HAL open science

Computing and interpreting the Number Needed to Treat for Cardiovascular Outcomes Trials

Lisa Ludwig, Patrice Darmon, Bruno Guerci

To cite this version:

Lisa Ludwig, Patrice Darmon, Bruno Guerci. Computing and interpreting the Number Needed to Treat for Cardiovascular Outcomes Trials. Cardiovascular Diabetology, 2020, 19 (1), 10.1186/s12933-020-01034-3 . hal-03158415

HAL Id: hal-03158415

https://hal.inrae.fr/hal-03158415v1

Submitted on 19 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Ludwig *et al. Cardiovasc Diabetol* (2020) 19:65
https://doi.org/10.1186/s12933-020-01034-3

Cardiovascular Diabetology

COMMENTARY

Open Access

Computing and interpreting the Number Needed to Treat for Cardiovascular Outcomes Trials

Perspective on GLP-1 RA and SGLT-2i therapies

Lisa Ludwig[1*], Patrice Darmon[2,3] and Bruno Guerci[1]

Abstract

The recent results of Cardiovascular Outcomes Trials (CVOTs) in type 2 diabetes have clearly established the cardiovascular (CV) safety or even the benefit of two therapeutic classes, Glucagon-Like Peptide-1 receptor agonists (GLP-1 RA) and Sodium-Glucose Co-Transporter-2 inhibitors (SGLT-2i). Publication of the latest CVOTs for these therapeutic classes also led to an update of ESC guidelines and ADA/EASD consensus report in 2019, which considers using GLP-1 RA or SGLT-2i with proven cardiovascular benefit early in the management of type 2 diabetic patient with established cardiovascular disease (CVD) or at high risk of atherosclerotic CVD. The main beneficial results of these time-to event studies are supported by conventional statistical measures attesting the effectiveness of GLP-1 RA or SGLT2i on cardiovascular events (absolute risk, absolute risk difference, relative risk, relative risk reduction, odds ratio, hazard ratio). In addition, another measure whose clinical meaning appears to be easier, the Number Needed to Treat (NNT), is often mentioned while discussing the results of CVOTs, in order to estimating the clinical utility of each drug or sometimes trying to establish a power ranking. While the value of the measure is admittedly of interest, the subtleties of its computation in time-to-event studies are little known. We provide in this article a clear and practical explanation on NNT computation methods that should be used in order to estimate its value, according to the type of study design and variables available to describe the event of interest, in any randomized controlled trial. More specifically, a focus is made on time-to-event studies of which CVOTs are part, first to describe in detail an appropriate and adjusted method of NNT computation and second to help properly interpreting NNTs with the example of CVOTs conducted with GLP-1 RA and SGLT-2i. We particularly discuss the risk of misunderstanding of NNT values in CVOTs when some specific parameters inherent in each study are not taken into account, and the following risk of erroneous comparison between NNTs across studies. The present paper highlights the importance of understanding rightfully NNTs from CVOTs and their clinical impact to get the full picture of a drug's effectiveness.

Keywords: Cardiovascular Outcome Trial, Type 2 diabetes, Number Needed to Treat, GLP-1 receptor agonist, SGLT-2 inhibitors

Background

During the last decade, Cardiovascular Outcomes Trials (CVOTs) have aroused considerable interest among diabetologists and cardiologists. In these randomized, controlled, time-to-event studies evaluating the cardiovascular (CV) safety of emerging antidiabetic drugs, the primary endpoint is most often a composite CV criterion

*Correspondence: L.LUDWIG@chru-nancy.fr
[1] Endocrinology, Diabetology & Nutrition, CHRU of Nancy, Brabois Hospital, Lorraine University, 54500 Vandoeuvre-lès-Nancy, France
Full list of author information is available at the end of the article

© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

called "3P-MACE" combining mortality from CV cause, non-fatal myocardial infarction and non-fatal stroke. CVOTs have notably clearly established the CV safety or even the benefit of two therapeutic classes, Glucagon-Like Peptide-1 receptor agonists (GLP-1 RA) and sodium-glucose co-transporter-2 inhibitors (SGLT-2i), and have positioned them as priority options in the management of type 2 diabetes, especially in patients with an established CV disease, heart failure and/or chronic kidney disease. The results of these CVOTs, confirmed by two recent meta-analyzes [1, 2] have led to the drafting of a new ADA/EASD consensus in 2018 [3], then revised at the end of 2019 [4], and to recommendations from the European Cardiology Society in 2019 [5]. The question then arises of the relative effectiveness of these molecules or classes. Indeed, the significant results obtained in the CVOTs question whether a ranking of molecules or classes could be established. In addition to the conventional statistical measures attesting the effectiveness of GLP-1 RA or SGLT2i in CVOTs (absolute risk and absolute risk difference, relative risk and relative risk reduction, odds ratio and hazard ratio), one may also refer to another measure whose clinical meaning appears to be easier: the Number Needed to Treat (NNT).

