Reflections on using non-inferiority randomised placebo controlled trials in assessing cardiovascular safety of new agents for treatment of type 2 diabetes

Denise Campbell-Scherer

10.1136/ebmed-2017-110685

Department of Family Medicine and Alberta Diabetes Institute, University of Alberta, Edmonton, Alberta, Canada

Correspondence to **Dr Denise Campbell-Scherer**, Alberta Diabetes Institute, Clinical Research Unit, 2-004 Li Ka Shing Ctr, 87 Ave and 112 St. Edmonton, Alberta T6G 2E1, Canada; denise.campbell-scherer@ualberta.ca

Abstract

The 2008 Food and Drug Administration (FDA) guidance to industry requires experimental evidence that new agents to treat type 2 diabetes do not have an unacceptable increase in cardiovascular risk. They specify this unacceptable increase to be a risk ratio of 1.3 in non-inferiority trials which may use placebo control. Clinically, this means that if a new agent achieves this threshold of not being 30% worse than placebo it is declared 'non-inferior'. This guidance was in response to safety concerns raised about medications approved on their basis of reducing glycated haemoglobin alone. There was concern that this FDA guidance would stifle new drugs coming to market. On the contrary, there have been a number of exciting new classes of agents approved with improved confidence that they reduce glycated haemoglobin, and that they also do not excessively increase cardiovascular risk. Cardiovascular safety trials have been conducted for a number of novel medications using a non-inferiority approach. However, clinicians need to recognise that the results of non-inferiority trials are not as credible as superiority trials. It is important to closely review the trials before accepting claims of 'non-inferiority' or 'cardiac neutrality' especially when these studies are often compared with placebo, and may be accepting estimates of effect which span potentially clinically meaningful harm. There are compelling reasons to further investigate agents showing promise in non-inferiority trials with superiority trials, which include prespecified subgroups, and with sufficient power and duration to provide robust estimates of harms and benefits to inform clinical decision-making.

The commentary on the trial by Marso *et al*[1] assessing the cardiovascular (CV) safety of a new agent for treating type 2 diabetes highlights a recognised dilemma in clinical research on CV risk in assessment of new therapeutic agents. That dilemma is that the baseline risk of CV events in the population is reduced with better adherence to measures known to reduce risk like statins and blood pressure control; as a result the incremental benefit of new agents is smaller.[2] Recognition of the limitations of surrogate end points has resulted in increased focus on reporting on patient-important outcomes in clinical trials.[2] This means that the cost of doing superiority efficacy studies for benefit on CV outcomes is increased.[3] This expense is justified for studying investigational agents that have been demonstrated to be safe with a good effect on surrogate outcomes, but is prohibitive in the preapproval phase.[3]

In response to safety concerns raised regarding diabetic agents approved on their ability to reduce glycated haemoglobin, the Food and Drug Administration (FDA) in the USA introduced new guidance to industry in 2008.[4] As per the FDA guidance, in order for investigational agents to be approved they added the requirement that they demonstrate no excess increase in CV risk in large safety outcome studies. They suggested non-inferiority (NI) studies of the new agent compared with placebo may be used. Clinicians are familiar with randomised controlled trials (RCTs), which compare a new and a standard treatment (or placebo) to determine whether the new treatment is superior. Another approach is the NI trial. In the NI RCT the question is whether the new treatment is not worse than the control by a defined amount called the NI margin. NI RCTs have the potential to lead to invalid conclusions.[2 5] It is important that clinicians understand the limitations of these trials and how to interpret their results.[6 7]

It is reasonable to consider the NI design if the new intervention is anticipated to have meaningful benefits aside from those related to increased effectiveness compared with standard therapy. Examples include treatments that are less costly, toxic or have fewer side effects, and where it would be unethical to use a placebo.[5 6] NI RCTs have historically not been used if the standard therapy has not been proven to have clinically meaningful efficacy, as demonstrated with adequate precision in robust trials.[5 6]

The choice of NI margin is critically important. Statistically, usually the NI margin is set to preserve 50% of the lower bound of the CI of the magnitude of the effect of the active control versus placebo as determined in systematic review meta-analysis of historic trials.[5] It is statistically reasonable to test for superiority in NI trials. From the clinicians' and patients' perspectives the choice of the NI margin has to make clinical sense. The choice of the NI margin drives the number of events required to show NI.[3] This is illustrated in table 1. From a cost perspective a generous NI margin means that fewer events are required. Small numbers of events can affect the ability of the trial to detect harms in important secondary outcomes and prespecified subgroups.

The FDA considered that there is evidence that improved glycated haemoglobin is associated with improved microvascular complications of type 2 diabetes, so they stipulated the need for confidence that there is not a 30% increase in excess CV morbidity and mortality compared with placebo, determined through independent, adjudicated evaluation.[4] In the preapproval phase it is sufficient to demonstrate there is not an 80% increase in excess CV morbidity and mortality.[4] However, if the studies in the preapproval phase do not exclude the excess risk of 30%, then the FDA requires a separate postmarketing randomised safety trial.

