



abstract **Objective:** To critically appraise the EMPA-REG OUTCOME trial and identify the appropriate place of the antidiabetic drug empagliflozin in therapeutics. **Methods:** A review was performed of the data provided in the main journal publication, supplementary appendix and different versions of the protocol, all published in The New England Journal of Medicine, as well as other public reports issued by major regulatory agencies. The manufacturer and the European Medicines Agency did not provide the Clinical Study Report. **Results:** A number of factors related to the development of the trial should be carefully considered before accepting the theoretical superiority of empagliflozin claimed by the investigators of the trial. More specifically, the substantial late amendments to the protocol and the statistical analysis plan, the breach of confidentiality of interim data, and the relevant conflicts of interest disclosed question the validity of the trial. In addition, the trial over-relies on merely exploratory data, the clinical relevance of the results is overestimated and its effectiveness in the European and American populations was very poor, which raises questions about the applicability of the results to our population. **Conclusions:** There is no solid evidence supporting the claims of superiority of empagliflozin over placebo for the prevention of macrovascular complications in diabetic patients with cardiovascular disease. The limitations of the EMPA-REG OUTCOME trial should be appropriately addressed before the drug's indications are extended. Also, a confirmatory trial should be conducted to verify the results of the first study.

The EMPA-REG OUTCOME trial (empagliflozin). A critical appraisal

The power of truth, the truth of power

LUIS CARLOS SAIZ
Pharmacist, Pharmacotherapy Research Coordinator
Navarre Health Service, Spain



OPEN ACCESS

«The power of true.»

ESLOGAN DE FLEISHMAN-HILLARD GROUP

«The truth is the truth, whether it is told by Agamemnon or his swineherd.

Agamemnon:

—I agree.

The swineherd:

—I don't completely agree.»

ANTONIO MACHADO / JUAN DE MAIRENA

Introduction

Diabetes Mellitus is characterized by the inability of the pancreas to produce insulin (type 1) or the inability of the body to respond adequately to the action of insulin (type 2). Combined with other factors, poor glycemic control may cause microvascular (at renal, neurological and retinal level) and macrovascular complications (primarily stroke, myocardial infarction and peripheral vascular disease) in the long-term. Diabetes Mellitus is a growing health problem that has become the focus of intense research.

Over 90% of diabetic patients have type 2 diabetes.¹ When exercise and a proper diet are not enough for achieving a good glycemic control, the administration of oral anti-diabetic drugs (OAD) is considered. Metformin is the first-choice drug for most patients, since it was conclusively proven to reduce morbidity and mortality in the UKPDS trial. Also, its profile of adverse events is well-known and acceptable in general terms.²

In the last decade, new families of antidiabetic drugs have been incorporated into the therapeutic arsenal for diabetes,³ such as GLP-1 agonists, dipeptidyl peptidase 4 (DPP-4) inhibitors and sodium-glucose cotransporter (type 2 SGLT) inhibitors. Empagliflozin -a type 2 SGLT inhibitor- was approved in Spain in June 2014 following dapagliflozin (November 2012) and canagliflozin (May 2014). These new drugs contribute to improving glycemic control by reducing the resorption of glucose by the kidneys.⁴

Effectiveness and use of empagliflozin

According to the World Health Organization,⁵ diabetes affects over 500 million people worldwide, which represents an 8.5% prevalence in the adult population and 1.5 million deaths directly attributable to diabetes every year. In Navarre, Spain, a total of 36,000 patients with diabetes have been identified through the Integrated Care Strategy for Chronic and Pluripathological Patients.⁶ Over half of them received at least an OAD by April 2016. Although SGLT2 inhibitors are a newly developed class of OADs (it is only prescribed to 3% of patients), there is a growing tendency to prescribe these drugs.

So far, the effectiveness of empagliflozin has been assessed by relying merely on changes in glycated hemoglobin (HbA1c) level. In the recent years, ten trials of a duration ranging from 24 to 78 weeks have been conducted with empagliflozin administered either alone or in combination with other drugs. In total, these trials included 11,250 patients and reported a mean 0.38% to 0.85% reduction in HbA1c levels with respect to placebo.⁷ Comparatively with other OADs, the only study published so far concluded that empagliflozin is not inferior to glimepiride in this intermediate outcome.⁷ Considering the above, the modest clinical results and data summarized in this paper do not support that empagliflozin provides an advantage over other older, more well studied, and cheaper OADs.

Rationale for the EMPA-REG OUTCOME trial

Ten years after the publication of the UKPDS 33 and 34 trials^{8,9}—which supported tight control of glycemia to reduce microvascular events and the use of metformin as first-choice drug for diabetes— an extensive therapeutic arsenal of glucose-lowering drugs is available nowadays. However, cases of cardiovascular events associated with the use of some of these molecules (such as tolbutamide) were reported, thus raising concerns about the cardiovascular safety of these agents. Finally, robust evidence was published in 2007 supporting the relationship between rosiglitazone and a higher risk of myocardial infarction and cardiovascular mortality¹⁰. Such warning resulted in the withdrawal of rosiglitazone from the market. This caused concern among physicians and regulatory authorities, since a drug for preventing cardiovascular events actually increased the risk of heart disease.