The NNT is a now common, statistical measure of the clinical utility of a treatment. After each reporting of CVOT results, a lively discussion ensued on the associated NNT and its impact, to a greater or lesser extent, on clinical practices or even on the efficiency of the drug itself. While the value of the measure is admittedly of interest, its computation remains controversial. Indeed, the seemingly simple calculation of the NNT must, however, consider some subtleties when derived from interventional studies where the occurrence of the primary endpoint is a function of time. Several methods to adjust the computation according to the design and assumptions of each study have been proposed by statisticians, some more consensual than others. Nevertheless, numerous calculation errors still occur, especially while reporting NNTs in these time-to-event studies, and of which CVOTs are an example. Indeed, a meta-analysis carried out on articles published in 4 major international journals (BMJ, JAMA, N Engl J Med, Lancet) between 2003 and 2005, and reporting results of randomized controlled trial (RCT) with a time-dependent main outcome highlighted that at least half of them reported incorrect NNT values because of an inappropriate computation method [6]. A second survey carried out in 2009 on the same journals reported 60% computation errors for studies where the occurrence of the event of interest is time-dependent [7]. The same issue, regarding complexity of NNTs' computation and interpretation, was found within more recent publication papers discussing data from other cardioprotective drug classes such as antiplatelet therapy, Proprotein Convertase subtilsin-kexin type 9 (PCSK-9) inhibitors, or angiotensin receptor–neprilysin inhibitor (ARNI), among others [8–11]. Hence, a wide range of therapeutic areas are concerned by the issue, which emphasizes the importance of paying careful attention to the analysis of NNTs.

This article aims firstly at clarifying in a practical manner how to calculate an NNT for time-to-event RCTs while avoiding some of the classical mistakes, and secondly at helping to interpret appropriately NNTs with the example of CVOTs results. One should not expect the development or demonstration of complex statistical models, but rather a key to simple computation and interpretation of NNTs in CVOTs. Also, data from observational studies, cohorts, case–control studies or even meta-analyzes are subject to specific NNT calculation methods and will not be addressed hereafter.

NNT computation methods

Introduced in 1988 by Laupacis et al., the NNT represents the number of patients to be treated during a given period of time to prevent the occurrence of one additional negative or unfavorable event [12]. A NNT can be calculated if the outcome of interest is binary, but unfortunately not if the main outcome associated data are continuous. The choice of the NNT computation method will depend in particular on the design of the study as well as on the type of variables describing the event of interest.

1. Computation of the absolute risk difference

First of all, before considering the NNT calculation, the absolute risk difference needs to be computed (Fig. 1) [13]. There are two possibilities:

First, consider the probability p of occurrence of a negative event in each group (also called "risk"), for example severe hypoglycemia. If the treatment is effective, the risk of a severe hypoglycemia should be lower in the experimental group as compared to the control group. The absolute risk difference (ARD) will be as follows: $ARD = p_{ctrl} - p_{exp}$. This is the most common case.

Then, consider a positive outcome of interest, e.g. proportion of patients free of retinopathy after 3 years of treatment. Its probability of occurrence p will rather be considered as a chance. In this case, if the treatment is effective, a greater proportion of patients should remain free of retinopathy, and thus the occurrence probability of the event should be higher in the treatment group as compared to the control group. The absolute risk difference formula should be inverted to obtain in this case: $ARD = p_{exp} - p_{ctrl}$.

	Negative Outcome	Positive Outcome	Total
Experimental group	a	c	a + c
Control group	b	d	b + d

$$p_{ctrl} = \frac{b}{b+d}$$

Probability of occurrence of the negative outcome (or « risk ») in the control group

$$p_{exp} = \frac{a}{a+c}$$

Probability of occurrence of the negative outcome (or « risk ») in the experimental group

Hence, the absolute risk difference is :

$$ARD = p_{ctrl} - p_{exp} = \frac{b}{b+d} - \frac{a}{a+c}$$

Fig. 1 Calculation of the absolute risk difference in practice. ARD: absolute risk difference

2. Binary outcome study with fixed and constant follow-up period

In RCT where the observed variables are binary ("event" vs "no event") and all the patients are followed for a predefined period of time, the risk that is the proportion of patients who present with the unfavorable event, is measured in each group [13]. The NNT can then be estimated according to the simple proportion formula as:

$$NNT = \frac{1}{ARD}$$

with:

- ARD: the absolute risk difference corresponding to:
  - ARD = $p_{ctrl} - p_{exp}$ if the event of interest is unfavorable
  - ARD = $p_{exp} - p_{ctrl}$ if the event of interest is favorable
- p_{ctrl}: the risk or occurrence probability of the event of interest in the control group
- p_{exp}: the risk or occurrence probability of the event of interest in the experimental group.

The more effective the treatment is, the greater the absolute risk difference will be, which in turn will translates in a lower NNT. A utopian goal would be to seek an NNT of 1: treatment would prevent the occurrence of the unfavorable event in all patients. On the opposite, if a treatment had no beneficial effect on the event of interest, the absolute risk difference would be close to zero, and therefore the NNT would be infinite: even if an infinite number of patients were treated, no beneficial effect on the event of interest would be observed. Finally, a negative value of NNT, which is a dystopia for clinical practice, should drive the clinician to consider the treatment as harmful for patients. In this case, one speaks of the Number Needed to Harm (NNH) as opposed to the Number Needed to Benefit (NNB) [14].