Linked

► http://dx.doi.org/10.1136/ebmed-2016-110652

Table 1 Estimation of the cardiovascular events required to fall below the FDA target cut-off in the design of a cardiovascular safety non-inferiority trial as a function of the expected true HR of the intervention (for 90% power)[8]

True HR	Upper HR boundary	<1.8	<1.3
0.70		48	110
0.75		55	139
0.80		64	179
0.85		75	233
0.90		88	311
0.95		103	428
1.0		122	611
1.05		145	921
1.1		174	1507

Non-inferiority margin.
FDA, Food and Drug Administration.

Clinically, this means that if a new agent achieves this threshold of not being 30% worse than placebo it is declared 'non-inferior'.

There are excellent reviews summarising considerations of some of the different approaches for doing this work.[3 4 8] The CV events required to achieve these generous FDA cut-offs for 'NI' are summarised in table 1.

The use of NI studies in assessing the CV safety of new agents compared with placebo is a significant departure from the usual rationale supporting their use. The use in this case is to provide a 'quick look' to rule out excessive CV risk of the agent compared with placebo. In order to do this expeditiously, high-risk populations with their associated higher CV event rates are selected for these preapproval studies.[3] There is incentive to have the event rate accumulate quickly as there is no expectation of superiority so there is no advantage to run the trial for a long period of time in order to enhance the differences in the treatment arms.[3]

One wrinkle is that these are event-driven trials.[8] In fact, the studies 'may employ group sequential designs that permit interim monitoring of accumulating data with the possibility of stopping for efficacy or futility', which is discussed in detail by Geiger *et al*.[8] It is well known that stopping trials early for benefit can result in implausible estimates of efficacy in superiority RCTs.[9] Unlike superiority RCTs where the numbers of CV death are typically low compared with myocardial infarctions and strokes, in these studies CV death rates are much higher. Consider the examples of the two CV safety NI trials that concluded superiority in the primary outcome (composite of CV death, non-fatal myocardial infarction and non-fatal stroke) compared with placebo in table 2. In both studies, all patients were encouraged to have optimal diabetes management, for example, with statins, good blood pressure control and good glycaemic control.

It is interesting to note that in EMPA-REG and LEADER, the significance of the primary composite outcome of time to event of non-fatal myocardial infarction, non-fatal stroke and CV death is being driven by CV death, with neither stroke nor myocardial infarction achieving significance. In fact, in EMPA-REG the CI on the HR for stroke extends to a 67% increase in non-fatal stroke compared with placebo at the upper bound. In the discussion of the EMPA-REG study the investigators comment, "Notably, reductions in the risk of cardiovascular death and death from any cause occurred early in the trial and the benefits persisted throughout the trial." Is it possible that in this high-risk population with a high rate of CV death that through random chance the death events happened earlier than

Table 2 Outcomes of the two CV safety NI studies that claim superiority

	Placebo (N=2333) Number (%)	Incident rate Number of events/1000 patient-years	Epaglifozin (N=4687) Number (%)	Incident rate Number of events/1000 patient-years	HR (95% CI)	p Value
Primary composite outcome	282 (12.1)	43.9	490 (10.5)	37.4	0.86 (0.74 to 0.99)	p<0.001 NI p=0.04 superiority
Non-fatal MI	121 (5.2)	18.5	213 (4.5)	16	0.87 (0.70 to 1.09)	0.22
Non-fatal CVA	60 (2.6)	9.1	150 (3.2)	11.2	1.24 (0.92 to 1.67)	0.16
CV death	137 (5.9)	20.2	172 (3.7)	12.4	0.62 (0.49 to 0.77)	<0.001
	Placebo (N=4672) Number (%)	Incident rate Number of events/ 100 patient-years	Liraglutide (N=4668) Number (%)	Incident rate Number of events/ 100 patient-years	HR (95% CI)	p Value (superiority)
Primary composite outcome	694 (14.9)	3.9	608 (13)	3.4	0.87 (0.78 to 0.97)	0.01
Non-fatal MI	317 (6.8)	1.8	281 (6)	1.6	0.88 (0.75 to 1.03)	0.11
Non-fatal CVA	177 (3.8)	1.0	159 (3.4)	0.9	0.89 (0.72 to 1.11)	0.30
CV death	278 (6)	1.6	219 (4.7)	1.2	0.78 (0.66 to 0.93)	0.007