The United States Food and Drug Administration (FDA) published new guidelines for the development of clinical trials on diabetes in 2008.¹¹ Some months later, a supplementary report specifically focused on the assessment of the cardiovascular safety of antidiabetic drugs was released.¹² These reports confirm that there is more solid evidence on the relationship between good control of HbA1c and a lower incidence of microvascular events than there is regarding macrovascular events.^{11,13}

Henceforth, for a new antidiabetic drug to receive initial approval, a pivotal trial—or meta-analysis of phase II and III trials—comparing the incidence of major cardiovascular events in the group receiving the new drug vs controls is required. In this case, the established upper limit of the 95% confidence interval estimate for the relative risk is < 1.8 . Once the drug is approved by regulators, a post-authorization study must be performed to prove that the upper limit of the 95% confidence interval actually is < 1.3 . In other words, antidiabetic drugs that increase major cardiovascular events by up to 80% with respect to controls could be initially approved. Afterwards, if they would confirm up to 30% increase with respect to controls will not be withdrawn from the market. This is a remarkably “generous” margin for new candidate antidiabetic drugs. In agreement with these guidelines, sitagliptin, saxagliptin and alogliptin have been reported not to increase cardiovascular risk, and liraglutide has even been claimed to reduce it.³

The EMPA-REG OUTCOME trial¹⁴ (hereinafter EMPA-REG) — registered in clinicaltrials.gov as NCT01131676 and technically labeled as BI-1245.25— had an “adaptive design” that, under FDA authorization, integrated these two phases into a single trial. In order to grant approval, the FDA considered the data of an interim analysis of the EMPA-REG trial along with a pooled analysis of data from previous studies.¹⁵ The trial would then continue until the scheduled completion date, where the final data obtained would eventually support (or not support) initial approval.

Therefore, this is a safety study which carried a primary goal of ruling out clinical damage rather than confirming clinical benefit. The objective of this article is to carry out a critical appraisal of the EMPA-REG trial and identify the decisive factors that may help determine the appropriate place of the antidiabetic drug empagliflozin in therapeutics.

Methods

This report is based on basic data from the EMPA-REG trial as published in *The New England Journal of Medicine* (NEJM) on 17 September 2015. The paper included a final results report, supplementary information and the different versions of the study protocol.¹⁴ The summary of the protocol published in *Cardiovascular Diabetology* was also considered for this report.¹⁵

A Clinical Study Report (CSR) is a comprehensive, un-abridged primary report of study conduct and results that sponsors are required to submit for thorough review by regulatory authorities.¹⁶ On 23 September 2015, we requested Boehringer Ingelheim (BI) (manufacturer of Jardiance® (empagliflozin)) via www.ClinicalStudy-DataRequest.com to provide the CSR and individual data obtained in the EMPA-REG trial for an independent reanalysis of data. A week later, the pharmaceutical company refused to supply the CSR and invited our team to

request the CSR later in November 2016. However, when we contacted again in November 2016 we were told that they were “not yet able to share further details on the study because access to clinical data and documents is normally only provided after regulatory review has been completed or after termination of the development program”. Twenty months after the completion of the study, only a 23 page clinical study synopsis is available at the corporate website.¹⁷

Simultaneously, in October 2015, we requested the European Medicines Agency (EMA) to provide all information they had on the EMPA-REG trial, with especial emphasis on the CSR. The EMA forwarded several documents related to the authorization of empagliflozin including a meta-analysis providing intermediate data from the 1245.25 trial that was used for approval of the drug. However, the EMA refused to provide the complete CSR of the study on the grounds that disclosure of the CSR would interfere with EMA's decision on an ongoing process of authorization (extension of therapeutic indications) and that an exception could not be made as it was not of particular public interest. We were invited to request the CSR once the process was concluded.

Description of the EMPA-REG OUTCOME trial

Main research question

In patients with type 2 diabetes, poor glycemic control and established cardiovascular disease, is empagliflozin as add-on to standard antidiabetic therapy safe in terms of cardiovascular health?¹⁸

Design

A randomized, multicenter, parallel-group, double-blind trial of an expected 6-8-year duration until a minimum of 691 confirmed primary outcome events are attained.

Location

590 sites in 42 countries. By descending order of number of patients recruited, the trial was performed in Europe, North America (plus Australia and New Zealand), Asia, Latin America and Africa.

Inclusion and exclusion criteria

The study included adult patients with type 2 diabetes, insufficient glycemic control (HbA_{1c} = 7-9% for drug-naive patients and 7-10% for subjects on baseline therapy) and established cardiovascular (CV) disease (acute myocardial infarction [AMI], stroke, unstable angina, peripheral vascular disease or coronary artery disease).

Prevalent reasons for exclusion were glomerular filtration rate <30 ml/min/1.73m², liver disease, uncontrolled hyperglycemia or recent cardiovascular event.

Intervention

A total of 7,028 patients were randomly allocated to three treatment groups, of whom 7,020 received at least one dose of study drug, namely: empagliflozin 10mg (n=2,345), empagliflozin 25mg (n=2,342) or placebo (n=2,333), once daily for all groups. Randomization was stratified by baseline HbA1c, body mass index, renal function and geographic origin.

Baseline antidiabetic therapy was continued unchanged for the first 12 weeks following randomization, except for cases of glycemia > 240 mg/dL. After this period, therapy could be adjusted to achieve desired glycemetic control and treat other CV risk factors at the investigator's discretion to achieve best standard of care according to local guidelines.

Outcomes

The study protocol established the following outcomes by decreasing order of relevance:

Primary endpoint

A composite outcome of time to first occurrence of CV death, non-fatal myocardial infarction (excluding silent AMI), or non-fatal stroke.

Key secondary outcome

A composite outcome that expands the primary composite outcome to include time to first occurrence of hospitalization for unstable angina.

Other secondary outcomes

Silent AMI, heart failure requiring hospitalization and other microvascular events. Diagnosis of silent AMI was exclusively based on ECG findings.