Take the example of the EXPLORER study (RCT), which aimed to assess the effect on wound closure of a new type of dressing (sucrose octasulfate) versus a control dressing in patients with a diabetic foot ulcer [15]. The primary endpoint was the proportion of patients with a closed wound at week 20. Forty-eight percent of patients in the experimental group had a closed wound at 20 weeks compared to 30% in the control group (ARD 18%, 95% CI 5–30). Wound closure is obviously beneficial to the patient, hence:

$$NNT = \frac{1}{0.48 - 0.30} = 6.$$

In addition, NNTs should always be presented with their confidence intervals (CI), but this is rarely done [14]. To this end, one should apply the same formula to the inverted bounds of the 95% CI of the absolute risk

difference (in our example 95% CI 5; 30) in order to compute the 95% CI of the NNT:

$$95\%\ CI(NNT) = \left[\frac{1}{0.30}; \frac{1}{0.05}\right] \leftrightarrow [4; 20]$$

After a median follow-up of 20 weeks, 6 patients (95% CI 4–20) would have to be treated with the new type of dressing to allow for one wound closure. NB: in this article, NNT will systematically be rounded upwards.

3. Time-to-event study with varying follow-up periods

In certain randomized controlled studies, the occurrence of the main endpoint is dependent on the duration of follow-up of each patient: these are referred to as *time-to-event* studies ("*the time before the event*"). One of the most evident examples of this type of time-dependent outcome may be the assessment of the effect of one specific intervention on survival or event-free survival. CVOTs are the best examples of these time-to-event studies. Indeed, the main outcome in CVOTs is often a 3P-MACE (a composite CV criterion combining mortality from CV cause, non-fatal myocardial infarction and non-fatal stroke), whose occurrence can't be predicted or controlled, and will eventually occur at a different time point for each patient. Hence, time-to-event studies do not allow to predefine a priori a duration of patient follow-up. Rather, methodologists will pre-specify a number of events to be achieved for the study to reach the required power and statistical significance. The duration of follow-up for each patient and the number of subjects remaining in the study will therefore vary over time. Consequently, the risk of presenting the event of interest will also vary over time and cannot be estimated in the same way as in a binary outcome study with a fixed follow-up period (see above). Survival analysis, adjusted or not, are performed and sometimes a Hazard Ratio (HR) is estimated as well. CVOTs usually report these data.

Thus, the computation of the NNT must take into account the changes in residual risk of the studied population. The NNT is a function of time, whose value varies over time. In other words, there is not a single global NNT representative for one study, but the NNT should be seen as a curve with one value corresponding to each time point t. A NNT at a specific time is interpreted as the average number of patients who will have to be treated to prevent the occurrence of one additional event in the experimental group as compared to the control group at a defined time. Often, the median follow-up time is chosen as the time of particular interest in CVOTs.

In 1999, Altman and Andersen, two renowned statisticians, proposed a method to simply adjust the calculation of NNT for time-to-event studies depending on the type of survival data available (in a publication paper for example) [16]. Basing the calculation on the use of survival analyzes has precisely the advantage of being adjusted according to varying follow-up times and censored data i.e. early dropouts of still "at risk" patients or competing events such as death from another cause than CV.

a. Calculating the NNT from survival probabilities of experimental and control groups

If a survival analysis has been performed, an estimate of the survival probability in each group at a given time point should be available and usually in CVOTs, Kaplan–Meier curves have been generated [16, 17]. The NNT is then calculated from the inverse of the survival probability difference between the two groups at a given time t:

$$NNT(t) = \frac{1}{\mathrm{S}_{exp}(\mathrm{t}) - \mathrm{S}_{ctrl}(\mathrm{t})}$$

with: $S_{exp}(t)$: Probability of event-free survival in the experimental group at time point t; $S_{ctrl}(t)$: Probability of event-free survival in the control group at time point t.

Hence, the absolute risk difference is: $ARD = S_{exp}(t) - S_{ctrl}(t)$.

Regarding the associated 95% CI of the NNT, it might be calculated from the standard error (SE) of each probability of survival by:

$$95\%\ CI(ARD) \Leftrightarrow ARD \pm SE(ARD) \Leftrightarrow \left[L_{Lo}; L_{Up}\right]$$

Hence: $95\%\ CI(NNT) \Leftrightarrow \left[\frac{1}{L_{Up}}; \frac{1}{L_{Lo}}\right]$

with: L_{Lo}, the lower limit of the 95% confidence interval; L_{Up}, the upper limit of the 95% confidence interval.

b. Calculating the NNT from the Hazard Ratio and the survival probability of the control group

If an adjusted survival analysis has been carried out, an estimate of the probability of survival in each group at a given time point is available, and also an estimate of the Hazard Ratio (HR) [16, 17]. The NNT can therefore be calculated from the Cox regression analysis commonly performed in CVOTs by:

$$NNT(t) = \frac{1}{[\mathrm{S}_{ctrl}(\mathrm{t})]^{HR} - \mathrm{S}_{ctrl}(\mathrm{t})}$$

with: HR: Hazard ratio; $S_{ctrl}(t)$: Probability of event-free survival in the control group at time t; $[S_{ctrl}(t)]^{HR}$:

Probability of event-free survival in the experimental group at time t.

