0.57 to 0.95)**
Insulin							
DEVOTE (25)	826	830	4.4 (−2.7 to 11.4)	874	2.4	46.0 (41.8 to 50.7)	0.97 (0.85 to 1.12)
ORIGIN (37)	2420	2413	−7.6 (−30.2 to 15.0)	2677	7.3	28.5 (26.8 to 30.3)	1.03 (0.94 to 1.12)
ORIGINALE (38)	2764	2745	−18.5 (−61.3 to 24.2)	3544	9.7	52.4 (50.1 to 54.8)	1.03 (0.96 to 1.09)
Metformin							
HOME (32)	1117	1138	21.1 (−84.8 to 127.0)	1466	4.0	99.0 (77.4 to 126.8)	0.93 (0.65 to 1.33)
Peroxisome proliferator activated–receptor agonists							
AleCardio (18)	989	991	1.9 (−10.0 to 13.7)	1072	2.9	52.0 (46.9 to 57.7)	0.96 (0.83 to 1.11)
Sodium–glucose cotransporter-2 inhibitors							
CANVAS Program (20)	2233	2265	**32.0 (4.4 to 59.5)**	2472	6.8	32.2 (29.3 to 35.3)	**0.82 (0.72 to 0.92)**
CREDENCE (22)	1164	1195	**31.3 (16.2 to 46.5)**	1276	3.5	61.4 (55.2 to 68.3)	**0.69 (0.59 to 0.82)**
DECLARE-TIMI 58 (24)	1371	1375	4.1 (−3.1 to 11.3)	1437	3.9	24.3 (22.6 to 26.1)	0.94 (0.85 to 1.04)
EMPA-REG OUTCOME (27)	1285	1302	**17.3 (2.2 to 32.3)**	1392	3.8	42.7 (38.0 to 48.0)	**0.85 (0.73 to 0.98)**
Thiazolidinediones							
PROactive (40)	966	973	6.4 (−8.6 to 21.4)	1093	3.0	86.0 (79.2 to 93.4)	0.91 (0.81 to 1.02)
PROFIT-J (42)	829	823	−6.5 (−24.7 to 11.6)	844	2.3	22.2 (12.3 to 40.0)	1.09 (0.46 to 2.56)
RECORD (43)	1968	1963	−4.3 (−32.3 to 23.7)	2135	5.8	28.4 (25.4 to 31.7)	1.00 (0.86 to 1.17)
TOSCA.IT (51)	1772	1778	5.5 (−12.2 to 23.2)	1827	5.0	14.0 (11.4 to 17.2)	0.89 (0.66 to 1.20)
Strategy-driven interventions							
ACCORD (13)	2089	2102	13.3 (−7.9 to 34.5)	2242	6.1	22.7 (20.5 to 25.1)	0.93 (0.80 to 1.07)
ACCORDION (14)	4089	4117	28.6 (−25.6 to 82.8)	4743	13.0	23.5 (22.1 to 25.1)	0.95 (0.87 to 1.04)
ADDITION (15)	2364	2378	14.4 (−16.0 to 44.8)	2487	6.8	15.5 (12.9 to 18.6)	0.87 (0.67 to 1.13)
ADVANCE (16)	1796	1815	**18.8 (0.8 to 36.8)**	2011	5.5	44.4 (41.9 to 47.2)	**0.90 (0.83 to 0.98)**
BARI 2D (19)	1536	1574	38.7 (−3.0 to 80.4)	1794	4.9	56.8 (50.6 to 63.8)	0.89 (0.75 to 1.05)
DCGP FU (23)	3626	3987	**361.7 (113.4 to 610.0)**	6839	18.7	82.8 (75.9 to 90.3)	**0.83 (0.74 to 0.93)**
HEART2D (31)	1001	1000	−0.2 (−64.2 to 63.8)	1376	3.8	138.1 (120.7 to 158.1)	0.97 (0.80 to 1.18)
J-DOIT3 (33)	3219	3258	38.7 (−10.4 to 87.8)	3423	9.4	13.9 (11.7 to 16.4)	0.81 (0.63 to 1.04)
Look AHEAD (35)	3344	3365	21.5 (−24.5 to 67.5)	3692	10.1	21.6 (19.8 to 23.6)	0.94 (0.83 to 1.07)
Steno-2 (46)	2004	2232	227.5 (−7.5 to 462.4)	2553	7.0	77.6 (55.7 to 108.1)	**0.47 (0.27 to 0.82)**
Steno-2 FU a (47)	2913	3607	**693.8 (218.3 to 1169.3)**	4479	12.3	76.3 (57.5 to 101.3)	**0.42 (0.26 to 0.68)**
UKPDS 33 (52)	3970	4101	**131.0 (7.4 to 254.6)**	5369	14.7	45.5 (41.4 to 49.9)	**0.89 (0.79 to 0.99)**
UKPDS 33 FU (54)	5179	5390	211.2 (−1.8 to 424.2)	8917	24.4	51.3 (47.6 to 55.4)	**0.91 (0.83 to 1.00)**
UKPDS 34 (53)	3838	4290	**452.1 (194.1 to 710.1)**	5538	15.2	57.7 (50.5 to 65.9)	**0.66 (0.53 to 0.82)**
UKPDS 34 FU (54)	5111	5634	**522.9 (123.4 to 922.5)**	8687	23.8	54.1 (47.9 to 61.0)	**0.82 (0.68 to 0.98)**
VADT (55)	1969	2029	59.7 (−17.5 to 136.9)	2484	6.8	66.6 (59.1 to 75.2)	0.87 (0.73 to 1.04)
VADT FU a (56)	2965	3097	**132.2 (11.6 to 252.8)**	3839	10.5	52.8 (47.1 to 59.3)	**0.82 (0.70 to 0.97)**
VADT FU b (57)	3773	3939	165.5 (−23.3 to 354.3)	5390	14.8	51.7 (46.5 to 57.5)	0.92 (0.79 to 1.07)

FU = follow-up; HR = hazard ratio; PY = person-years; RCT = randomized controlled trial; RMST = restricted mean survival time; t* = time horizon (minimum of the largest observed event time in the trial groups).
† Differences are between intervention and control: Positive values indicate a postponement of the outcome (as defined in Supplement Table 7, available at Annals.org) comparing intervention with control. Statistically significant ($P < 0.050$) differences are shown in bold.
‡ Within each group, RCTs are sorted in alphabetical order. Only studies with available information (Supplement Table 3, available at Annals.org) are shown. Full trial names are provided in Supplement Table 5 (available at Annals.org).

In UKPDS 34, the primary outcome was postponed by approximately 450 days over 15 years and 520 days over 25 years. Postponements of smaller magnitude also were found for the primary outcomes in UKPDS 33, DCGP FU, and the follow-up of VADT.

Our methods for this example had several limitations. Because direct access to trial data was not possible, individual-level survival times and outcomes were reconstructed by using published Kaplan–Meier curves of varying quality, particularly for older studies. However, agreement was seen between estimated and reported HRs (differences >0.05 in ratios were found in only 4 of 71 cases; Supplement Figure 4, available at Annals.org), and our estimations are in line with those of previous studies using the same methodology (61–64). Moreover, the Kaplan–Meier estimates often become imprecise at the end of the survival curve, because the number of patients at risk may be small. The inability to access individual-level data also limited the possibility of estimating the RMST differences by individual patient characteristics, such as sex and ethnicity, and to account for possible competing risk (65). We used a specific time horizon, which was the minimum of the largest observed event time in both intervention and control groups. However, the RMST also may be estimated continuously, as the increasing difference between the areas under the survival curves with increasing follow-up, up to the end of the study period (multiple t*).

Example 2: Use of the RMST in a Clinical Context

To illustrate the use of the RMST in the setting of a clinical consultation, we consider a 70-year-old man who has had atherosclerotic cardiovascular disease and poorly controlled hyperglycemia while receiving metformin and

Table 2. RMST Difference for All-Cause Mortality†

RCT (Reference)‡	RMST, *d*			Follow-up (t*)		Estimated Event Rate in Control Group per 1000 PY (95% CI)	Estimated HR (95% CI)
	Control	Treatment	Difference (95% CI)	Days	Years		
Specific drug treatments							
Dipeptidyl peptidase-4 inhibitors							
EXAMINE (28)	846	852	6.6 (−2.2 to 15.4)	888	2.4	39.6 (34.0 to 46.1)	0.89 (0.71 to 1.11)
TECOS (50)	1402	1399	−2.6 (−10.6 to 5.4)	1466	4.0	23.8 (21.8 to 25.9)	1.01 (0.89 to 1.14)
Glucagon-like peptide-1–receptor agonists							
EXSCEL (29)	1718	1730	**11.5 (1.1 to 21.9)**	1811	5.0	22.5 (20.7 to 24.5)	**0.88 (0.78 to 1.00)**
LEADER (34)	1538	1549	10.4 (−0.0 to 20.8)	1615	4.4	25.1 (22.9 to 27.5)	**0.85 (0.74 to 0.97)**
REWIND (44)	2049	2066	**17.2 (2.4 to 32.1)**	2175	6.0	22.9 (21.1 to 24.9)	**0.88 (0.79 to 1.00)**
Insulin							
ORIGIN (37)	2474	2474	0.1 (−20.0 to 20.3)	2694	7.4	26.1 (24.5 to 27.8)	0.99 (0.90 to 1.08)
ORIGINALE (38)	2766	2747	−18.8 (−60.6 to 23.0)	3542	9.7	51.9 (49.7 to 54.1)	1.02 (0.96 to 1.08)
Sodium–glucose cotransporter-2 inhibitors							
CANVAS Program (20)	2288	2311	**23.3 (1.6 to 44.9)**	2436	6.7	19.3 (17.1 to 21.7)	**0.85 (0.73 to 0.99)**
CREDENCE (22)	1171	1180	9.5 (−2.5 to 21.4)	1234	3.4	35.2 (30.7 to 40.4)	0.82 (0.67 to 1.00)
DECLARE-TIMI 58 (24)	1402	1404	1.8 (−3.5 to 7.0)	1440	3.9	15.3 (14.0 to 16.7)	0.97 (0.86 to 1.10)
EMPA-REG OUTCOME (27)	1363	1384	**21.4 (9.6 to 33.3)**	1433	3.9	28.5 (24.7 to 32.7)	**0.67 (0.55 to 0.80)**
Thiazolidinediones							
PROactive FU (41)	3850	3875	24.2 (−56.1 to 104.5)	4749	13.0	38.9 (36.3 to 41.6)	0.94 (0.85 to 1.04)
RECORD (43)	2110	2120	9.3 (−6.9 to 25.5)	2174	6.0	11.7 (10.0 to 13.8)	0.88 (0.69 to 1.11)
Strategy-driven interventions							
ACCORD (13)	2189	2173	**−16.1 (−31.4 to −0.7)**	2265	6.2	11.1 (9.6 to 12.7)	**1.24 (1.03 to 1.49)**
ACCORDION (14)	4310	4291	−19.8 (−67.0 to 27.3)	4860	13.3	20.6 (19.4 to 21.9)	1.01 (0.93 to 1.10)
ADDITION (15)	2341	2345	3.7 (−20.6 to 27.9)	2423	6.6	12.1 (9.8 to 14.9)	0.94 (0.71 to 1.25)
ADVANCE (16)	1895	1897	2.4 (−9.4 to 14.3)	1983	5.4	19.6 (18.0 to 21.4)	0.93 (0.82 to 1.06)
ADVANCE-ON (17)	3313	3313	−0.3 (−31.0 to 30.3)	3652	10.0	22.7 (21.3 to 24.1)	1.00 (0.92 to 1.09)
BARI 2D (19)	1712	1712	0.6 (−28.5 to 29.6)	1826	5.0	27.0 (23.0 to 31.7)	0.95 (0.75 to 1.19)
DCGP FU (23)	4293	4358	64.4 (−188.0 to 316.8)	6990	19.1	61.4 (56.0 to 67.4)	0.94 (0.83 to 1.07)
J-DOIT3 (33)	3388	3392	3.5 (−23.6 to 30.5)	3450	9.4	4.7 (3.6 to 6.3)	1.01 (0.68 to 1.51)
Steno-2 FU a (47)	3916	4156	240.2 (−113.1 to 593.6)	4764	13.0	49.7 (36.8 to 67.0)	**0.50 (0.30 to 0.82)**
Steno-2 FU b (48)	4730	5372	**642.6 (69.1 to 1216.1)**	6585	18.0	49.0 (37.6 to 63.8)	**0.56 (0.37 to 0.84)**
UKPDS 33 FU (54)	6756	6959	**203.4 (17.7 to 389.1)**	9112	24.9	29.8 (27.3 to 32.4)	**0.87 (0.79 to 0.97)**
UKPDS 34 FU (54)	6681	7191	**509.9 (149.8 to 870.1)**	9157	25.1	32.6 (28.5 to 37.2)	**0.74 (0.60 to 0.91)**
VADT (55)	2338	2316	−22.3 (−65.2 to 20.6)	2470	6.8	19.5 (15.9 to 23.9)	1.08 (0.81 to 1.44)
VADT FU a (56)	3971	3949	−22.1 (−135.5 to 91.3)	4639	12.7	30.2 (26.7 to 34.1)	1.04 (0.88 to 1.23)
VADT FU b (57)	4691	4663	−27.7 (−186.9 to 131.6)	5882	16.1	35.7 (32.2 to 39.5)	1.01 (0.88 to 1.17)

FU = follow-up; HR = hazard ratio; PY = person-years; RCT = randomized controlled trial; RMST = restricted mean survival time; t* = time horizon (minimum of the largest observed event time in the trial groups).
† Differences are between intervention and control: Positive values indicate a postponement of the outcome comparing intervention with control. Statistically significant ($P < 0.050$) differences are shown in bold.
‡ Within each group, RCTs are sorted in alphabetical order. Only studies with available information (Supplement Table 3, available at Annals.org) are shown. Full trial names are provided in Supplement Table 5 (available at Annals.org).

Table 3. HR and Common Metrics of Absolute Risk Reduction for Time-to-Event Data in Clinical Trials

Effect Measure	Time	Meaning	Example Point Estimate
HR (Cox regression)	Generally the entire follow-up	Ratio of the 2 hazard rates, which represent instantaneous risk at a given time point in previously event-free participants	For an HR of 0.5, the treatment reduces the instantaneous risk for an event by 50% vs. the control.
Difference in survival probabilities/absolute risk reduction	Selected time horizon	Difference in the probabilities of being event-free at the selected time point between the 2 groups	For 2 survival probabilities–0.75 in the treatment group and 0.5 in the control group at 1 y–the probability of being free from the outcome at 1 y is 0.25 (i.e., 25 percentage points) higher in the treatment than the control group.
NNT	Selected time horizon	Number of patients who would need to be transferred to the other group for 1 patient (not) to experience the event by the selected time point. Corresponds to 1/absolute risk reduction	For an NNT of 5 estimated at 1 y and favoring the treatment, 5 patients may need to receive the treatment instead of the control for 1 y to prevent 1 event.
Median survival time difference	Single time point in each group	Difference between the time points at which the survival probability in each group is 50%	For 2 median survival times–1 y in the treatment group and 0.5 y in the control group–the event will have occurred in half the patients, on average, 0.5 y later in the treatment vs. the control group.
RMST	Selected time horizon	Mean delay or anticipation of an event in the treatment vs. the control group over the selected period	For 2 RMSTs estimated over a time horizon of 6 y–5 y and 10 d in the treatment group and 5 y in the control group–a participant delays an event, on average, by 10 d over 6 y through treatment vs. control.

HR = hazard ratio; NNT = number needed to treat; RMST = restricted mean survival time.

for whom current guidelines recommend empagliflozin (66). On the basis of EMPA-REG OUTCOME, the HRs for the composite primary outcome and all-cause mortality are 0.86 and 0.68 (27). The corresponding RMST differences are 17 and 21 days postponement over 4 years. The 2 metrics present different messages that might influence the decision of the health care professional and the patient. Referring to the RMST difference, the health care professional could advise the patient that empagliflozin, on average, would delay a cardiovascular event (death from cardiovascular causes, nonfatal myocardial infarction, or nonfatal stroke) by 17 days and prolong his life by 21 days over 4 years. We believe clinicians and patients would more readily understand these results than if they were told that the relative effect of empagliflozin was a reduction in the hazards of the primary outcome and all-cause mortality by 14% and 32%, respectively.

By quantifying the effects of an intervention in this way, the RMST difference allows the patient and the clinician to decide whether the particular therapy offers a meaningful benefit in the context of the patient's values and preferences. This may be especially salient when the decision to start a treatment is complicated by such issues as a limited time frame for delivering the intervention, important potential side effects, or high costs. In the preceding example of a patient with a life expectancy of only a few years, both the patient and clinician may be better positioned to gauge the expected benefit of starting empagliflozin treatment if the benefits are presented in terms of likely days of event postponement during the next 4 years. The RMST difference might also facilitate "weighing" benefits against potential drawbacks, such as genital infections, when considering whether to withdraw empagliflozin or deintensify treatment (67, 68). We recommend that when deciding whether to start or continue a treatment, both the relative and absolute effects be weighed carefully. To this end, an easily interpretable metric of the absolute effect, such as the RMST difference, should be routinely communicated to the patient.

Advantages of the RMST Difference Over Other Measures

The RMST difference offers several advantages over other commonly used measures, such as the HR, absolute risk reduction, number needed to treat, and median survival time difference (**Table 3**).

Hazard Ratio

The HR is the ratio between the hazard rates of the treatment and control groups. These rates represent the instantaneous risk for participants who have been event-free up to the given time point (**Figure**). Unlike the HR estimated by Cox regression, the RMST difference avoids the assumption of a constant relative benefit (or risk) of the treatment during the study period (proportional hazards; **Supplement Figure 5**, available at Annals.org). It also is less susceptible to uncertainty than the HR when the number of events is small, yielding narrower CIs that make it particularly useful to assess the safety of an intervention or the treatment effect in noninferiority trials (1, 69–71). Similar to other metrics that can be estimated from the survival curves, it also has a better causal interpretation than the HR (1). Substantial evidence exists of incorrect HR interpreta-

Table 3–Continued

Advantages	Disadvantages	Best Suited for
Captures whole survival curve Calculation possible for all major statistical software	Proportional hazards assumption Wide CIs for low event rates Difficult to interpret for patients	Research: analysis of trials with proportional hazards and designs other than noninferiority
Easily obtained from (Kaplan–Meier) survival curves	Insensitive to distribution of events within selected time horizon	Clinical practice and research: explaining the impact of a treatment in terms of avoiding or causing an event
Easily calculable from (Kaplan–Meier) survival curves	Difficult to interpret for patients, especially when CI encompasses benefit and harm Insensitive to distribution of events within selected time horizon	Research: additional measure of statistically significant harm or benefit
Easily obtained from (Kaplan–Meier) survival curves Robust against early or late events	Difficult to interpret for patients Not suitable for low event rates Unstable for small sample sizes	Research: additional measure for large trials with high event rates and/or long duration
Intuitive interpretation for patients Can yield narrow CIs despite low event rate in noninferiority trials	Cannot assess effects beyond selected time point Insensitive to distribution of events within selected time horizon	Clinical practice and research: explaining the impact of a treatment in terms of delaying or anticipating an event

tions resulting in an exaggerated interpretation of the benefit (or risk) of an intervention (2, 3). Hazard ratios may erroneously be interpreted as risk ratios and do not clearly indicate the absolute magnitude of a treatment effect, unless the rate of at least 1 randomized group is reported. For example, large HRs may translate to negligible absolute clinical differences or, conversely, small differences to large absolute benefits when the risk for the outcome is high (2). By contrast, the RMST difference can reflect such differences due to the event rates (Figure). Although the RMST must refer to a stated time horizon, the HR frequently is not reported, or thought of, in relation to the duration of follow-up, thus lending itself to incorrect inferences that the benefit (or risk) extends beyond the study follow-up when the hazard rates might cease to be proportional.

