Review Article

Cardio vascular outcomes trials (CVOTS) with anti-hyperglycemic agents: demystifying statistical complexities

Abstract

There is an explosion of cardiovascular outcome trials (CVOTs) in the last decade or so. Each of these studies generated more questions than providing clinically useful answers. The specific aims of the study were blurred by a lot of attention directed towards secondary end points sometimes diverting our attention from the primary hypothesis being tested. One of the principle reasons behind this phenomenon has been our poor understanding of statistical analysis. This drawback resulted in a non-systematic approach to analyzing the CV outcome trial and the resulting confusion. It is after a very long time that endocrinologists and diabetologists have found some meaning to managing diabetes after UKPDS helped us in becoming one. Instead of basking in the glory of the new found treasure and taking it forwards with intense vigor we saw our fraternity going into a shell. Criticism of the data is important as it helps in improving the trial designs further in the long run. However this should not shift our focus away from the robust positive findings at hand. This review attempts at demystifying certain aspects of statistical analysis frequently encountered in CV outcome trial and develop a personal understanding on the same instead of depending on conflicting interpretations floated at different continuing medical educations (CMEs) and Diabetes Congresses.

Keywords: statistics, diabetes, CVOTS

Samit Ghosal,[1] Binayak Sinha,[2] Kingshuk Bhattacharjee[3]
[1]Nightingale Hospital, India
[2]AMRI Hospitals, India
[3]Kingshuk Bhattacharjee Senior MSL, India

Correspondence: Samit Ghosal, Nightingale Hospital, 11 Shakespeare Sarani, Kolkata, India,
Email ramdasghosal@gmail.com

Introduction

The rosiglitazone story was a turning point to the way regulatory authorities started viewing anti-hyperglycemic agents.[1] The focus shifted from glucose lowering to safety. Strict glycemic control was associated with impressive reduction in microvascular end points as documented in the UKPDS trial.[2] But the holy grail of type 2 diabetes has always been reducing cardiovascular end points. Unfortunately none of the glucose lowering trials was able to document this benefit. UKPDS was conducted in young, recently diagnosed type 2 diabetic patients and failed to demonstrate any CV benefit during the trial period.[2] There were CV benefits in the 10years follow-up phase but not during the trial period. In the older, longer duration of diabetes cohort there was either CV harm or no benefit with stringent glycemic control.[3–5] Hence the whole attention was turned towards the impact of molecules used just like the pleiotropic benefits associated with statins or RAAS blockade.

After the adverse CV outcomes documentation with rosiglitazone, regulatory authorities focused on safety issues associated with anti-hyperglycemic agents. The search for CV benefits with use of these drugs was not the primary focus. FDA came out with a set of rules for the pharmaceutical industry whereby assessing CV safety became mandatory.[6] After a string of non-inferiority trials we were suddenly gifted with a CV superiority data, which came as a sweet surprise.[7] This was followed by a couple of more similar data.[8,9] There was both excitement as well as suspicion in the endocrinologists mind regarding these data. These trial data were being presented at every major diabetes congress with varying interpretations. To a physician this came as a rude shock to see the same data being presented in both positive and negative vein. Unfortunately the only way to decode this jargon was to have a solid understanding of statistical analysis.

This review is an attempt to simplify some aspects of the regulatory requirements for trials and its interpretation.

CVOT Trial designs

We have been traditionally tuned to the fact that CV benefits can be demonstrated only by a very long-term follow up trial. This concept was created after the positive CV effects seen in the 10years follow up data from UKPDS.[10] However CV benefits in SUSTAIN 6 trail was documented in a very short period of time (2.1years).[9] How was it possible?

CV outcome trials were designed in such a way so as to accumulate a large number of events in a short span of time. According to the FDA recommendations the number of events should be predicted in advance using a prediction model. At the end of the analysis, the upper limit of two-sided 95% confidence interval (CI) for the hazard ratio or risk ratio (relative risk) should be less than 1.8 to satisfy the non-inferiority margin (Table 1).[6] However the pharmaceutical company in question must conduct a dedicated CV outcome trial. If the data analysis points at 95% upper limit CI below 1.3, it is not mandatory to perform a CV outcome trial.[6] With this criterion only most of the gliptins could have opted out of performing a dedicated CV outcome study. Here lies a minor catch. Most of the pooled phase 2 & 3 trials with gliptins had very small number of CV events and hence it cannot be comprehensively stated that the molecule is safe or unsafe.[11,12] Hence all the gliptins (except Vildagliptin) underwent a dedicated CV outcome trial. It was calculated that a minimum of 611 events were needed to satisfy the 1.3 upper limit 95% CI bar and 122 events for <1.8.[6] This aspect was factored in by most of the study designs with a background prevalence of CV events in the population being tested. Most of the CVOTs based their calculations taking a background CV

prevalence rate of 2%-3% (Table 2). Increasing the sample size and recruiting exclusively those with established CVD could reduce the number of years of follow-up required to reach the magic figure.