NB: the terms "survival" and "mortality" are often confused in the scientific literature. It is the probability of event-free survival S(t) that must be used in this formula, and not the probability of the event p [17]. As a reminder, note that $S(t) = 1 - p$

The calculation of the 95% CI of the NNT should be done using the same formula stated above, by simply replacing the value of the HR with the values of its CI.

Let's take the example of the EMPAREG-OUTCOME study, a CVOT assessing the effect of the SGLT-2 inhibitor empagliflozin on major adverse CV events (3P-MACE) (Fig. 2) [18]. In this trial, 12.1% of patients in the placebo group experienced the primary outcome, which means that 87.9% remained free of event throughout the study. The reported HR was 0.86 (95% CI 0.74–0.99). Using Altman and Andersen's calculation method, the NNT value for the occurrence of 3P-MACE after a median observation time of 3.1 years in the study would be estimated at 63 patients and the associated 95% CI [34–882]. Beware however that a NNT at 63 does not mean that 1 patient will get the full benefit of the treatment while 62 patients will not benefit at all from it; the benefit might actually be shared by all 63 patients, that is each patient will benefit to some extent from the treatment.

4. Highlight on the confidence interval of the NNT

As with all statistical measures of effect, the uncertainty associated with the measure should be reported through its confidence interval (CI) [14]. As discussed above, NNT CIs values shall be calculated by the reciprocal of the confidence interval of the absolute risk difference depending again on the type of data available. Quite often, these CIs are not reported for NNTs, which may be a concern. Indeed, apparent abnormalities may occur especially when statistically significant difference between the groups has not been reached: if the treatment effect is not significant at the 5% alpha risk as compared to placebo, the 95% CI of the absolute risk difference would include the zero value and therefore, the 95% CI of the NNT would include infinity.

Consider the example of the EXSCEL CVOT, conducted with exenatide extended release [19]. In this trial, 11.4% of patients in the experimental group experienced a 3P-MACE, versus 12.2% in the placebo group. The HR

Endpoint	Placebo (N=2333)		Empagliflozin (N=4687)		Hazard Ratio (95%CI)
	No.(%)	No./1000 person-year	No.(%)	No./1000 person-year	
3P-MACE	282 (12.1)	43.9	490 (10.5)	37.4	0.86 (0.74-0.99)

S(t) = 1 – M(t) = 1 – 0.121 = 0.879

$$NNT(t) = \frac{1}{[S_{ctrl}(t)]^{HR} - S_{ctrl}(t)} = \frac{1}{0.879^{0.86} - 0.879} = 63$$

Median observation time : 3.1 years

Fig. 2 Calculation of the NNT from the EMPAREG-OUTCOME study in practice. N: number of patients; HR: Hazard Ratio; 95% CI: 95% confidence interval; S(t): survival rate at time t; M(t): mortality rate at time t

for this comparison was 0.91 (95% CI 0.83–1.00), which was consequently not statistically significant. The NNT value and its 95% confidence interval for the main outcome of the EXSCEL study would therefore be, according to Altman and Andersen's computation method, as follows:

$$NNT(t) = \frac{1}{[S_{ctrl}(t)]^{HR} - S_{ctrl}(t)}$$

$$\text{hence} \quad NNT = \frac{1}{0.878^{0.91} - 0.878} = 97.$$

And the 95% CI: $\left[\frac{1}{0.878^{0.83}-0.878}; \frac{1}{0.878^{1}-0.878}\right] = [51; \infty]$.

The NNT associated with the 3P-MACE of the EXSCEL study with extended release exenatide would then be 97 (95% CI, $51 - \infty$) after a median follow-up period of 3.2 years. Thus the 95% CI of the NNT might imply that exenatide provides no CV benefit to patients because an infinite number of patients could be treated without avoiding one 3P-MACE. Of course, this result must be weighed with the numerous methodological limitations of the study, which could have minimized the size of the effect (number of study dropouts, more concomitant treatments with CV effectiveness in the placebo group).

Critical interpretation of NNTs in CVOTs

Through its simplicity and practicality, the NNT has been increasingly used by the scientific community to account for a therapy's clinical utility. Since 2001, the CONSORT group (CONsolidated Standards Of Reporting Trials) recommends reporting the NNT in the results of randomized controlled trials with binary or time-to-event outcome, in addition to other usual effect measures [20]. However, miscalculations on the one hand and misinterpretation on the other hand may occur while discussing results of CVOTs. Indeed, interpreting an NNT value requires to consider 3 related factors that are not constant across CVOTs: baseline risk, study duration and outcome [21].