EMPA-REG outcome[10]—epagliflozin N=7020, mean duration of observation 3.1 years, NI margin 1.3, sample size 691 events required for NI with 90% power assuming HR 1.0. Duration of trial until reached adjudicated outcome of at least 691 patients. Primary outcome is the time to event of a composite of death from CV causes (CV death), non-fatal MI or non-fatal CVA.
LEADER[11]—liraglutide N=9340, mean duration of follow-up 3.8 years, NI margin 1.3, sample size 611 events required for a NI with 90% power assuming a HR 1.0. Duration of the trial would be a follow-up of minimum of 42 weeks and maximum of 60 months receiving intervention and 30 days of follow-up afterward. The data safety monitoring board could terminate for adverse events or clear evidence of benefit in the liraglutide arm (significance level determined p<0.001). Primary outcome is the time to event of a composite of death from CV causes (CV death), non-fatal MI or non-fatal CVA.
CV, cardiovascular; CVA, stroke; NI, non-inferiority; MI, myocardial infarction.

in other NI trials of other agents with the same composite outcome that did not demonstrate this finding? It would be interesting to see superiority RCTs of these agents to see if the result would be repeated.

This NI approach has resulted in large-scale, adequately powered, end point adjudicated trials that have demonstrated that some new agents do not reduce risks of CV events and identified adverse events and side effects, which may have previously become apparent only with postmarketing surveillance.[3]

Menon and Lincoff[3] highlight that the downside of this NI approach is that patients with low-risk and early disease might actually have greater potential to reverse and mitigate the damage caused by diabetes mellitus but are not the focus for NI trials. This regulation may disincentivise conducting long-term trials designed to focus on benefit. This reduces rich data on the long-term effects of treatment benefit in important subgroups, and adequate power to show robust estimates of harms.

Clinically, a rational consumer might entertain a small risk that the new agent is actually slightly less effective, if it provided other tangible benefits. It is more difficult to understand the appeal of an improvement in glycaemic control (associated with improved microvascular complications) in exchange for any worsening of CV risk compared with placebo. As in other areas like HIV research, which use NI trials extensively, there are calls for follow-up superiority trials for promising agents.[12]

With the FDA guidance, the NI approach has resulted in a number of exciting new agents approved for use in type 2 diabetes, with confidence that they are not associated with excessive CV harms compared with placebo. Agents showing promise with this approach should be followed by appropriate superiority trials, with prespecified subgroups with sufficient power and duration to provide robust estimates of harms and benefits to inform clinical decision-making.

Competing interests DCS has received support from the Canadian Obesity Network for an unrestricted educational grant for primary care obesity training supported by Novo Nordisk.

Provenance and peer review Commissioned; internally peer reviewed.

References

1. **Marso SP**, Bain SC, Consoli A, *et al*. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. *New Engl J Med* 2016;**375**:1834–44.
2. **Snapinn SM**, Noninferiority trials. *Trials* 2000;**1**:19.
3. **Menon V**, Lincoff AM. Cardiovascular safety evaluation in the development of new drugs for diabetes mellitus. *Circulation* 2014;**129**:2705–13.
4. **Food and Drug Administration.** Guidance for industry—diabetes mellitus evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes. http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/UCM071627.pdf (accessed 13 Feb 2017).
5. **Fleming TR**. Current issues in non-inferiority trials. *Statist Med* 2008;**27**:317–32.
6. **Mulla SM**, Scott IA, Jackevicius CA, *et al*. How to use a noninferiority trial: users' guides to the medical literature. *JAMA* 2012;**308**:2605–11.
7. **Piaggio G**, Elbourne DR, Pocock SJ, *et al*. Reporting of noninferiority and equivalence randomized control trials. Extension of the CONSORT 2010 statement. *JAMA* 2012;**308**:2594–604.
8. **Geiger MJ**, Mehta C, Turner JR, *et al*. Clinical development approaches and statistical methodologies to prospectively assess the cardiovascular risk of new antidiabetic therapies for type 2 diabetes. *Ther Innov Regul Sci* 2015;**49**:50–64.
9. **Bassler D**, Montori VM, Devereaux P, *et al*. Randomized trials stopped early for benefit. In: Guyatt G, Rennie D, Meade MO, Cook DJ. eds. Users' guides to the medical literature: a manual for evidence-based clinical practice. 3rd edn. New York, NY: McGraw-Hill, 2015;135–42.
10. **Zinman BZ**, Wanner C, Lachin JM, *et al*. Epagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med* 2015;**373**:2117–28.
11. **Marso SP**, Daniels GH, Brown-Frandsen K, *et al*. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2016;**375**:311–22.
12. **Kanters S**, Mills EJ, Thorland K, *et al*. Antiretroviral therapy for initial human immunodeficiency virus/AIDS treatment: critical appraisal of the evidence from over 100 randomized trials and 400 systematic reviews and meta-analyses. *Clin Microbiol Infect* 2014;**20**:114–22.

Reflections on using non-inferiority randomised placebo controlled trials in assessing cardiovascular safety of new agents for treatment of type 2 diabetes

Denise Campbell-Scherer

Evid Based Med 2017 22: 54-56 originally published online March 8, 2017
doi: 10.1136/ebmed-2017-110685