Table 1 USFDA Recommendations to the Industry

Upper bound of two-sided 95% confidence interval (CI)	No of CV required events	Approval recommendation
<1.8	122	Dedicated CVOT may be required
<1.3	611	Dedicated CVOT not required

Table 2 CVOT Trials designs: Background annual CV prevalence rates

Study	Background estimated annual cv event rate(%)
SAVOR-TIMI 53[13]	2.8
EXAMINE[14]	3.5
TECOS[15]	2.5-3.0
EMPA-REG[7]	1.5
CANVAS[16]	2.25
LEADER[8]	1.8
SUATAIN 6[9]	1.98

Getting approval

It would take a long time for the anti-hyperglycemic agent to hit the market if we had to wait for the CV outcome trial to end. Analyzing CV events from pooled phase 2&3 data would not be valid due to very small number of events. However if we could get hold of an interim analysis whereby the safety of the drug is established and continue with the CV outcome trial simultaneously, then the drug would be available for use earlier. There are three possible strategies by which the above-mentioned end could be achieved:

a. Meta-analysis of pooled phase 2&3 trials - In case of evaluating an agent of a group of drugs which has already demonstrated CV benefit, it might be reasonable to include more high-risk participants in the phase 3 trials. The limitation of this approach is that the number of CV events would still be insufficient to discharge the 1.3 upper bound CI. Hence the regulatory authorities could still request for a separate CV outcomes trial post approval.

Example: Saxagliptin performed a meta-analysis of pooled phase 2 & 3 trials prior to marketing approval. This retrospective analysis satisfied the upper bound CI bar of <1.3. However the number of CV events was only 41.[11] Hence it was logical that a separate CVOT was conducted. Vildagliptin followed a similar pre-approval analysis with an exception. The phase 3 trials were pre-adjudicated. The pooled phase 2 & 3 data comfortably passed the <1.3 upper bound CI bar for the 50mg twice-daily dose (n=81).[17] However the upper bound CI for vildagliptin 50mg once daily arm was 2.11(n=10).[17] As evident from the discussions above, these discordant notes cropped up due to the small numbers being evaluated and hence a dedicated CVOT could settle the issue conclusively.

b. Design a single large CVOT- After around 24 weeks, open up the study for an interim analysis of efficacy and demonstrating MACE+ upper bound CI of <1.8. This would be sufficient to get a marketing approval. The study continues to collect CV events after the interim analysis until sufficient numbers of CV events have been reached to discharge the <1.3 upper bound CI bar.

Example: The EXAMINE Trial. It was an unique trial since both the non-inferiority upper bound CI of <1.8 as well as <1.3 were tested in a single trial without the need to perform an interim analysis with combined pooled phase 2 & 3 data.[18] There were 4 interim analysis planned prospectively. If the upper bound CI of <1.8 was not met at any of these analysis the study would be stopped. Once the 1.8 bar was satisfied and the drug forwarded for marketing approval, the trial was continued until the next analysis and the final interim analysis would be conducted when the critical number of CV events to analyze for the 1.3 bar was procured.

c. A meta-analysis & CVOT: In case the drug in question is the first in class to design a CVOT this approach is commonly employed. The CVOT is initiated at a time when phase 3 trials are on. A meta-analysis of pooled data from phase 2 & 3 trials with or without CV event enrichment from the CVOT is used for discharging the <1.8 upper bound CI and submitting the data for marketing approval. MACE data from the dedicated CVOT is utilized to demonstrate <1.3 upper bound CI for non-inferiority and then <1.0 upper bound CI for superiority as part of sequential analysis. MACE+ can also be used as a primary end point if the sole purpose of the study is to achieve non-inferiority. The very recent CVOT summit of EASD endorsed the view that 3P-MACE should be used as the primary end point.[19]