The first factor to consider is the baseline risk of the studied population. Indeed, the NNT will inversely vary with the baseline risk, which means that if the baseline risk of a study population is low, one should expect a greater NNT. In most CVOTs, populations with high or very high CV risk were selected, in order to ensure a high absolute risk and therefore a high probability of CV events over a short time period (Table 1; Fig. 3) [18, 19, 22–29]. For example, in the HARMONY-Outcomes study evaluating the CV safety of albiglutide, all patients had very high CV risk or established CV disease, as reflected by the annual placebo primary outcome rate of 5.9 per 100 patient-year [24]. The NNT associated with the 3P-MACE, was 53 (95% CI, 36–116) after a median duration of follow-up of 1.6 years. In contrast, patients enrolled in the REWIND study with dulaglutide were overall at a lower CV risk as evidenced by the annual placebo primary outcome rate of 2.7 per 100 patient-year, even though the proportion of patients at very high and high CV risk is unknown [25, 30]. The NNT associated with 3P-MACE was 67 (95% CI, 38–803) with a median follow-up of 5.4 years. Given the large difference in the absolute risk level of these two populations at baseline

Table 1 Summary patient and study characteristics influencing the NNT value in GLP-1 RA and SGLT-2i CVOTs

CVOT	Drug	Primary outcome	Annual placebo primary outcome rate (N/100 patient-year)	Median follow-up (years)	NNT (according to Altman & Andersen's formula) with 95% CI
GLP-1 RA					
LEADER [22]	Liraglutide	3P-MACE	3.9	3.8	56 [33–243]
SUSTAIN-6 [23]	Semaglutide	3P-MACE	4.4	2.1	45 [28–235]
HARMONY-Outcomes [24]	Albiglutide	3P-MACE	5.9	1.6	53 [36–116]
REWIND [25]	Dulaglutide	3P-MACE	2.7	5.1	67 [38–803]
EXSCEL [19]	Exenatide	3P-MACE	4.0	3.2	Not significant
ELIXA [26]	Lixisenatide	4P-MACE	6.3	2.1	Not significant
PIONEER-6 [27]	Semaglutide (oral)	3P-MACE	3.7	1.3	Not significant
SGLT-2i					
EMPAREG-Outcome [18]	Empagliflozin	3P-MACE	4.4	3.1	63 [34–882]
DECLARE-TIMI58 [28]	Dapagliflozin	CV death or hospitalization for heart failure	1.5	4.2	104 [66–355]
		3P-MACE	2.4	4.2	Not significant
CANVAS [29]	Canagliflozin	3P-MACE	3.15	2.4	Not calculable*

CVOT Cardiovascular Outcomes Trial, *CV* cardiovascular, *3P-MACE* 3 points Major Adverse Cardiovascular Events

*Required data for calculation were not available in the publication paper or supplementary appendix

(See figure on next page.)

Fig. 3 Graphic illustration of annual placebo primary outcome rates and associated NNTs in GLP-1 RA (**a**) and SGLT-2i (**b**) CVOTs. GLP-1 RA: Glucagon Like Peptide-1 receptor agonists; SGLT-2i: Sodium-Glucose Co-Transporter-2 inhibitors; NNT: Number Needed to Treat; CVOTs: cardiovascular outcomes trials; N/100 patient-years: number per 100 patient-years; 95% CI: 95% confidence interval; CV: cardiovascular; HHF: hospitalization for heart failure; NS: not significant; NC: not calculable because required data for calculation were not available in the publication paper or supplementary appendix. *median study follow-up in years; Primary outcome was a 3-points MACE (Major Adverse Cardiovascular Events) for all studies, except ELIXA (4-points MACE) and DECLARE-TIMI58 (co-primary endpoint: 3P-MACE and CV death or HHF); Dark grey bars represent annual placebo primary outcome rates; Light grey bars represent NNTs with 95% CI; regarding data from the REWIND and EMPAREG-Outcome studies, a vertical arrow and 2 slash signs were used to represent the upper limit of their respective 95% confidence intervals for NNTs on a sensible scale

(HARMONY-Outcomes and REWIND), considering an indirect comparison of the two drugs and seeking establishing a power ranking between albiglutide and dulaglutide based on respective NNTs, would be highly inappropriate (or even wrong).