Other exploratory outcomes

Overall mortality, CV mortality and other macrovascular or microvascular outcomes. Data on severe adverse effects were also collected, and potential renal, liver or genitourinary tract events, bone fractures, and thromboembolic episodes of hypoglycemia were monitored.

Statistical Analysis

Statistical analysis included a hierarchical testing strategy. The main hypothesis of the study was that empagliflozin administered at a dose of 10mg or 25mg (although dose groups were analyzed jointly) was not inferior to placebo regarding the primary outcome. Non-inferiority was demonstrated if the upper limit of the adjusted 95% confidence

Both, the manufacturer and the EMA denied access to the Clinical Study Report of the trial, which delays the performance of an independent analysis of data

interval for the hazard ratio between empagliflozin and placebo was <1.3 . In other words, the non-inferiority of empagliflozin vs placebo was proven if the increase in the risk of occurrence of primary outcome events did not exceed 30%.

Once non-inferiority was demonstrated, hierarchical testing for superiority was performed with respect to the primary outcome and the key secondary outcome. On such purpose, the new and single requirement to demonstrate the superiority of empagliflozin was that the upper limit of the adjusted 95% confidence interval for the hazard ratio was <1.0 . This means that statistical—but not clinical—superiority was required to be demonstrated. The primary analysis selected was a modified intention-to-treat analysis, which included all patients who received at least one dose of the study drug.

Results

A total of 772 primary outcome events were reported to have occurred in over 7,020 treated patients after a median duration of treatment of 2.6 years and 3.1-year follow-up, much earlier than expected. Twenty-five percent of patients prematurely discontinued study medication. The main outcomes related to CV safety are summarized in Table 1.

Differences in HbA1c between empagliflozin and placebo peaked at 12 weeks (0.60%) and dropped at the end of the study (0.24%). According to the Figure 1 provided in the paper, minimum differences in HbA1c coincided with the greatest differences in major outcomes. Slight reductions were also achieved with empagliflozin in weight, waist circumference, systolic blood pressure and uric acid, whereas mild increases were observed in c-LDL and c-HDL.

Authors' conclusions

In the authors' words, this trial provides evidence supporting the long-term use of empagliflozin and offer "strong evidence" of a reduction of CV risk. Patients with type 2 diabetes and high CV risk who received empagliflozin as add-on therapy to standard-of-care drugs as compared to placebo showed a lower rate of occurrence of primary CV outcomes and a lower overall mortality.

Critical review of the EMPA-REG OUTCOME trial

Choice of an adaptive study design

The EMPA-REG trial is one of the first examples of a clinical trial with an adaptive design focused on a standard-of-care therapy in primary care. At present, the role of traditional clinical trials is a matter of controversy. A number of scholars and regulators have questioned the appropriateness of using clinical trials as the main source of scientific evidence^{19,20} in the light of their inefficient design and the fact that it is not easy to extrapolate the results obtained to real practice.

Unfortunately, the solution proposed eliminates the main strength of clinical trials i.e. their high reliability based on a high internal consistency. Instead, adaptive, flexible and enabling trial designs that allow rapid access to new drugs have been proposed. Yet, they are accompanied by a greater uncertainty about their effectiveness and safety.

Traditional clinical trials have a range of limitations that should be minimized. However, the use of drugs in early clinical development as advocated by the adaptive method has significant weaknesses that should be considered²¹, namely: a) If expedited approval mechanisms are used, pharmacovigilance problems may arise in the future; b) sometimes (not in the EMPA-REG case) the adaptive pathways model relies on surrogate endpoints excessively, which may compromise the assessment of drug effectiveness; c) a large number of mandated post-authorization studies are eventually not performed; d) early market access to innovative drugs does not ensure that they are more effective. In sum, the adaptive model does not genuinely promote true innovation.²² Therefore, accelerated approval schemes entail the risk of serving the interests of the industry without regard to patient safety.

Conflicts of interests in the trial

The EMPA-REG trial was co-financed by BI, which is the manufacturer of empagliflozin, and Lilly. Of the 12 listed authors, seven are employees of BI and five are affiliated with several universities. The latter reported receiving consulting fees from BI¹⁵ and were paid for serving on the Steering Committee of the EMPA-REG trial.¹⁴

The author that chaired the Data Monitoring Committee—a supposedly independent body with regular access to data produced by the trial²³—received consulting fees from BI for the amount of \$40,000 between 2013 and 2015 while the EMPA-REG trial was still in progress.²⁴

Table 1. Main results of the EMPA-REG OUTCOME trial.

	No patients (%)		Hazard ratio (IC 95%)	ARR (%) (IC 95%)	NNT (IC 95%)
	Placebo (N=2333)	Empagliflozin (N=4687)			
PRIMARY OUTCOME					
Cardiovascular mortality + Non-fatal AMI + Non-fatal stroke	282 (12.1)	490 (10.5)	0.86 (0.74-0.99)	1.63 (0.05 to 3.22)	61 (31 to 2152)
KEY SECONDARY OUTCOME					
Cardiovascular mortality + Non-fatal AMI + Non-fatal stroke + Hospitalization for unstable angina	333 (14.3)	599 (12.8)	0.89 (0.78-1.01)	1.49 (-0.22 to 3.20)	67 (31 to -459)
ADDITIONAL SECONDARY OUTCOMES					
Silent AMI	15 (0.6)	38 (0.8)	1.28 (0.70-2.33)	-0.17 (-0.58 to 0.25)	-596 (406 to -171)
Hospitalization for heart failure	95 (4.1)	126 (2.7)	0.65 (0.50-0.85)	1.38 (0.46 to 2.31)	72 (43 to 219)
EXPLORATORY OUTCOMES					
All-cause mortality	194 (8.3)	269 (5.7)	0.68 (0.57-0.82)	2.58 (1.27 to 3.88)	39 (26 to 79)
Cardiovascular mortality	137 (5.9)	172 (3.7)	0.62 (0.49-0.77)	2.20 (1.11 to 3.30)	45 (30 to 90)
Fatal and non-fatal AMI (excluding silent AMI)	121 (5.2)	213 (4.5)	0.87 (0.70-1.09)	0.64 (-0.44 to 1.72)	156 (58 to -228)
Fatal and non-fatal stroke	69 (3.0)	164 (3.5)	1.18 (0.89-1.56)	-0.54 (-1.41 to 0.32)	-185 (308 to -71)