Example: EMPA-REG Outcome Trial. In a single trial both efficacy of the agent against standard of care, non-inferiority upper bound CI bar of <1.8 at an interim analysis and subsequent sequential non-inferiority upper bound CI of <1.3 and <1.0 respectively were analyzed.[7]

MACE versus MACE+

It is confusing to note that these two terms were used interchangeably as primary end-points in several CV outcome trials (Table 3). It is now acceptable to include hospitalization due to unstable angina as the plus component of MACE+. The major difference between MACE & MACE+ is that the fourth component in the later is extremely subjective. Hence MACE+ should be used to increase the number of CV events required to prove the null hypothesis i.e. as a part of the interim analysis to prove the <1.8 upper bound CI bar. However MACE should be universally used as the primary composite end point testing hard end points and sequentially prove both the <1.3 non-inferiority bar as well as <1.0 superiority bar. In order words MACE should be used to disprove the null hypothesis whereas MACE+ used to prove it. The very recent CVOT summit of EASD endorsed the view that 3P-MACE should be used as the primary end point.[19]

Example: EMPA-REG Outcomes Trial. Hard-core CV primary end point (MACE) met both the <1.3 & <1.0 upper bound CI bar (95.02% confidence interval, 0.74 to 0.99; P<0.001 for non-inferiority and P=0.04 for superiority). While MACE+ proved the null hypothesis (95% CI, 0.78 to 1.01).[7]

Table 3 Primary end-points in CVOT trials

Studies	Primary MACE (3-P MACE)	Primary MACE+
SAVOR-TIMI 53[13]	Composite of CV death, nonfatal myocardial infarction, or nonfatal ischemic stroke.	
EXAMINE[14]	Composite of CV death, nonfatal myocardial infarction, or nonfatal ischemic stroke.	
TECOS[15]		Composite of CV death, nonfatal myocardial infarction, nonfatal ischemic stroke or hospitalization due to unstable angina.
EMPA-REG[7]	Composite of CV death, nonfatal myocardial infarction, or nonfatal ischemic stroke.	
CANVAS[16]	Composite of CV death, nonfatal myocardial infarction, or nonfatal ischemic stroke.	
LEADER[8]	Composite of CV death, nonfatal myocardial infarction, or nonfatal ischemic stroke.	
SUATAIN 6[9]	Composite of CV death, nonfatal myocardial infarction, or nonfatal ischemic stroke.	

Primary versus secondary end points

A large number of end points are analyzed as part of secondary end points. In most of the articles published on CVOT we find larger space allocated to secondary end points, sub-group and post hoc analysis. What is the importance of such end point analysis? Are they hypothesis generating or have clinical significance?

The study design section stresses on the importance of pre-adjudicating the primary composite end points as well as making adequate predictions on the study population to induce sufficient power in the study. Both these strategies are employed to take care of type I & type II errors (Table 4).[20] If we analyze secondary end point data, which were neither pre-specified nor pre-adjudicated, they are riddled with type II errors and lack the power to predict accurately. This can be overcome by increasing the study sample, which takes care of the type II errors at the expense of worsening type II error.

Table 4 Type I & type II errors: Definition.[20]

	Type I Error	Type II Error
Trial data	Difference between interventions	No difference between interventions
In the background population	No difference between interventions	Difference exists between interventions

However none of the CVOT trials makes calculation on the pre-study sample population based on individual secondary end points.

Secondary end points have ingrained type I error associated with them. One way to circumvent this problem is to correct the p value for the number of secondary variables included in the trial. This can be done using several methods (Bonferroni correction, False Discovery Rate etc.).[21] However once you correct for type I error there is automatic worsening of type II error.[21]

It is practically impossible to come to a definitive conclusion from secondary end-points and sub-group analysis. Hence they are at best utilized for hypothesis generation and tested specifically at a subsequent study as a primary end point (Table 5).