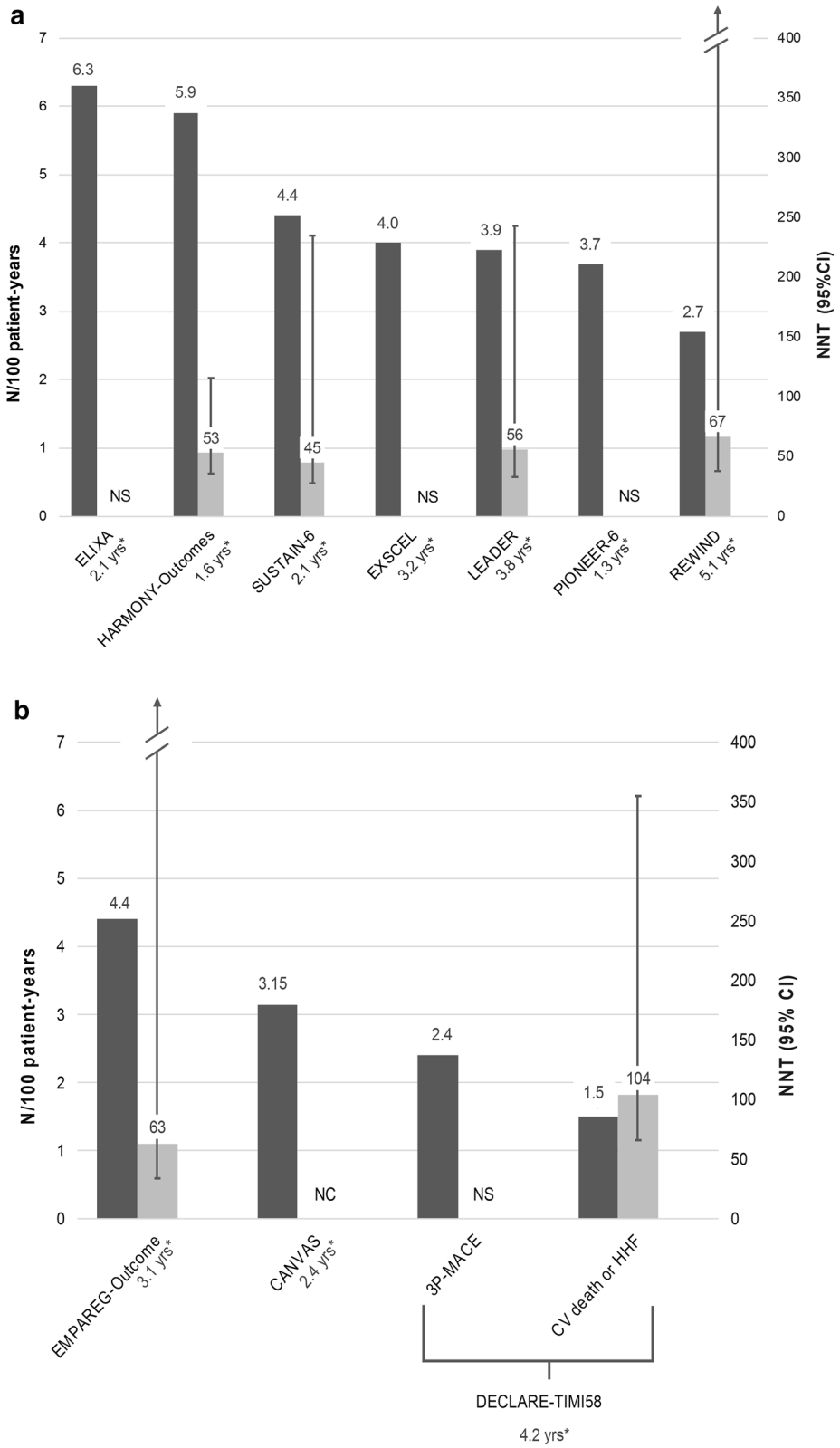
The second factor that must be taken into account is the duration of the study. Each NNT is associated to a specific duration, usually the median follow-up time point. A certainly tempting error would be to seek to standardize study follow-up durations to be able to compare NNTs on a standardized time period [7, 21]. For example, one could imagine converting each specific NNTs of each CVOTs into a standardized 1-year period of follow-up. Again, this would be incorrect because when the follow-up duration increases, the NNT will accordingly tend to decrease since the absolute event rate gets higher. However, such projections to different time frames have been proposed, for instance with ARNI on the basis of data from the PARADIGM-HF trial (27 months median follow-up) in order to estimate the 5-year NNT [10]. Despite the use of a sophisticated statistical model, data generated should be considered as exploratory and take the limitations underlined by the authors into account. Besides, CVOTs are typically long duration studies, which could potentially leave competing events, such as a death from another cause, come into play and influence the occurrence of the event of interest [31]. Thus, as NNT values will vary non-linearly over time, extrapolating some NNT results to a different time horizon, shorter or longer, would be inappropriate. It is common sense for any clinician to say that treating 60 patients for 3 years would not be as effective as treating 180 patients for 1 year.

And thirdly, the outcome itself plays a role. A NNT is specific to a defined study endpoint, so that the NNT of each endpoint of interest should be taken into account to interpret the overall benefit/risk balance of a treatment Take the example of the DECLARE-TIMI58 study with dapagliflozin designed with two co-primary endpoints: a 3P-MACE and a composite of CV mortality and hospitalization for heart failure [28]. The associated NNT were respectively, 160 and 104 after 4.2 years of treatment (Note: the comparison of the two groups regarding the 3P-MACE endpoint was not significant, which questions the interpretation of the associated NNT CI. See previous section). Finally, one might also imagine calculating a NNT for safety parameters in addition to the efficacy ones. In the example of the DECLARE-TIMI58 study, a serious adverse event would occur every 38 patients treated.