AMI = Acute myocardial infarction. ARR = Absolute risk reduction. NNT = Number needed to treat. NS = Non-significant.

Finally, the study was published in *The New England Journal of Medicine*, the journal with the greatest impact factor within the field of general medicine.²⁵ The drafting of the manuscript was carried out with the assistance of a communication and marketing company—Fleishman-Hillard Group—²⁶, which was directly hired by BI. Some authors have raised the alarm about the pharmaceutical industry's approach to legitimize ghostwriting by changing its definition, wherein ghostwriters are not to be considered ghostwriters so long as their name appears in the fine print of the manuscript, such as in the acknowledgements.²⁷

Interim data and confidentiality

FDA approval of empagliflozin required previous performance of a meta-analysis that incorporates all evidence available on its cardiovascular safety, which was conducted by the manufacturer. The meta-analysis included several studies, although most CV event records (74%) were collected from an interim analysis performed while the EMPA-REG trial was still in progress.

The validity of interim analyses lies on a strict maintenance of confidentiality so that interim data disclosure do not affect the progress of the trial.²⁸ However, strict maintenance of confidentiality does not seem to be a realistic goal when about 230 individuals—primarily BI employees—had access to confidential results sometime during the conduct of the trial.²⁹ Indeed, two Lilly employees committed a breach of confidentiality, as recognized by BI and reported in the corresponding FDA Statistical Review.³⁰ Although the date of breach is not provided, it occurred before the 4th version of the protocol was published (October 2013), whereby several substantial modifications were made (see below).

Nevertheless, the breach of confidentiality in the EMPA-REG trial is not the first case described in the literature. A review of the LIGHT trial—which assessed the CV safety of naltrexone-bupropion for the treatment of obesity—was recently published. The LIGHT trial was prematurely terminated due to a serious case of malpractice.³¹ In conclusion, the use of interim data from ongoing trials should be thoroughly questioned.³²

Amendments to the protocol

Completion of the EMPA-REG trial took five years, from May 2010 to the publication of the study in September 2015 (Table 2). The original protocol underwent continuous changes and five versions of the protocol are attached to the paper as supplementary material.³³

The original version and first amendment to the protocol were completed before patient recruitment started and, together with the second amendment as of April 2011, do not involve substantial changes.

The third amendment to the protocol—which was published more than a year after patient recruitment had started—establishes that silent AMI patients would be excluded from the primary composite outcome (page 113 of 165) and included in the outcome “non-fatal AMI” instead (page 118 of 165), showing a serious lack of coherence between the two outcomes. The exclusion of silent AMI patients from the primary outcome might have been due to the study design, which was intended to discard safety problems instead of demonstrating the efficacy of the drug.²⁹ This is of paramount importance, as statistical superiority would have disappeared if silent infarction had been included in the primary outcome.²⁹

The fourth and last amendment to the protocol was published in October 2013, three years after patient recruitment had started and after an interim analysis had been conducted due to the breach of confidentiality mentioned above. In the previous version of the protocol, the statistical analysis plan established that the trial would end when 691 events of the primary outcome were reached, which was enough to demonstrate the margin of non-inferiority requested (<1.3) with a statistical power of 90%.

However, the text in the new version was modified, and the number of events was raised to “a minimum of 691 events” (pages 162-163 of 165). The interim analysis showed that the result of the primary outcome for empagliflozin nearly reached—but did not reach—statistical superiority. [HR = 0.74 (95% CI 0.53-1.03)]. Finally, 772 events were evaluated and superiority in relation to the primary outcome very closely brushed the limit of statistical significance [HR = 0.86 (95% CI 0.74-0.99)]. In the light of the above, the hypothesis that by increasing the sample size a posteriori the investigators actually sought to reach statistical significance and demonstrate the superiority rather than the non-inferiority of empagliflozin in reducing macrovascular events—something unseen in antidiabetic drugs—seems plausible. This practice has been documented before.³⁴

Finally, if changes to the protocol were noticeable, amendments to the Statistical Analysis Plan were not less striking, given that it was last modified in May 2015, when the study had already been completed and was about to be published. Apart from including many of the amendments made to the protocol, the final version of the Statistical Plan also incorporated new outcomes post-hoc that were used later in other published papers on heart failure³⁵ or renal effects.³⁶

Table 2. Primary milestones of the EMPA-REG OUTCOME

2010 May	Original version of the protocol.
2010 September	First randomized subject.
2011 December	3rd amendment to the protocol: a) Silent infarctions are excluded from the primary outcome. b) Target: 691 events of the primary outcome.
2012 June	Deadline for interim analysis and meta-analysis.
2012 August	Final Statistical Analysis Plan.
2013 March	Application for approval of empagliflozin by FDA and EMA.
2013 April	Last randomized subject.
2013 October	FDA detects a breach of confidentiality of interim data.
	4th amendment to the protocol: a) Target of primary outcome events is extended to a minimum of 691. b) The authors open up the possibility of incorporating new outcomes.
2014 May	EMA approves empagliflozin.
2014 August	FDA approves empagliflozin.
2015 April	Last visit of the last active subject.
2015 May	Revision of the "final" Statistical Analysis Plan: a) New cardiovascular and renal outcomes are incorporated. b) The method of analysis of mortality outcomes is modified.
2015 September	The paper is published in <i>The New England Journal of Medicine</i> .