Table 5 End-points significance

Primary end points	Hypothesis generating end points (subgroup analysis, secondary end points, post-hoc analysis etc.)
Pre-adjudicated	Not pre-adjudicated (most of the times)
Pre-specified	Not pre-adjudicated (most of the times)
Adjusted for type I & II errors	Not adjusted for type I & II errors

Table 6 Types of alpha spending.[25]

	Pocock	O'brien-fleming
Alpha spent per interim analysis using 50% of the total CV events	0.0155	0.0015

Examples:

(1). Is Saxagliptin associated with heart failure? Can we derive at a definitive conclusion from the SAVOR-TIMI 53 trial? The original article documented a significant hospitalization due to heart failure rate with saxagliptin (hazard ratio=1.27, 95% Confidence Interval 1.07–1.51, p=0.007).[22]

However this end point was part of several end points and a Bonferroni correction (p value of significance/number of secondary end points) was performed keeping into consideration 6 secondary end points, which were to be assessed. Hence the new p-value of significance was 0.05/6=0.0083. The original p-value (0.007) was still below the new Bonferroni corrected p-value and hence this signal remains significant.[23]

However after the study was published we saw that there were 10 secondary end points under consideration. Hence the true p-value of significance was 0.05/10=0.005. Hence the null hypothesis remains intact and the heart failure signal disappears.[23] Even if we make such a correction we cannot come to a definitive conclusion since correcting for type I error would have worsened type II errors. FDA in their analysis took the reverse route. They included hospitalization due to heart failure or all cause death as a composite end point to analyze the heart failure signal.[24] It was akin to correcting for the type II errors by increasing the numbers but at the cost of increasing the type I errors.

Either way the analysis does not give us a definite answer to the

question. Hence the only way out would be to conduct a separate trial whereby hospitalization due to heart failure is made a part of the primary composite.

Example: Does semaglutide increase retinopathy rates?

The calculations were made from a host of secondary end points and on 50 patients in the semaglutide & 29 in the standard care arm.[9] The p-value of significance was 0.02 (HR: 1.76; 95% CI, 1.11 to 2.78).[9] There were two groups of pre-specified and pre-adjudicated secondary composite end-points and two individual nephropathy & retinopathy secondary end points taking the count to four. If we make a Bonferroni correction here the new p value of significance is 0.05/4=0.0125. Hence the null hypothesis remains intact and it is no longer a significant adverse effect. However as mentioned above we cannot arrive at a definitive conclusion unless a dedicated trial is performed utilizing retinopathy complications as a primary end point. Other examples: Stroke signal with empagliflozin, bone fracture with canagliflozin, and renal-protection with empagliflozin.

Spending the alpha at interim analysis

This was a term we got introduced to after EMPA-REG trial was published. We have seen how the designs of CVOT were modified as time went by. One of the major reasons being the enormous cost associated with conducting such trials.

A pragmatic way to go forwards were to design a single trial to answer multiple questions especially those related to CV outcome. The trial should be pre-adjudicated to multiple interim analyses. The advantages associated with such an approach includes the following:

i. To satisfy the <1.8 upper bound CI bar and submit the drug for marketing approval.

ii. To decide on the futility of continuing with the trial if the predicted power for superiority at that time was <20%, the study can be stopped saving lots of fund.[25]

iii. If however the non-inferiority margin is met and there is a ≥20% power to achieve superiority the trial is allowed to continue and accumulate 611 CV events then it is cost effective to go ahead.

There is however a price to pay for opening up the trial multiple times for interim analysis (type I error). To overcome this, the alpha spending function must be invoked to control for the type I error induced (Table 6).[25] Of the three types of alpha spending methods Pocock approach is used when the intention is to obtain a closure at non-inferiority. This is because Pocock method spends a lot of alpha (0.0155) shifting the hazard ratio towards more stringent cut-off targets.[25] However when the study design is aimed at performing multiple interim analyses with both non-inferiority and subsequent superiority targets it is advisable to spend lesser alpha (O'Brien-Fleming or Haybittle-Peto).

Example: The CANVAS Trial. Plan is to perform three interim analyses.

There were 18,000 participants divided into two cohorts (Cohort A: 4,000 participants & Cohort B 14,000 participants). Cohort B would not be involved until an interim meta-analysis with MACE+ was done with Cohort a (contributing 80% of the events). This interim data satisfied the <1.8 upper bound CI bar (95% CI: 0.91-1.22)[26] and also had the power of >20% to look into superiority. Hence the Cohort B was planned to get activated and the final results are awaited. However in view of the multiple interim analyses alpha had to be spent to adjust for type I errors. Judiciously the O'Brien-Fleming

method was chosen to avoid spending too much of the total available alpha (0.025).[27]

And finally the robustness of a study outcome

It is a common question whether the end outcome(s) tested in a trial (positive, negative, neutral) is replicable or a result by chance. There are several ways to answer this question:

i. Repeat the same trial

ii. Design the trial with two different dosages of the same drug resulting in an in-built replication process.

iii. Persuasive p-value.

iv. Alternative analysis of the same data.

v. Sensitivity analysis.

Repeating the same study all over again is practically impossible from a logistic and monetary point of view. Waiting for similar data from another drug belonging to the same group is also riddled with problems. The study population, designs are different and we have seen how different molecules from the same drug class can behave differently from a CV outcomes perspective (ELIXA Vs. LEADER).[28] Hence the answer lies with the remaining methods of evaluation.

Example: The EMPA-REG Trial.

The EMPA-REG trial was performed with two different dosage of the same drug (10mg and 25mg). While involving two doses of the same drug in comparison to standard of care 137 events would be required to satisfy the <1.8 upper bound CI with 88% power for one dose and 687 CV events to discharge <1.3 with 88% power for both the doses (Bonferrroni adjustment).[25] The upper bound CI for 10mg & 25mg empagliflozin doses were 1.01 & 1.02 respectively.[7] In view of the smaller number of events in the individual arm superiority could not be demonstrated individually.

According to USFDA a p value of <0.001 is considered as a persuasive p-value reflecting replication of data.[30] In EMPA-REG trial cardiovascular & all-cause mortality satisfied this criteria. EMPA-REG trial also underwent an alternate analysis with Bayesian statistical method using the same set of data and came out with an equally robust outcome.[29]

Sensitivity analysis on the other hand is pre-specified for certain outcomes in most of the trials and aims to investigate the robustness of certain outcomes. It is predominantly utilized on the primary end points.[30] In EMPA-REG trial, sensitivity analysis was performed with these primary end point components and they were consistent with the initial data results.[7]

Conclusion

With CVOT becoming the order of the day and the regulatory authorities making several changes to the methodological aspects associated with these trials it is essential to have a basic understanding on the same. The way to approach CVOT starts with understanding the study design, background prevalence of the event(s) in question, pre-specified and pre-adjudicated end points, primary and pre-specified secondary end points chosen, aim of the study (non-inferiority with/ without superiority), alpha spending (if applicable) and ability to replicate the data. There are further analytical issues pertaining to the softer secondary end points like correcting for heterogeneity, adjusting for multiple variables etc., which are beyond the scope of this review.

A very pertinent question, which comes up from most of these

CVOT is whether the data is applicable to a primary CV prevention cohort. This is an issue not covered by these study types nor the mechanisms behind benefit or harm. An event driven study takes into consideration the total number of events collected over a period of time and analyses it. For example the pooled phase 2 & 3 data along with a large number of interim data from an ongoing CVOT is a mixing of cohort (both primary & secondary CV). This is done to reduce the number of years required to attain the magical number to satisfy the <1.8, <1.3 or <1.0 upper bound CI bars. Another area of concern is whether these data gives us any idea about the disease free interval from the day of initiation of the drug with a positive CV outcomes data. Unfortunately it does not. In order to get a specific answer on the same we need to conduct trials taking into account the time to event endpoints. With the constant evolution of strategies related to designing of CVOT we might see several such changes occurring in the future.

Acknowledgements

Conflict of interest

Author declares that there is no conflict of interest.

References

1. Nissen SE, Wolsky K. Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes. *N Engl J Med.* 2007;356(24):2457–71.

2. UK Prospective Diabetes Study Group: Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet.* 1998;352(9131):837–853.

3. Gerstein HC, Miller ME, Byington RP, et al. Effects of Intensive Glucose Lowering in Type 2 Diabetes The Action to Control Cardiovascular Risk in Diabetes Study Group. *N Engl J Med.* 2008;358(24):2545–2559.

4. Intensive Blood Glucose Control and Vascular Outcomes in Patients with Type 2 Diabetes The ADVANCE Collaborative Group. *N Engl J Med.* 2008;358(24):2560–72.

5. Duckworth W, Abraira C, Moritz T, et al. Glucose Control and Vascular Complications in Veterans with Type 2 Diabetes. *N Engl J Med.* 2009;360(2):129–139.

6. Hirshberg B, Raz I. Impact of the U.S. Food and Drug Administration Cardiovascular Assessment Requirements on the Development of Novel Antidiabetes Drugs. *Diabetes Care.* 2011;34(Suppl 2):S101–S106.

7. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *N Engl J Med.* 2015;373(22):2117–28.

8. Marso SP, Daniels GH, Brown-Frandsen K, et al. (Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med.* 2016;375(4):311–322.

9. Marso SP, Bain SC, Consoli A, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *N Engl J Med.* 2016;375(19):1834–1844.

10. Holman RR, Paul SK, Bethel MA, et al. 10Year Follow-up of Intensive Glucose Control in Type 2 Diabetes. *N Engl J Med.* 2008;359(15):1577–1589.

11. FDA Briefing Materials-Saxagliptin. 2009.

12. OE Johansen, D Neubacher, M von Eynatten, et al. Cardiovascular safety with linagliptin in patients with type 2 diabetes mellitus: a pre-specified, prospective, and adjudicated meta-analysis of a phase 3 programme. *Cardiovasc Diabetol.* 2012;11:3 .

13. Scirica BM, Bhatt DL, Braunwald E, et al. The design and rationale of the Saxagliptin Assessment of Vascular Outcomes Recorded in patients with diabetes mellitus–Thrombolysis in Myocardial Infarction (SAVOR-TIMI) 53 Study. *Am Heart J.* 2011;162(5):818–825.

14. White WB, Bakris GL, Bergenstal RM, et al. EXamination of cAardiovascular OutcoMes with AlogliptIN versus Standard of CarE in Patients with Type 2 Diabetes Mellitus and Acute Coronary Syndrome (EXAMINE): A cardiovascular safety study of the dipeptidyl peptidase 4 inhibitor alogliptin in patients with type 2 diabetes with acute coronary syndrome. *Am Heart J.* 2011;162(4):620–626.

15. Green JB, Bethel MA, Paul SK, et al. Rationale, design, and organization of a randomized, controlled Trial Evaluating Cardiovascular Outcomes with Sitagliptin (TECOS) in patients with type 2 diabetes and established cardiovascular disease. *Am Heart J.* 2013;166(6):983–989.

16. Neal B, Perkovic V, de Zeeuw D, et al. Rationale, design, and baseline characteristics of the Canagliflozin Cardiovascular Assessment Study (CANVAS)--A randomized placebo-controlled trial. *Am Heart J.* 2013;166(2):217–223.

17. Schweizer A, Dejager S, Foley JE, et al. Assessing the cardio–cerebrovascular safety of vildagliptin: meta-analysis of adjudicated events from a large Phase III type 2 diabetes population. *Diabetes Obes Metab.* 2010;12(6):485–494.

18. White WB, Cannon CP, Heller SR, et al. Alogliptin after Acute Coronary Syndrome in Patients with Type 2 Diabetes. *N Engl J Med.* 2013;369(14):1327–1335.

19. Schnell O, Standl E, Catrinoiu D, et al. Report from the 1st Cardiovascular Outcome Trial (CVOT) Summit of the Diabetes & Cardiovascular Disease (D&CVD) EASD Study Group. *Cardiovasc Diabetol.* 2016;15:33.

20. Harris M, Taylor G, Jackson D. *Clinical Evidence Made Easy.* 1st ed. © *Scion Publishing Limited*; 2014.

21. Why is multiple testing a problem? 2008.

22. Scirica BM, Bhatt DL, Braunwald E, et al. Saxagliptin and Cardiovascular Outcomes in Patients with Type 2 Diabetes Mellitus. *N Engl J Med.* 2013;369(14):1317–1326.

23. Cebrián-Cuenca AM, Orozco-Beltrán D, Navarro-Pérez J, et al. Saxagliptin and risk of heart failure hospitalization: Concern or miscalculation?. *Int J Cardiol.* 2016;220:573–574.

24. Briefing Material, NDA 22350: Saxagliptin (Onglyza). NDA 200678: Saxagliptin/Metformin (Kombiglyze XR). 2015.

25. 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. *Therapeutic Innovation & Regulatory Science.* 2015;49(1):50–64.

26. Janssen Research & Development LLC. Endocrinology and Metabolic Drug Advisory Committee.

27. CANVAS-Canagliflozin Cardiovascular Assessment Study.

28. Pfeffer MA, Claggett B, Diaz R, et al. Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome. *N Engl J Med.* 2015;373(23):2247–2257.

29. Kaul S. Is the Mortality Benefit With Empagliflozin in Type 2 Diabetes Mellitus Too Good To Be True? *Circulation.* 2016;134(2):94–96.

30. Thabane L, Mbuagbaw L, Zhang S, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol.* 2013;13:92.