In conclusion, an NNT should not be considered as an absolute measure of the overall clinical benefit of a treatment. A difference in NNT across trials may be attributable to a true difference in treatment effectiveness as much as a difference in patients' baseline risk or any other CVOT characteristic. Further, the treatment benefit based on an estimated NNT should be, as always, weighed with its toxicity, and possibly its cost for an overall assessment of the drug's efficiency [32]. Finally, indirect comparisons between NNTs of two separate studies should be avoided since they are not adjusted in the same way [21].

Conclusion

Given the great medical impact of CVOTs, a correct calculation of the NNT is a fundamental point, and the first step to an appropriate interpretation. In the absence of cardio-renal events studies conducted under the same conditions, there is to date no justification for establishing a strong ranking of power or efficiency between the different GLP-1 RA and SGLT-2i based on NNTs. However, NNTs could provide input to the design of a medico-economic model taking efficacy and safety parameters as well as cost into account, in order to have a full picture of these drugs cost-effectiveness.


a
N/100 patient-years
NNT (95%CI)
6.3
5.9
4.4
4.0
3.9
3.7
2.7
53
45
56
67
NS
NS
NS
ELIXA 2.1 yrs*
HARMONY-Outcomes 1.6 yrs*
SUSTAIN-6 2.1 yrs*
EXSCEL 3.2 yrs*
LEADER 3.8 yrs*
PIONEER-6 1.3 yrs*
REWIND 5.1 yrs*
b
N/100 patient-years
NNT (95% CI)
4.4
3.15
2.4
1.5
63
104
NC
NS
EMPAREG-Outcome 3.1 yrs*
CANVAS 2.4 yrs*
3P-MACE
CV death or HHF
DECLARE-TIMI58
4.2 yrs*

Abbreviations
ADA/EASD: American Diabetes Association/European Association for the Study of Diabetes; ARD: Absolute risk difference; ARNI: Angiotensin receptor–neprilysin inhibitor; CV: Cardiovascular; CVD: Cardiovascular disease; CVOT: Cardiovascular Outcomes Trial; ESC: European Society of Cardiology; GLP-1 RA: Glucagon-like-peptide receptor agonist; HR: Hazard ratio; NNT: Number Needed to Treat; PCSK-9: Proprotein convertase subtilsin-kexin type 9; RCT: Randomized controlled trial; SGLT-2i: Sodium-glucose co-transporter-2 inhibitors; T2D: Type 2 diabetes; 3P-MACE: 3-Point major adverse cardiovascular events.

Acknowledgements
Not applicable.

Authors' contributions
LL has drafted the manuscript, table and figures. PD and BG have substantively revised the manuscript. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
Dr. LUDWIG reports personal fees from ELI LILLY and Company, outside the submitted work; Dr. DARMON reports personal fees from Eli Lilly, personal fees from Novo Nordisk, personal fees from Sanofi, personal fees from Astra Zeneca, personal fees from Merck Sharpe & Dohme, personal fees from Boehringer Ingelheim, personal fees from MundiPharma, personal fees from Novartis, personal fees from Abbott, outside the submitted work; Dr. GUERCI reports grants, personal fees and non-financial support from Sanofi, personal fees from Orkyn (provider), grants, personal fees and non-financial support from Vitalaire/air liquide (provider), personal fees and non-financial support from Abbott, personal fees and non-financial support from Novo Nordisk, grants, personal fees and non-financial support from Eli Lilly, personal fees and non-financial support from Boehringer Ingelheim, personal fees and non-financial support from roche diagnostic, personal fees and non-financial support from Astra Zeneca, personal fees and non-financial support from MSD, grants and personal fees from medtronic, personal fees and non-financial support from dinno santé (provider), outside the submitted work.

Author details
[1] Endocrinology, Diabetology & Nutrition, CHRU of Nancy, Brabois Hospital, Lorraine University, 54500 Vandoeuvre-lès-Nancy, France. [2] Endocrinology, Metabolic Diseases and Nutrition Department, Assistance Publique Hôpitaux de Marseille, Marseille, France. [3] INSERM, INRA, C2VN, Aix Marseille University, 13005 Marseille, France.