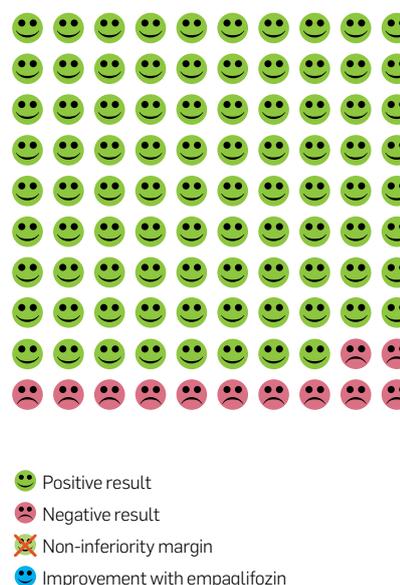
FDA: Food and Drug Administration. EMA: European Medicines Agency.

Clinical vs. Statistical Superiority

The main hypothesis was confirmed, since the reported data met the minimal clinically important difference, which means that empagliflozin (both doses combined) was non-inferior to placebo for the primary composite outcome. Yet, did data prove that empagliflozin was superior to placebo?

The EMPA-REG trial reports the statistical superiority of empagliflozin, which did not clearly translate into clinical benefit. The percentage of primary outcome events in the control group was 12% (Fig. 1). In turn, for non-inferiority to be demonstrated, a generous 30% relative margin (4% in absolute terms) with respect to the control group was established (Fig. 2). However, the 30% non-inferiority margin should also have been applied to the comparator (placebo).³⁷ Given that the difference in the incidence of primary outcome events was 2% in favor of empagliflozin (Fig. 3), the magnitude of the benefit was substantially lower than the non-inferiority margin established. Therefore, it cannot be concluded that the clinical superiority of empagliflozin was demonstrated.

Figure 1. Incidence of primary outcome events in the control group (in red).



From a different perspective, the NICE guideline for hypertension³⁸ establishes a minimal clinically important difference of 10% to 15% when interventions are compared in terms of mortality, incidence of stroke, AMI or composite primary outcome events. In contrast, in the EMPA-REG trial, estimated differences ranged from 1% to 26% [0.86 (95% CI 0.74-0.99)], which cannot be considered conclusive. On the other hand, it is disappointing that the non-inferiority of empagliflozin vs. placebo—which was the main hypothesis of the study—is not mentioned in the abstract, but only its superiority.

Exploratory outcomes established as primary outcomes

All-cause and cardiovascular mortality—measured separately—were initially included as tertiary outcomes and subsequently re-categorized as “additional outcomes”. In any case, these outcomes are merely exploratory and were not assessed with the statistical adjustments necessary to test a hypothesis with multiple primary outcomes. Yet, this crucial fact is not mentioned in the NEJM paper, and the reduction of mortality is reported as a key finding by including it in the title, discussion and conclusions, which places it at the same level as the primary outcome, without any further clarification.¹⁴

This scarcely rigorous appraisal of data is also remarked in the latest report on empagliflozin issued by the FDA in response to a BI application for approval of the proposed additional indication to reduce the risk of cardiovascular death.²⁹ The FDA argues that mortality is not included in the main analysis strategy, 40% of deaths labeled as “CV death” had an uncertain origin, the mechanism of action of the supposed benefit is unknown, and two confirmatory trials would be required prior to approval. Regardless of these objections, on the basis that sensitivity tests were positive for these outcomes, the FDA recommended the approval of the proposed additional indication.²⁹

Interestingly, the bid for a new indication for empagliflozin to reduce hospital stay for heart failure was rejected on the grounds that it is an exploratory outcome. However, CV death is also an exploratory outcome—with all its limitations—and the fact that it forms part of the primary composite outcome does not change its character.

Non-confirmatory sensitivity tests

The aim of a sensitivity test is to determine the robustness of a measurement tool by exploring to what extent changes in the methodology initially selected for the study affect the results. The different sensitivity tests conducted resulted in the loss of statistical superiority regarding the primary outcome, which indicates that the benefit obtained is unstable and weak. Additional sensitivity analyses were conducted by the FDA, with the same non-conclusive results (Table 3).

The lack of effectiveness observed in Europe and America seriously hinders the applicability of the study to our population

Figure 2. Excess of primary outcome events (crossed-out faces) allowed to demonstrate the non-inferiority of empagliflozin.

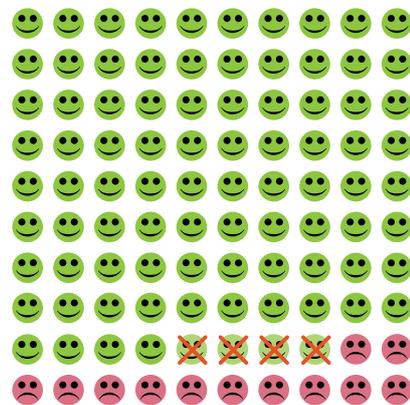
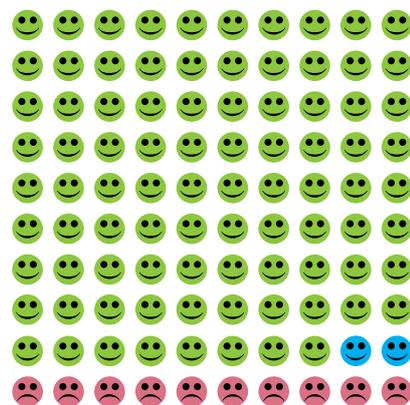


Figure 3. Percentage of prevented primary outcome events (in yellow) with primary endpoint.



- Positive result
- Negative result
- Non-inferiority margin
- Improvement with empagliflozin

Table 3. Results of the different sensitivity tests on the primary outcome of the EMPA-REG OUTCOME trial.

Principal Outcome (Cardiovascular death + Non-fatal AMI + Non-fatal stroke)	Hazard ratio (IC 95%)	Observations
Main analysis	0.86 (0.74-0.99)	Intention-to-treat analysis was selected.
Per protocol analysis	0.86 (0.75-1.00)	The most appropriate analysis in non-inferiority studies.
Analysis including silent AMI	0.92 (0.79-1.06)	Data from 4052 patients with available information.
Analysis excluding "non-assessable deaths"	0.90 (0.77-1.06)	"Non-assessable deaths" (27% of total) were labeled as cardiovascular deaths.
Analysis including events up to 30 days after the last dose	0.87 (0.74-1.02)	The paper misreports that the result is consistent with the main analysis.

AMI = Acute Myocardial Infarction. Data collected from references 11 and 26.

Differences between geographical subgroups

In clinical trials on cardiovascular disease, there is a tendency to include populations of different geographical origins, which has raised concerns on technical and ethical issues.³⁹

In the EMPA-REG trial, although a significant relationship was not observed between geographical regions and the primary outcome, the results obtained in Latin America and Asia (35% of the sample) were remarkably better as compared to those in Europe and North America. The German Health Technology Assessment Agency (IQWiG) noticed this difference⁴⁰ and pointed out that the supposed superiority of empagliflozin was determined by the benefit observed in Latin American and Asian study subjects, whereas no benefit was achieved in European subjects. In the same vein, the report issued by the FDA to respond to the request for a new indication of empagliflozin also remarks that empagliflozin did not provide any benefit to the North American subgroup.²⁹ However, the authors did not provide any explanation to the differences observed in clinical benefit between population subgroups and did not justify the generalization of results in their conclusions.

Randomization and masking

Randomization was performed using a validated electronic system that yields allocation numbers that are unpredictable and reproducible. Although this randomization method is commonplace in research and minimizes investigator bias, it is completely controlled by the study sponsor. Given that this is a very sensitive aspect, it is the regulatory agency that should perform randomization.

The method of masking is not described in the protocol of the trial. In addition, the protocol does not specify if the investigators had access to the blood and urine test results of their patients, which could have facilitated the identification of the treatment group assigned to them.

Mechanism of action "in search and seizure"

The publication of the EMPA-REG trial has caused speculations on the unexpected results obtained, which are especially surprising regarding stroke and non-fatal AMI. Some authors have theorized that these striking results may be caused by alterations in vascular reactivity and cardio-renal function,⁴¹ a relationship between albumin and uric acid,⁴¹ and even by an energy metabolism dysfunction.⁴²

The reduction observed in HbA1c associated with empagliflozin is too modest to be considered the most conclusive factor. It is also of note that poor baseline glycemic control did not improve homogeneously throughout the trial, as expected.^{29,40} Thus, persistent differences were observed in baseline glycemic control and other biochemical parameters (weight, systolic blood pressure, waist circumference, HDL-c) in favor of the groups receiving empagliflozin.¹⁴ This situation might have affected the results of the study, in favor of empagliflozin.

In accordance with the evidence provided in the literature so far, the underlying cause of the reduction in CV mortality might be the diuretic effect of empagliflozin in heart failure.⁴³ This hypothesis is supported by the positive result obtained for the secondary outcome "hospitalization for heart failure" and the fact that, of all CV death categories, "worsening of heart failure" was more frequent in the placebo group.

However, data are not consistent with this construct. On the one hand, empagliflozin did not have any effect on cardiogenic shock mortality, which is also linked to heart failure, whereas it did reduce the incidence of sudden death, which is not related to heart failure. But, even more importantly, the study design does not allow to assess the severity of baseline heart failure of subjects or the degree of adequacy of the treatment administered, which hinders the interpretation of results.²⁹ Nevertheless, this is the main hypothesis that the manufacturer is going to test shortly by two specific studies.⁴⁴

Adverse reactions... with class effect?

The overall balance of serious adverse events in the EMPA-REG study was favorable to empagliflozin basically due to the reduction in CV mortality.¹⁴ As to non-CV safety, no severe adverse reactions were reported, except for genital tract infections (6.4% for empagliflozin, 1.8% for placebo). Urinary infections —already reported in previous studies— were not found to be more frequent in the empagliflozin than in the control group.

As to the other two SGLT2 agents, warning alerts have been continuous. The most recent warning issued by the FDA was that canagliflozin and dapagliflozin can cause acute renal failure.⁴⁵ Regulators had previously reported cases of non-traumatic amputation of the lower limbs in subjects who had received canagliflozin,⁴⁶ which raised alerts on other SGLT2 inhibitors, and of ketoacidosis in all SGLT molecules.^{47,48} The EMA Risk Management Plan also includes collecting data on bone fractures.

Empagliflozin should be used with extreme caution even though the EMPA-REG trial only showed a higher frequency of renal adverse events associated with empagliflozin in the first three months of the study²⁹ to decrease later, and a very low rate of ketoacidosis without differences between groups.¹⁴ In fact, an increase of the hematocrit was observed in the empagliflozin group linked to the depletion of plasma volume,⁴³ a factor that might have clinical effects.

Other factors of interest

Apart from the aspects thoroughly addressed in this report, other factors should also be considered when assessing the contribution of empagliflozin in the management of diabetic patients, namely:

The statistical superiority of the primary composite variable was achieved by adding the populations of the two empagliflozin groups, who received different doses of the drug (10mg and 25 mg). When assessed separately, none of the empagliflozin groups reached superiority vs. placebo.¹⁴

The drop-out rate was very high (23.4% for empagliflozin, 29.2% for placebo) and 3% did not complete the study. Although vital status was assessed in 99% of patients, missing data on adverse events affected the primary composite variable.²⁹

Figures 1 and 2 of the paper published in NEJM deserve special attention. In Figure 1, the vertical axis (% of patients who suffered an event) does not show 100% of patients, which artificially widens the gap between treatment groups. Additionally, survival curves are remarkably divergent in the last months due to a disproportionately higher incidence of events in the placebo group that was not explained by the authors. In Figure 2, along with results by primary outcome subgroups (predefined study), results are also provided for the exploratory outcome "CV mortality" (*post-hoc* study). This may give a false impression of homogeneity, as substantially broader confidence intervals are used for this outcome.

Acknowledgements

We are grateful to Peter Doshi of the University of Maryland, Baltimore (USA) for reviewing the text.

Conclusions

Empagliflozin has been proven non-inferior to placebo in the primary outcome composite variable (CV mortality + non-fatal infarction + non-fatal stroke).

The superiority of empagliflozin claimed by the authors is questionable due to significant limitations of the trial, which include: relevant late amendments to the protocol and the statistical analysis plan, the breach of confidentiality of interim data, important conflicts of interest, excessive value given to merely exploratory outcomes, and the magnification of the clinical relevance based on a modest statistical superiority.

The potential association between empagliflozin and severe adverse reactions such as renal failure, ketoacidosis and amputations—which have already been reported for other similar drugs of the same family— should be closely monitored.

Empagliflozin has not been appropriately demonstrated to be effective in preventing macrovascular complications in patients with type 2 diabetes and established CV disease.

The limitations of the EMPA-REG OUTCOME trial should be appropriately addressed before its indications are extended. Also, a confirmatory trial should be conducted to verify the results of the first study.

Bibliografía

1. Kharroubi AT, Darwish HM. Diabetes mellitus: The epidemic of the century. *World J Diabetes* [Internet]. 2015 Jun 25 [cited 2016 Jun 19];6(6):850–67.
2. Lafita J, Áriz M. Tratamiento no insulínico de la diabetes tipo 2. Solución pendiente. *BIT Navarra*. 2009;17(2):13–28.
3. Ingelfinger JR, Rosen CJ. Cardiac and Renovascular Complications in Type 2 Diabetes. Is There Hope? *N Engl J Med*. 2016;375(4):380–2.
4. Jardiance: European Public Assessment Report (EPAR) [Internet]. 2014.
5. World Health Organization. Global Report on diabetes. Geneva; 2016.
6. Ibarrola C, Basterra G, Celorrio JJ, De la Fuente M, De la Rosa P, Escriche L, et al. Evaluación y Despliegue. Estrategia Navarra de atención integrada a pacientes crónicos y pluripatológicos. 2016.
7. Informe de Posicionamiento Terapéutico de empagliflozina (Jardiance) [Internet]. 2016.
8. UK Prospective Diabetes Study (UKPDS) 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 Sep 12;352(9131):837–53.
9. UK Prospective Diabetes Study (UKPDS) Group. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet*. 1998 Sep 12;352(9131):854–65.
10. Nissen SE, Wolksi K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Eng J Med* 2007;356:2457–71.
11. Guidance for Industry. Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention [Internet]. 2008.
12. Guidance for Industry Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes [Internet]. 2008.
13. Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical Update: Cardiovascular Disease in Diabetes Mellitus: Atherosclerotic Cardiovascular Disease and Heart Failure in Type 2 Diabetes Mellitus – Mechanisms, Management, and Clinical Considerations. *Circ* [Internet]. 2016;133(24):2459–502.
14. 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* [Internet]. 2015;373(22):2117–28.
15. Zinman B, Inzucchi SE, Lachin JM, Wanner C, Ferrari R, Fitchett D, et al. Rationale, design, and baseline characteristics of a randomized, placebo-controlled cardiovascular outcome trial of empagliflozin (EMPA-REG OUTCOMETM). *Cardiovasc Diabetol* [Internet]. 2014;13:102.
16. Doshi P, Jefferson T. Clinical study reports of randomised controlled trials: an exploratory review of previously confidential industry reports. *BMJ Open* 2013;3:e002496 doi:10.1136/bmjopen-2012-002496.
17. Clinical Study Synopsis for Public Disclosure [Internet]. Consultado el 08/12/2016.
18. BI 10773 (Empagliflozin) Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME). [Internet]. ClinicalTrials.gov. 2010 [cited 2016 Aug 17].

19. Bhatt DL, Mehta C. Adaptive Designs for Clinical Trials. *N Engl J Med*. 2016;375:65–74.
20. Eichler H-G, Oye K, Baird LG, Abadie E, Brown J, Drum CL, et al. Adaptive licensing: taking the next step in the evolution of drug approval. *Clin Pharmacol Ther* [Internet]. 2012;91(3):426–37.
21. Davis C, Lexchin J, Jefferson T, Gøtzsche P, McKee M. “Adaptive pathways” to drug authorisation: adapting to industry? *BMJ* [Internet]. 2016;4437(August):i4437.
22. A PRIME example of how EMA is pushing for accelerated market approvals, but at what cost for patients? [Internet]. 2015.
23. Marston L, McKenzie DR, Freemantle N. Pursuit of truly independent data monitoring committees in research. *BMJ* [Internet]. 2016;353:i2309.
24. U.S. Department of Health & Human Services. Open Payments [Internet]. [cited 2016 Jul 5].
25. The New England Journal of Medicine. Media Center Fact Sheet [Internet]. [cited 2016 Jun 28].
26. Fleishman-Hillard. The power of true [Internet]. [cited 2016 Jun 28].
27. Matheson A. Ghostwriting: the importance of definition and its place. *BMJ* [Internet]. 2016;4578(August):1–7.
28. Fleming TR. Protecting the Confidentiality of Interim Data: Addressing Current Challenges. *Clin Trials*. 2015;12(1):5–11.
29. FDA Briefing Document. Endocrine and Metabolic Drug Advisory Committee Meeting. June 28, 2016. [Internet]. 2016.
30. Charles JK. Empagliflozin: Statistical Review and Evaluation [Internet]. 2013.
31. Sharfstein JM, Psaty BM. Evaluation of the Cardiovascular Risk of Naltrexone-Bupropion A Study Interrupted. *JAMA J Am Med Assoc* [Internet]. 2016;315(10):984–6.
32. Naci H, Lehman R, Wouters OJ, Goldacre B, Yudkin JS. Rethinking the Appraisal and Approval of Drugs for Type 2 Diabetes. *Bmj* [Internet]. 2015;351(October):h5260.
33. EMPA-REG OUTCOME Investigators. Protocol for the EMPA-REG OUTCOME trial. *N Engl J Med*. 2016;297.
34. Erviti López J. Ensayo IMPROVE-IT (ezetimiba añadida a simvastatina en el síndrome coronario agudo). Una lectura crítica. *BIT Navarra* [Internet]. 2015;23(5):1–8.
35. Fitchett D, Zinman B, Wanner C, Lachin JM, Hantel S, Salsali A, et al. Heart failure outcomes with empagliflozin in patients with type 2 diabetes at high cardiovascular risk: results of the EMPA-REG OUTCOME® trial. *Eur Heart J* [Internet]. 2016;ehv728–.
36. Wanner C, Inzucchi SE, Lachin JM, Fitchett D, Eynatten M Von, Mattheus M, et al. Empagliflozin and Progression of Kidney Disease in Type 2 Diabetes. 2016;1–12.
37. DiNubile MJ, Sklar P, Lupinacci RJ, Eron Jr JJ. Paradoxical interpretations of noninferiority studies: violating the excluded middle. *Future Virol*. 2012;7(11):1055–63.
38. Clinical Guideline 127. Hypertension. The clinical management of primary hypertension in adults [Internet]. 2011.
39. Mentz RJ, Kaski JC, Dan GA, Goldstein S, Stockbridge N, Alonso-Garcia A, et al. Implications of geographical variation on clinical outcomes of cardiovascular trials. *Am Heart J*. 2012;164(3):303–12.
40. Empagliflozin – Benefit assessment according to § 35a Social Code Book V [Internet]. 2016.
41. Ingelfinger JR, Rosen CJ. Cardiovascular Risk and Sodium-Glucose Cotransporter 2 Inhibition in Type 2 Diabetes. *N Engl J Med*. 2015;373(22):2178–9.
42. Mudaliar S, Aljoju S, Henry RR. Can a Shift in Fuel Energetics Explain the Beneficial Cardiorenal Outcomes in the EMPA-REG OUTCOME Study? A Unifying Hypothesis. *Diabetes Care* [Internet]. 2016;(March):dc160542.
43. Empagliflozine: Diabète de type 2: pas d'emballement. *La Rev Prescrire*. 2016;36(389):168–73.
44. Jardiance (empagliflozin) to be studied for the treatment of people with chronic heart failure [Internet]. 2016 [cited 2016 Jul 11].
45. FDA Drug Safety Communication: FDA strengthens kidney warnings for diabetes medicines canagliflozin (Invokana, Invokamet) and dapagliflozin (Farxiga, Xigduo XR) [Internet]. 2016.
46. Canagliflozina: posible incremento del riesgo de amputación no traumática en miembros inferiores [Internet]. 2016.
47. Recomendaciones sobre el riesgo de cetoacidosis diabética asociada al uso de canagliflozina, dapagliflozina y empagliflozina [Internet]. 2016.
48. Farjo PD, Kidd KM, Reece JL. A Case of Euglycemic Diabetic Ketoacidosis Following Long-term Empagliflozin Therapy. *Diabetes Care* [Internet]. 2016;(April):dc160728.



**Servicio Navarro de Salud
Osasunbidea**

ISSN

1138-1043

COPYRIGHT

NA-1263/1997

INFORMATION AND SUSCRIPTION

Servicio Navarro de Salud / Osasunbidea

Plaza de la Paz, s/n

31002 Pamplona

T 848429047

F 848429010

E-mail

farmacia.atprimaria@cfnavarra.es

Web site

www.dtb.navarra.es

EDITORIAL BOARD

CHAIRMAN

Antonio López

MEMBERS

Cristina Agudo

M^a José Ariz

Miguel Ángel Imízcoz

Víctor Napal

Idoia Gaminde

M^a Mar Malón

Rodolfo Montoya

Luis Carlos Saiz

Javier Garjón

Iván Méndez

Gabriela Elizondo

EDITOR

Juan Erviti