Measures of Absolute Risk Reduction

As for the RMST difference, measures of absolute risk reduction do not assume proportional hazards and may allow for a more accurate interpretation of the benefit of a treatment (3). However, ideally they should be interpreted with reference to both survival curves. Because they refer to specific time horizons, information about the distribution of the individual events included in the analysis—that is, the distribution of events within the selected time horizon—is best captured with several measurements or by plotting the metric itself as a function of time, as for the RMST difference in the Figure.

Among these measures, the difference between the survival probabilities, or absolute risk reduction, for a given time horizon provides a potentially useful alternative to the HR. It effectively captures events avoided or caused by the treatment compared with the control (that is, probability) over a given period. As such, it provides information complementary to the RMST difference, which presents effects in terms of the delay or anticipation of an event (that is, time) over a given time.

A popular, related metric is the number needed to treat. Because it equates to the reciprocal of the absolute risk reduction, it is also readily obtained from the survival curves (72). However, both the number needed to treat and absolute risk reduction may be misinterpreted by health care professionals and patients, because these metrics may indicate very different effects for the same value (for example, an absolute risk reduction of 5 percentage points, equating to a number needed to treat of 20, may result from a 95% to 90% reduction or a 10% to 5% reduction in the absolute risk) (3, 73, 74). This problem is exacerbated if the CI of the number needed to treat encompasses both negative and positive values, representing harm and benefit at either extreme. Moreover, evidence exists that the number needed to treat is frequently reported without reference to key parameters, including the CIs and selected time horizon (75). The number needed to treat therefore may be most useful under specific circumstances, such as to help interpret clinically a statistically significant benefit of a treatment in a trial when key parameters are clearly stated.

The median survival time difference represents the difference between the time points at which the survival probability in each group is 50%. Although reading off the survival curves is easy, the use of this metric is limited to certain situations. It can be obtained only when the survival probabilities actually drop below 50%,

which may not always be the case in trials that are short or have low event rates. Likewise, only with larger sample sizes do CIs become sufficiently narrow, and only with exponential survival distributions does this measure accurately reflect the HR (76).

By contrast, the RMST difference is flexible in its application for designing and analyzing trial data. By presenting treatment effects in terms of a delay over a specific time, it permits a clear and immediate interpretation of results for both health care professionals and patients. It therefore is preferable to other metrics when these are inappropriate from a methodological point of view, and is always valuable in supplementing them to reduce the risk for overstating or understanding treatment effects.

LIMITATIONS TO THE USE OF THE RMST

Although the RMST may be useful in addressing clinical uncertainty and facilitates decision making, its statistical characteristics pose some important challenges that potentially limit its applicability, particularly if it is considered in isolation from other measures of treatment effect. Because the RMST difference reflects the difference in the areas under 2 survival curves, the same difference may be obtained from diverse combinations of survival curves. For example, an RMST difference of 1 year over 5 years may be the result of an RMST of either 4 years in the treatment group and 3 years in the control group or 3 years in the treatment group and 2 years in the control group. As with the aforementioned measures of absolute risk reduction, we therefore advise interpretation of the RMST along with the survival curves.

As shown in the **Figure**, the survival curves of the intervention and control groups, hence their RMST difference, depend on the length of the follow-up and the outcome rates in the control and treatment groups, which are determined by the characteristics of included participants and the effect of the intervention, respectively. This implies that the RMST difference should not be used to directly compare the efficacy or safety of interventions across heterogeneous studies. For example, in the hypothetical consultation described earlier, the health care professional might suggest alternatives to empagliflozin, such as canagliflozin (20, 66). At first glance, the postponements for canagliflozin are larger: 32 days for the primary outcome and 23 days for all-cause mortality, compared with 17 and 21 days, respectively, for empagliflozin. However, the estimates for canagliflozin are restricted to the trial duration—approximately 7 years, compared with 4 years in the case of empagliflozin—and thus refer to different time horizons. To overcome this problem, previous analyses calculated the RMST differences by using an identical time point across studies (61). Even then, a straightforward comparison between the 2 medications would not be possible because the estimated event rates in the control groups differ markedly: around 19 deaths per 1000 person-years in the CANVAS Program versus 29 deaths per 1000 person-years in EMPA-REG OUTCOME. Of note, the same problems apply to other measures, such as the number needed to treat, because these are equally dependent on the time horizon and the event rates. Other analytic strategies, such as pairwise or network meta-analysis of HRs, may be more suitable for summarizing and comparing interventions, given the evidence of constant HR across heterogeneous characteristics of RCT participants (77–79).

In view of the relationship between follow-up and magnitude of postponement, an argument may be made that the smaller effects observed in many drug-specific trials are related mainly to the shorter follow-up of those studies. Because no data are available on the long-term effect for specific drugs, our results cannot indicate that these interventions do not deliver a clinically meaningful impact on the outcomes beyond the duration of the trial; conversely, long-term data are available for glucose-lowering strategies, such as in the Steno-2 study and UKPDS 33 and 34. By estimating the treatment effect beyond the trial period, several studies have shown a way to improve the usefulness of the RMST for short trials. Using individual-level data, Claggett and colleagues (80) quantified the years of life gained for persons of different ages receiving empagliflozin in EMPA-REG OUTCOME. The life expectancy was 4.5 years longer at the age of 45 years and decreased progressively to 1 year longer at the age of 80 years, assuming a lifelong exposure to the active treatment until the age of 90 years (that is, 45 and 10 years of treatment, respectively) (80). Through a different analysis using aggregated data, years of life gained were estimated by applying the HR of death reported in statin trials to the mortality rates obtained from national registries. The gain in life ranged from 1.3 years at the age of 30 years to 0.7 year at the age of 70 years, assuming a lifelong exposure to statin treatment until the age of 100 years (81). These differences are only apparently larger than our estimates. In fact, in both analyses, the follow-up was extended to ages much older than those observed in the trials, up to 90 or 100 years in persons whose mean age at the baseline assessment across these trials was about 60 years and who were followed for a median of 5 years. Moreover, both studies assume that patients remain on the same treatment for their entire life and that treatment effects are constant over time. In addition, these methods assume that the HR obtained from a highly selected group (that is, persons participating in a trial) may be applied to an external population that may be very different from trial participants. Although these long-term effects are possible, presence and magnitude of a treatment effect are proven for the study population and for the follow-up of the trial. Predictions for subjects with different characteristics or for a longer follow-up should be considered carefully along with analytic assumptions and a trial's generalizability (82); therefore, health care professionals should make patients aware of these caveats when referring to RMST differences that are predicted beyond the available trial data.

SOFTWARE

Because commands for estimating the RMST are emerging frequently, users should check for new and updated packages and indicate the chosen commands with, where available, their version. At the time of writing, the estimation of the HR and the assessment of the proportionality of hazards were implemented in all major statistical software, whereas commands to calculate the RMST difference have been developed mainly in Stata and R (R Foundation for Statistical Computing). Stata commands now include *standsurv* (83), *strmst2* (12), *strmst* (84), and *stpmean* (85); all allow for adjustment for covariates, and–along with the difference–*standsurv* and *strmst2* also report the RMST ratio. In R, available packages are *survRM2adapt* (86) and *survRM2* (87), the latter of which also allows adjustment for covariates and estimation of the RMST ratio. The number of observations in a trial commonly is smaller than that of an observational study, and less frequently, adjustment for covariates is required. Yet, the estimation of the RMST difference may be computationally intensive for large databases or complex models.

SUMMARY

Here, we discuss the RMST difference as a convenient effect measure in helping to answer research questions using clinical trial data and in communicating treatment effects during patient consultations. Applied to studies examining the effects of glucose-lowering treatments or strategies on macrovascular outcomes or death postponement in patients with type 2 diabetes (example 1), the RMST difference helps one appreciate clinically important benefits of intensive glucose control. In the setting of a patient consultation (example 2), it may be part of the information to support a health care professional and patient in deciding whether to start or discontinue a treatment.

We also highlight advantages of the RMST difference over the HR and several measures of absolute risk difference. Unlike those metrics, the RMST difference is applicable to any time-to-event data and is less liable to errors in interpretation. However, we also draw attention to several limitations to the use of the RMST. Of most importance, the RMST difference should not be used to compare heterogeneous trials, and any estimations beyond the duration of follow-up of a study should be interpreted with caution. We recommend routine use of the RMST difference in future RCTs with time-to-event outcomes to complement other measures of treatment effect.

From Leicester Real World Evidence Unit, Leicester Diabetes Centre, Leicester General Hospital, Leicester, United Kingdom (D.E.K., K.K., F.Z.); and Diabetes Research Centre, Leicester Diabetes Centre, Leicester General Hospital, Leicester, United Kingdom (D.E.K., M.J.D., K.K., F.Z.).

Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, National Institute for Health Research (NIHR), or Department of Health.

Financial Support: From NIHR Applied Research Collaborations East Midlands and NIHR Leicester Biomedical Research Centre. The funding bodies had no role in the design, conduct, or analysis of the study or in the decision to submit the manuscript for publication.

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M19-3286.

Corresponding Author: David E. Kloecker, MPhil, Leicester Real World Evidence Unit, Leicester Diabetes Centre, Leicester General Hospital, Gwendolen Road, Leicester LE5 4PW, United Kingdom; e-mail, dek9@student.le.ac.uk or davidekloecker@gmail.com.

Current author addresses and author contributions are available at Annals.org.

References

1. **Stensrud MJ, Aalen JM, Aalen OO, et al.** Limitations of hazard ratios in clinical trials. Eur Heart J. 2019;40:1378-1383. [PMID: 30500891] doi:10.1093/eurheartj/ehy770
2. **Spiegelhalter D.** Risk and uncertainty communication. Annu Rev Stat Appl. 2017;4:31–60.
3. **Zipkin DA, Umscheid CA, Keating NL, et al.** Evidence-based risk communication: a systematic review. Ann Intern Med. 2014;161:270-80. [PMID: 25133362] doi:10.7326/M14-0295
4. **Keene ON.** Alternatives to the hazard ratio in summarizing efficacy in time-to-event studies: an example from influenza trials. Stat Med. 2002;21:3687-700. [PMID: 12436464]
5. **Kristensen ML, Christensen PM, Hallas J.** The effect of statins on average survival in randomised trials, an analysis of end point postponement. BMJ Open. 2015;5:e007118. [PMID: 26408281] doi:10.1136/bmjopen-2014-007118
6. **Kim DH, Uno H, Wei LJ.** Restricted mean survival time as a measure to interpret clinical trial results. JAMA Cardiol. 2017;2:1179-1180. [PMID: 28877311] doi:10.1001/jamacardio.2017.2922
7. **Uno H, Claggett B, Tian L, et al.** Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol. 2014;32:2380-5. [PMID: 24982461] doi:10.1200/JCO.2014.55.2208
8. **Trinquart L, Jacot J, Conner SC, et al.** 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;34:1813-9. [PMID: 26884584] doi:10.1200/JCO.2015.64.2488
9. **A'Hern RP.** Restricted mean survival time: an obligatory end point for time-to-event analysis in cancer trials? [Editorial]. J Clin Oncol. 2016;34:3474-6. [PMID: 27507871] doi:10.1200/JCO.2016.67.8045
10. **Higgins JP, Altman DG, Gøtzsche PC, et al; Cochrane Bias Methods Group.** The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928. [PMID: 22008217] doi:10.1136/bmj.d5928
11. **Wei Y, Royston P.** Reconstructing time-to-event data from published Kaplan-Meier curves. Stata J. 2017;17:786-802. [PMID: 29398980]
12. **Cronin A, Tian L, Uno H.** Strmst2 and Strmst2pw: new commands to compare survival curves using the restricted mean survival time. Stata J. 2016;16:702-16.
13. **Gerstein HC, Miller ME, Byington RP, et al; Action to Control Cardiovascular Risk in Diabetes Study Group.** Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med. 2008;358:2545-59. [PMID: 18539917] doi:10.1056/NEJMoa0802743

14. **ACCORD Study Group.** Nine-year effects of 3.7 years of intensive glycemic control on cardiovascular outcomes. Diabetes Care. 2016; 39:701-8. [PMID: 26822326] doi:10.2337/dc15-2283
15. **Griffin SJ, Borch-Johnsen K, Davies MJ, et al.** Effect of early intensive multifactorial therapy on 5-year cardiovascular outcomes in individuals with type 2 diabetes detected by screening (ADDITION-Europe): a cluster-randomised trial. Lancet. 2011;378:156-67. [PMID: 21705063] doi:10.1016/S0140-6736(11)60698-3
16. **Patel A, MacMahon S, Chalmers J, et al; ADVANCE Collaborative Group.** Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. N Engl J Med. 2008;358:2560-72. [PMID: 18539916] doi:10.1056/NEJMoa0802987
17. **Zoungas S, Chalmers J, Neal B, et al; ADVANCE-ON Collaborative Group.** Follow-up of blood-pressure lowering and glucose control in type 2 diabetes. N Engl J Med. 2014;371:1392-406. [PMID: 25234206] doi:10.1056/NEJMoa1407963
18. **Lincoff AM, Tardif JC, Schwartz GG, et al; AleCardio Investigators.** Effect of aleglitazar on cardiovascular outcomes after acute coronary syndrome in patients with type 2 diabetes mellitus: the AleCardio randomized clinical trial. JAMA. 2014;311:1515-25. [PMID: 24682069] doi:10.1001/jama.2014.3321
19. **Frye RL, August P, Brooks MM, et al; BARI 2D Study Group.** A randomized trial of therapies for type 2 diabetes and coronary artery disease. N Engl J Med. 2009;360:2503-15. [PMID: 19502645] doi:10.1056/NEJMoa0805796
20. **Neal B, Perkovic V, Mahaffey KW, et al; CANVAS Program Collaborative Group.** Canagliflozin and cardiovascular and renal events in type 2 diabetes. N Engl J Med. 2017;377:644-657. [PMID: 28605608] doi:10.1056/NEJMoa1611925
21. **Rosenstock J, Perkovic V, Johansen OE, et al; CARMELINA Investigators.** Effect of linagliptin vs placebo on major cardiovascular events in adults with type 2 diabetes and high cardiovascular and renal risk: the CARMELINA randomized clinical trial. JAMA. 2019; 321:69-79. [PMID: 30418475] doi:10.1001/jama.2018.18269
22. **Perkovic V, Jardine MJ, Neal B, et al; CREDENCE Trial Investigators.** Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. N Engl J Med. 2019;380:2295-2306. [PMID: 30990260] doi:10.1056/NEJMoa1811744
23. **Hansen LJ, Siersma V, Beck-Nielsen H, et al.** Structured personal care of type 2 diabetes: a 19 year follow-up of the study Diabetes Care in General Practice (DCGP). Diabetologia. 2013;56:1243-53. [PMID: 23549519] doi:10.1007/s00125-013-2893-1
24. **Wiviott SD, Raz I, Bonaca MP, et al; DECLARE–TIMI 58 Investigators.** Dapagliflozin and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2019;380:347-357. [PMID: 30415602] doi:10.1056/NEJMoa1812389
25. **Marso SP, McGuire DK, Zinman B, et al; DEVOTE Study Group.** Efficacy and safety of degludec versus glargine in type 2 diabetes. N Engl J Med. 2017;377:723-732. [PMID: 28605603] doi:10.1056/NEJMoa1615692
26. **Pfeffer MA, Claggett B, Diaz R, et al; ELIXA Investigators.** Lixisenatide in patients with type 2 diabetes and acute coronary syndrome. N Engl J Med. 2015;373:2247-57. [PMID: 26630143] doi:10.1056/NEJMoa1509225
27. **Zinman B, Wanner C, Lachin JM, et al; EMPA-REG OUTCOME Investigators.** Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med. 2015;373:2117-28. [PMID: 26378978] doi:10.1056/NEJMoa1504720
28. **White WB, Cannon CP, Heller SR, et al; EXAMINE Investigators.** Alogliptin after acute coronary syndrome in patients with type 2 diabetes. N Engl J Med. 2013;369:1327-35. [PMID: 23992602] doi:10.1056/NEJMoa1305889
29. **Holman RR, Bethel MA, Mentz RJ, et al; EXSCEL Study Group.** Effects of once-weekly exenatide on cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2017;377:1228-1239. [PMID: 28910237] doi:10.1056/NEJMoa1612917
30. **Hernandez AF, Green JB, Janmohamed S, et al; Harmony Outcomes committees and investigators.** Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. Lancet. 2018;392:1519-1529. [PMID: 30291013] doi:10.1016/S0140-6736(18)32261-X
31. **Raz I, Wilson PW, Strojek K, et al.** Effects of prandial versus fasting glycemia on cardiovascular outcomes in type 2 diabetes: the HEART2D trial. Diabetes Care. 2009;32:381-6. [PMID: 19246588] doi:10.2337/dc08-1671
32. **Kooy A, de Jager J, Lehert P, et al.** Long-term effects of metformin on metabolism and microvascular and macrovascular disease in patients with type 2 diabetes mellitus. Arch Intern Med. 2009;169: 616-25. [PMID: 19307526] doi:10.1001/archinternmed.2009.20
33. **Ueki K, Sasako T, Okazaki Y, et al; J-DOIT3 Study Group.** Effect of an intensified multifactorial intervention on cardiovascular outcomes and mortality in type 2 diabetes (J-DOIT3): an open-label, randomised controlled trial. Lancet Diabetes Endocrinol. 2017;5: 951-964. [PMID: 29079252] doi:10.1016/S2213-8587(17)30327-3
34. **Marso SP, Daniels GH, Brown-Frandsen K, et al; LEADER Steering Committee.** Liraglutide and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2016;375:311-22. [PMID: 27295427] doi:10.1056/NEJMoa1603827
35. **Wing RR, Bolin P, Brancati FL, et al; Look AHEAD Research Group.** Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. N Engl J Med. 2013;369:145-54. [PMID: 23796131] doi:10.1056/NEJMoa1212914
36. **Gantz I, Chen M, Suryawanshi S, et al.** A randomized, placebo-controlled study of the cardiovascular safety of the once-weekly DPP-4 inhibitor omarigliptin in patients with type 2 diabetes mellitus. Cardiovasc Diabetol. 2017;16:112. [PMID: 28893244] doi:10.1186/s12933-017-0593-8
37. **Gerstein HC, Bosch J, Dagenais GR, et al; ORIGIN Trial Investigators.** Basal insulin and cardiovascular and other outcomes in dysglycemia. N Engl J Med. 2012;367:319-28. [PMID: 22686416] doi:10.1056/NEJMoa1203858
38. **ORIGIN Trial Investigators.** Cardiovascular and other outcomes postintervention with insulin glargine and omega-3 fatty acids (ORIGINALE). Diabetes Care. 2016;39:709-16. [PMID: 26681720] doi:10.2337/dc15-1676
39. **Husain M, Birkenfeld AL, Donsmark M, et al; PIONEER 6 Investigators.** Oral semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med. 2019;381:841-851. [PMID: 31185157] doi:10.1056/NEJMoa1901118
40. **Dormandy JA, Charbonnel B, Eckland DJ, et al; PROactive Investigators.** Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. Lancet. 2005;366:1279-89. [PMID: 16214598]
41. **Erdmann E, Harding S, Lam H, et al.** Ten-year observational follow-up of PROactive: a randomized cardiovascular outcomes trial evaluating pioglitazone in type 2 diabetes. Diabetes Obes Metab. 2016;18:266-73. [PMID: 26592506] doi:10.1111/dom.12608
42. **Yoshii H, Onuma T, Yamazaki T, et al; PROFIT-J Study Group.** Effects of pioglitazone on macrovascular events in patients with type 2 diabetes mellitus at high risk of stroke: the PROFIT-J study. J Atheroscler Thromb. 2014;21:563-73. [PMID: 24477028]
43. **Home PD, Pocock SJ, Beck-Nielsen H, et al; RECORD Study Team.** Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (RECORD): a multicentre, randomised, open-label trial. Lancet. 2009;373:2125-35. [PMID: 19501900] doi:10.1016/S0140-6736(09)60953-3
44. **Gerstein HC, Colhoun HM, Dagenais GR, et al; REWIND Investigators.** Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. Lancet. 2019;394:121-130. [PMID: 31189511] doi:10.1016/S0140-6736(19)31149-3
45. **Scirica BM, Bhatt DL, Braunwald E, et al; SAVOR-TIMI 53 Steering Committee and Investigators.** Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. N Engl J Med. 2013;369:1317-26. [PMID: 23992601] doi:10.1056/NEJMoa1307684

46. **Gaede P, Vedel P, Larsen N, et al.** Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. N Engl J Med. 2003;348:383-93. [PMID: 12556541]
47. **Gaede P, Lund-Andersen H, Parving HH, et al.** Effect of a multifactorial intervention on mortality in type 2 diabetes. N Engl J Med. 2008;358:580-91. [PMID: 18256393] doi:10.1056/NEJMoa0706245
48. **Gæde P, Oellgaard J, Carstensen B, et al.** Years of life gained by multifactorial intervention in patients with type 2 diabetes mellitus and microalbuminuria: 21 years follow-up on the Steno-2 randomised trial. Diabetologia. 2016;59:2298-2307. [PMID: 27531506] doi:10.1007/s00125-016-4065-6
49. **Marso SP, Bain SC, Consoli A, et al; SUSTAIN-6 Investigators.** Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med. 2016;375:1834-1844. [PMID: 27633186]
50. **Green JB, Bethel MA, Armstrong PW, et al; TECOS Study Group.** Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2015;373:232-42. [PMID: 26052984] doi:10.1056/NEJMoa1501352
51. **Vaccaro O, Masulli M, Nicolucci A, et al; Thiazolidinediones Or Sulfonylureas Cardiovascular Accidents Intervention Trial (TOSCA.IT) study group.** Effects on the incidence of cardiovascular events of the addition of pioglitazone versus sulfonylureas in patients with type 2 diabetes inadequately controlled with metformin (TOSCA.IT): a randomised, multicentre trial. Lancet Diabetes Endocrinol. 2017;5:887-897. [PMID: 28917544] doi:10.1016/S2213-8587(17)30317-0
52. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet. 1998;352:837-53. [PMID: 9742976]
53. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). UK Prospective Diabetes Study (UKPDS) Group. Lancet. 1998;352:854-65. [PMID: 9742977]
54. **Holman RR, Paul SK, Bethel MA, et al.** 10-year follow-up of intensive glucose control in type 2 diabetes. N Engl J Med. 2008;359:1577-89. [PMID: 18784090] doi:10.1056/NEJMoa0806470
55. **Duckworth W, Abraira C, Moritz T, et al; VADT Investigators.** Glucose control and vascular complications in veterans with type 2 diabetes. N Engl J Med. 2009;360:129-39. [PMID: 19092145] doi:10.1056/NEJMoa0808431
56. **Hayward RA, Reaven PD, Wiitala WL, et al; VADT Investigators.** Follow-up of glycemic control and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2015;372:2197-206. [PMID: 26039600] doi:10.1056/NEJMoa1414266
57. **Reaven PD, Emanuele NV, Wiitala WL, et al; VADT Investigators.** Intensive glucose control in patients with type 2 diabetes - 15-year follow-up. N Engl J Med. 2019;380:2215-2224. [PMID: 31167051] doi:10.1056/NEJMoa1806802
58. **Olivarius NF, Beck-Nielsen H, Andreasen AH, et al.** Randomised controlled trial of structured personal care of type 2 diabetes mellitus. BMJ. 2001;323:970-5. [PMID: 11679387]
59. **Sarwar N, Gao P, Seshasai SR, et al; Emerging Risk Factors Collaboration.** Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet. 2010;375:2215-22. [PMID: 20609967] doi:10.1016/S0140-6736(10)60484-9
60. **Cefalu WT, Kaul S, Gerstein HC, et al.** Cardiovascular outcomes trials in type 2 diabetes: where do we go from here? Reflections from a diabetes care editors' expert forum. Diabetes Care. 2018;41:14-31. [PMID: 29263194] doi:10.2337/dci17-0057
61. **Kaneko M, Narukawa M.** Time-matched evaluation of cardiovascular risks associated with drugs for type 2 diabetes mellitus. Clin Drug Investig. 2019;39:469-476. [PMID: 30887418] doi:10.1007/s40261-019-00770-z
62. **Kaneko M, Narukawa M.** Effects of sodium-glucose cotransporter 2 inhibitors on amputation, bone fracture, and cardiovascular outcomes in patients with type 2 diabetes mellitus using an alternative measure to the hazard ratio. Clin Drug Investig. 2019;39:179-186. [PMID: 30506378] doi:10.1007/s40261-018-0731-4
63. **Kaneko M, Narukawa M.** Assessment of cardiovascular risk with glucagon-like peptide 1 receptor agonists in patients with type 2 diabetes using an alternative measure to the hazard ratio. Ann Pharmacother. 2018;52:632-638. [PMID: 29424239] doi:10.1177/1060028018757407
64. **Kaneko M, Narukawa M.** Assessment of the risk of hospitalization for heart failure with dipeptidyl peptidase-4 inhibitors, saxagliptin, alogliptin, and sitagliptin in patients with type 2 diabetes, using an alternative measure to the hazard ratio. Ann Pharmacother. 2017;51:570-576. [PMID: 28622738] doi:10.1177/1060028017698496
65. **Austin PC, Fine JP.** Accounting for competing risks in randomized controlled trials: a review and recommendations for improvement. Stat Med. 2017;36:1203-1209. [PMID: 28102550] doi:10.1002/sim.7215
66. **Davies MJ, D'Alessio DA, Fradkin J, et al.** Management of hyperglycaemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia. 2018;61:2461-2498. [PMID: 30288571] doi:10.1007/s00125-018-4729-5
67. **Sussman JB, Kerr EA, Saini SD, et al.** Rates of deintensification of blood pressure and glycemic medication treatment based on levels of control and life expectancy in older patients with diabetes mellitus. JAMA Intern Med. 2015;175:1942-9. [PMID: 26502220] doi:10.1001/jamainternmed.2015.5110
68. **Seidu S, Kunutsor SK, Topsever P, et al.** Deintensification in older patients with type 2 diabetes: a systematic review of approaches, rates and outcomes. Diabetes Obes Metab. 2019;21:1668-1679. [PMID: 30938038] doi:10.1111/dom.13724
69. **Hernán MA.** The hazards of hazard ratios. Epidemiology. 2010;21:13-5. [PMID: 20010207] doi:10.1097/EDE.0b013e3181c1ea43
70. **Dehbi HM, Royston P, Hackshaw A.** Life expectancy difference and life expectancy ratio: two measures of treatment effects in randomised trials with non-proportional hazards. BMJ. 2017;357:j2250. [PMID: 28546261] doi:10.1136/bmj.j2250
71. **Uno H, Wittes J, Fu H, et al.** Alternatives to hazard ratios for comparing the efficacy or safety of therapies in noninferiority studies. Ann Intern Med. 2015;163:127-34. [PMID: 26054047] doi:10.7326/M14-1741
72. **Altman DG, Andersen PK.** Calculating the number needed to treat for trials where the outcome is time to an event. BMJ. 1999;319:1492-5. [PMID: 10582940]
73. **Ghosh AK, Ghosh K.** Misinterpretation of numbers: a potential source of physician's error. Lancet. 2004 Sep 25-Oct 1;364:1108-9. [PMID: 15451207]
74. **Harrell F.** Problems with NNT. Accessed at https://discourse.datamethods.org/t/problems-with-nnt/195 on 24 May 2019.
75. **Mendes D, Alves C, Batel-Marques F.** Number needed to treat (NNT) in clinical literature: an appraisal. BMC Med. 2017;15:112. [PMID: 28571585] doi:10.1186/s12916-017-0875-8
76. **Saad ED, Zalcberg JR, Péron J, et al.** Understanding and communicating measures of treatment effect on survival: can we do better? J Natl Cancer Inst. 2018;110:232-240. [PMID: 29933439] doi:10.1093/jnci/djx179
77. **Kent DM, Paulus JK, van Klaveren D, et al.** The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. Ann Intern Med. 2019. [PMID: 31711134] doi:10.7326/M18-3667
78. **Kent DM, van Klaveren D, Paulus JK, et al.** The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement: explanation and elaboration. Ann Intern Med. 2019. [PMID: 31711094] doi:10.7326/M18-3668
79. **Kent DM, Nelson J, Dahabreh IJ, et al.** Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. Int J Epidemiol. 2016;45:2075-2088. [PMID: 27375287] doi:10.1093/ije/dyw118
80. **Claggett B, Lachin JM, Hantel S, et al.** Long-term benefit of empagliflozin on life expectancy in patients with type 2 diabetes mellitus and established cardiovascular disease [Letter]. Circulation. 2018;

138:1599-1601. [PMID: 30354516] doi:10.1161/CIRCULATIONAHA.118.033810
81. **Carstensen B.** Do statins only postpone death by three or four days? Accessed at https://bmjopen.bmj.com/content/5/9/e007118.responses#do-statins-only-postpone-death-by-three-or-four-days qj on 24 May 2019.
82. **Yeh RW, Valsdottir LR, Yeh MW, et al; PARACHUTE Investigators.** Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. BMJ. 2018;363:k5094. [PMID: 30545967] doi:10.1136/bmj.k5094
83. **Lambert PC.** standsurv. Accessed at https://pclambert.net/software/standsurv on 15 January 2020.
84. **Royston P.** Estimating the treatment effect in a clinical trial using difference in restricted mean survival time. Stata J. 2015;15:1098-117.
85. **Overgaard M, Andersen PK, Parner ET.** Regression analysis of censored data using pseudo-observations: an update. Stata J. 2015; 15:809-21.
86. **Horiguchi M, Uno H.** Package 'survRM2adapt.' Accessed at https://cran.r-project.org/web/packages/survRM2adapt/survRM2adapt.pdf on 15 January 2020.
87. **Uno H, Tian L, Cronin A, et al.** Package 'survRM2.' Accessed at https://cran.r-project.org/web/packages/survRM2/survRM2.pdf on 15 January 2020.

Current Author Addresses: Mr. Kloecker: Leicester Real World Evidence Unit, Leicester Diabetes Centre, Leicester General Hospital, Gwendolen Road, Leicester LE5 4PW, United Kingdom.
Drs. Davies, Khunti, and Zaccardi: Leicester Diabetes Centre, Leicester General Hospital, Gwendolen Road, Leicester LE5 4PW, Leicestershire, United Kingdom.

Author Contributions: Conception and design: F. Zaccardi.
Analysis and interpretation of the data: D.E. Kloecker, F. Zaccardi.
Drafting of the article: D.E. Kloecker, F. Zaccardi.
Critical revision for important intellectual content: D.E. Kloecker, M.J. Davies, K. Khunti, F. Zaccardi.
Final approval of the article: D.E. Kloecker, M.J. Davies, K. Khunti, F. Zaccardi.
Provision of study materials or patients: D.E. Kloecker.
Statistical expertise: D.E. Kloecker, F. Zaccardi.
Administrative, technical, or logistic support: D.E. Kloecker.
Collection and assembly of data: D.E. Kloecker.