Received: 5 April 2020 Accepted: 7 May 2020
Published online: 13 May 2020

References
1. Kristensen SL, Rørth R, Jhund PS, Docherty KF, Sattar N, Preiss D, et al. Cardiovascular, mortality, and kidney outcomes with GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. Lancet Diabetes Endocrinol. 2019;7:776–85.
2. Zelniker TA, Wiviott SD, Raz I, Im K, Goodrich EL, Furtado RHM, et al. Comparison of the effects of glucagon-like peptide receptor agonists and sodium-glucose Cotransporter 2 inhibitors for prevention of major adverse cardiovascular and renal otcomes in type 2 diabetes mellitus. Circulation. 2019;139:2022–31.
3. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, et al. Management of hyperglycemia in type 2 diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care. 2018;2018(41):2669–701.
4. Buse JB, Wexler DJ, Tsapas A, Rossing P, Mingrone G, Mathieu C, et al. 2019 Update to: management of hyperglycemia in type 2 diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care. 2020;2020(43):487–93.
5. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J. 2020;41:255–323.
6. Hildebrandt M, Vervölgyi E, Bender R. Calculation of NNTs in RCTs with time-to-event outcomes: a literature review. BMC Med Res Methodol. 2009;9:21.
7. Suissa S. The Number Needed to Treat: 25 years of trials and tribulations in clinical research. Rambam Maimonides Med J. 2015;6:e0033.
8. Seidu S, Kunutsor SK, Sesso HD, Gaziano JM, Buring JE, Roncaglioni MC, et al. Aspirin has potential benefits for primary prevention of cardiovascular outcomes in diabetes: updated literature-based and individual participant data meta-analyses of randomized controlled trials. Cardiovasc Diabetol. 2019;18:70.
9. Wong ND, Shapiro MD. Interpreting the findings from the recent PCSK9 monoclonal antibody cardiovascular outcomes trials. Front Cardiovasc Med. 2019;6:14.
10. Srivastava PK, Claggett BL, Solomon SD, McMurray JJV, Packer M, Zile MR, et al. Estimated 5-year Number Needed to Treat to prevent cardiovascular death or heart failure hospitalization with angiotensin receptor-neprilysin inhibition vs standard therapy for patients with heart failure with reduced ejection fraction: an analysis of data from the PARADIGM-HF Trial. JAMA. 2018;3:1226–31.
11. Sposito AC, Berwanger O, de Carvalho LSF, Saraiva JFK. GLP-1RAs in type 2 diabetes: mechanisms that underlie cardiovascular effects and overview of cardiovascular outcome data. Cardiovasc Diabetol. 2018;17:157.
12. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318:1728–33.
13. Cook RJ, Sackett DL. The Number Needed to Treat: a clinically useful measure of treatment effect. BMJ. 1995;310:452–4.
14. Altman DG. Confidence intervals for the number needed to treat. BMJ. 1998;317:1309–12.
15. Edmonds M, Lázaro-Martínez JL, Alfayate-García JM, Martini J, Petit J-M, Rayman G, et al. Sucrose octasulfate dressing versus control dressing in patients with neuroischaemic diabetic foot ulcers (Explorer): an international, multicentre, double-blind, randomised, controlled trial. Lancet Diabetes Endocrinol. 2018;6:186–96.
16. 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.
17. Bowry SK, Schoder V, Apel C. An inadvertent but explicable error in calculating Number Needed to Treat for reporting survival data. J Am Soc Nephrol. 2014;25:875–6.
18. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med. 2015;373:2117–28.
19. Holman RR, Bethel MA, Mentz RJ, Thompson VP, Lokhnygina Y, Buse JB, et al. Effects of once-weekly exenatide on cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2017;377:1228–39.
20. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. 2010;63:e1–37.
21. McAlister FA. The « number needed to treat » turns 20–and continues to be used and misused. Can Med Assoc J. 2008;179:549–53.
22. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2016;375:311–22.
23. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, et al. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med. 2016;375:1834–44.

24. Hernandez AF, Green JB, Janmohamed S, D'Agostino RB, Granger CB, Jones NP, et al. Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. Lancet. 2018;392:1519–29.
25. Gerstein HC, Colhoun HM, Dagenais GR, Diaz R, Lakshmanan M, Pais P, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. Lancet. 2019;394:121–30.
26. Pfeffer MA, Claggett B, Diaz R, Dickstein K, Gerstein HC, Køber LV, et al. Lixisenatide in patients with type 2 diabetes and acute coronary syndrome. N Engl J Med. 2015;373:2247–57.
27. Husain M, Birkenfeld AL, Donsmark M, Dungan K, Eliaschewitz FG, Franco DR, et al. Oral semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med. 2019;38:841–51.
28. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2019;380:347–57.
29. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. N Engl J Med. 2017;377:644–57.
30. Gerstein HC, Colhoun HM, Dagenais GR, Diaz R, Lakshmanan M, Pais P, et al. Design and baseline characteristics of participants in the researching cardiovascular events with a weekly INcretin in Diabetes (REWIND) trial on the cardiovascular effects of dulaglutide. Diabetes Obes Metab. 2018;20:42–9.
31. Gouskova NA, Kundu S, Imrey PB, Fine JP. Number Needed to Treat for time-to-event data with competing risks. Stat Med. 2014;33:181–92.
32. Chuang LH, Verheggen BG, Charokopou M, Gibson D, Grandy S, Kartman B. Cost-effectiveness analysis of exenatide once-weekly versus dulaglutide, liraglutide, and lixisenatide for the treatment of type 2 diabetes mellitus: an analysis from the UK NHS perspective. J Med Econ. 2016;19:1127–34.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC