Interpretación de los aspectos claves estadísticos y diferenciales de los Ensayos de Seguridad Cardiovascular (CVOTs)

Jose Ríos

Coord. Medical Statistics (IDIBAPS) Hosp. Clínic de Barcelona

Prof. Associado de la Fac. de Medicina de la UAB

Datos del Facilitador de la reunión

PP-DG-ES-0513

Objetivo

Conocer los aspectos generales y clave para la interpretación estadística de un CVOT o de su agrupación en metaanálisis

Ejemplos de los arGLP1

CVOTs, cardiovascular outcomes trials

Agenda

- **Distintos tipos de CVOT y diseños:**
  - **Diseños de no-inferioridad vs superioridad.**
  - **Cómo se establece el tamaño de la muestra y el margen de no-inferioridad**
- **Población estudiada en los CVOT**
- **Meta-análisis y sus posibles fuentes de heterogeneidad**

CVOTs, cardiovascular outcomes trials

• Preguntas

- **Diseño de los CVOT**
  - ¿Qué diferencias en el diseño de los CVOTs conoceis?
  - ¿Pensáis que implica diferente nivel de evidencia?
  - ¿Pensáis que afecta a la extrapolación de resultados a la práctica habitual?
  - ¿Tenemos en cuenta si el diseño es de superioridad o no a la hora de elegir un GLP u otro?

Nota importante

Statistics Notes

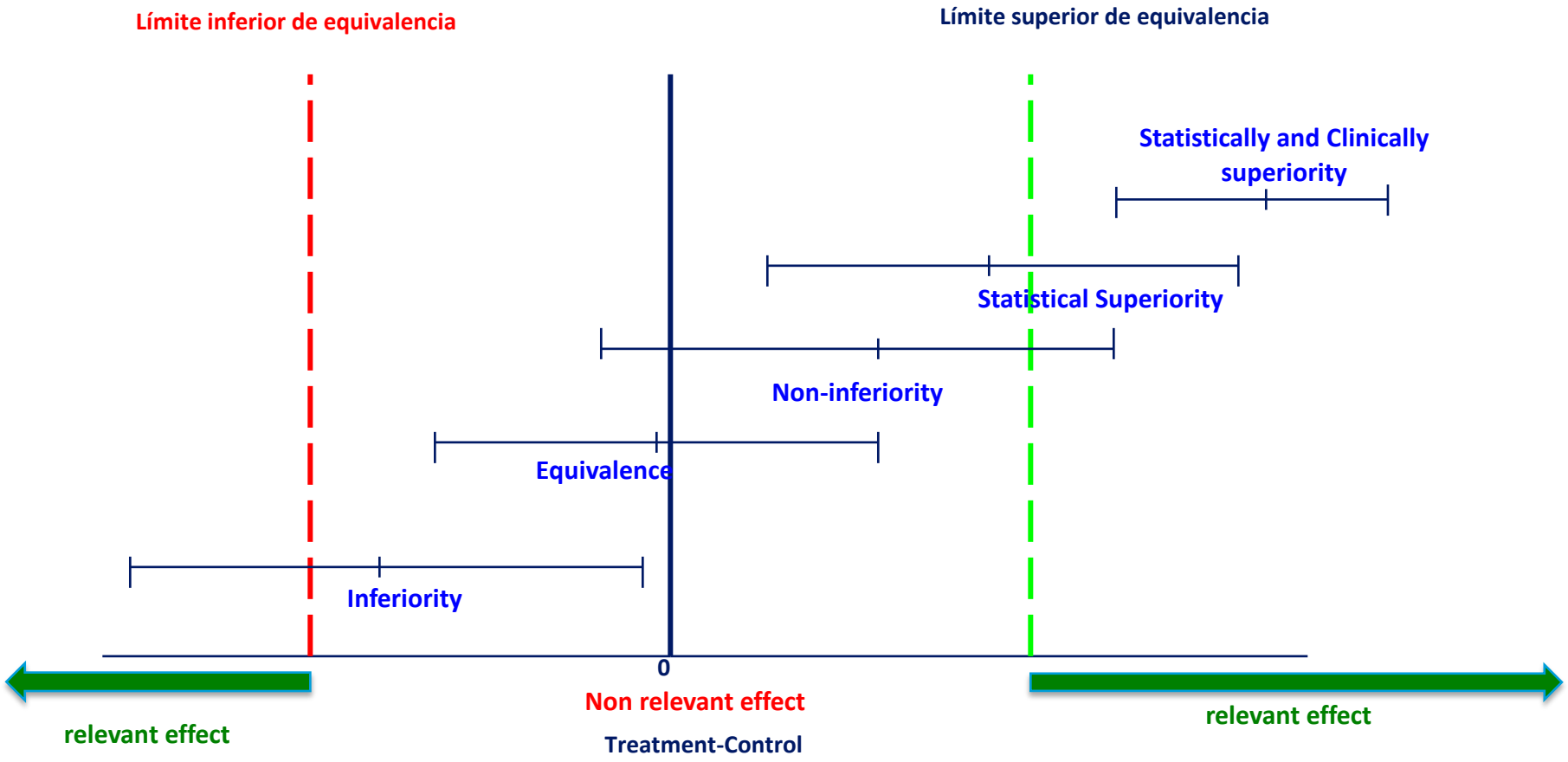
Absence of evidence is not evidence of absence

Douglas G Altman, J Martin Bland

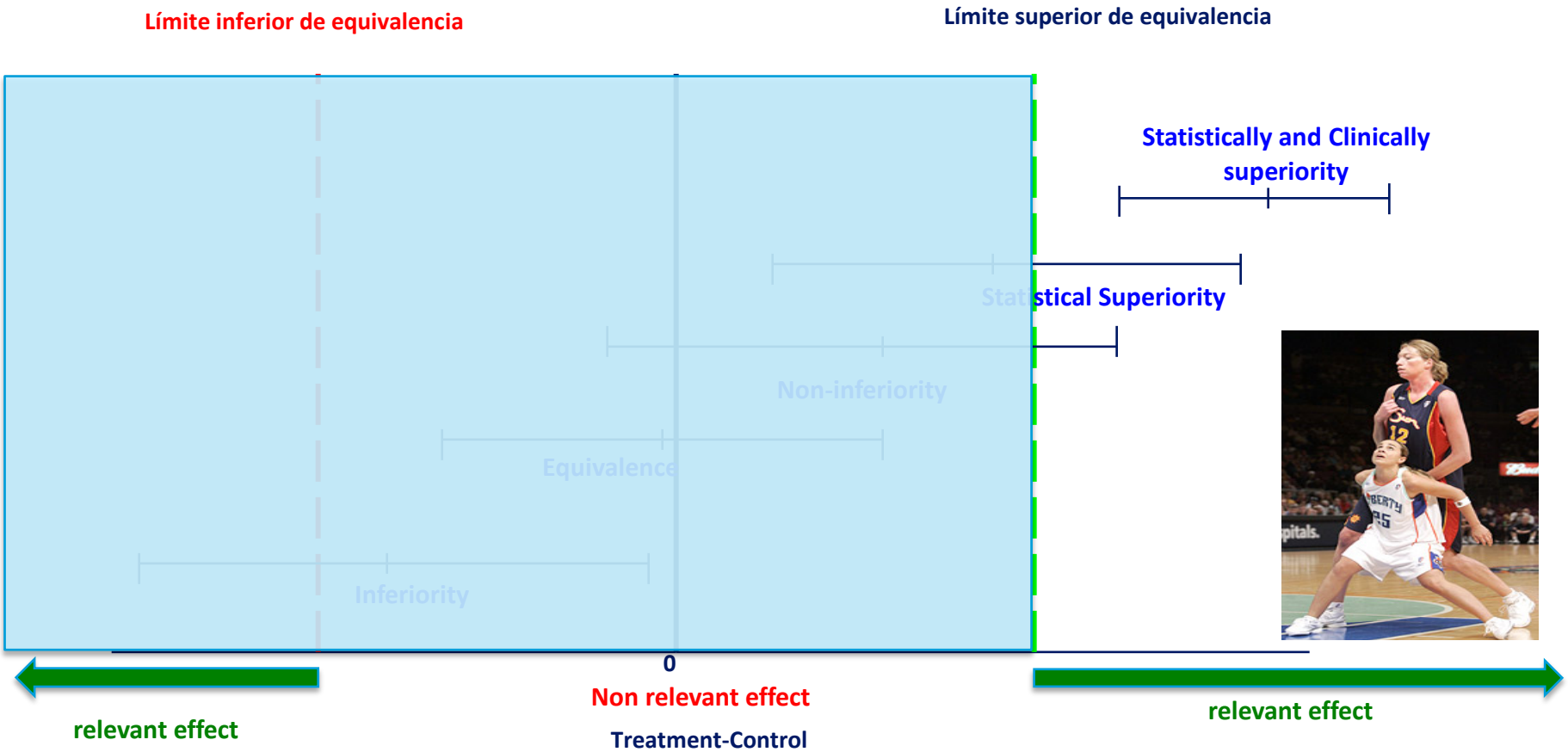
BMJ 1995;311:485

Consecuencias: Los métodos estadísticos no sirven para mostrar igualdad, ya que la igualdad entre dos cosas no existe.

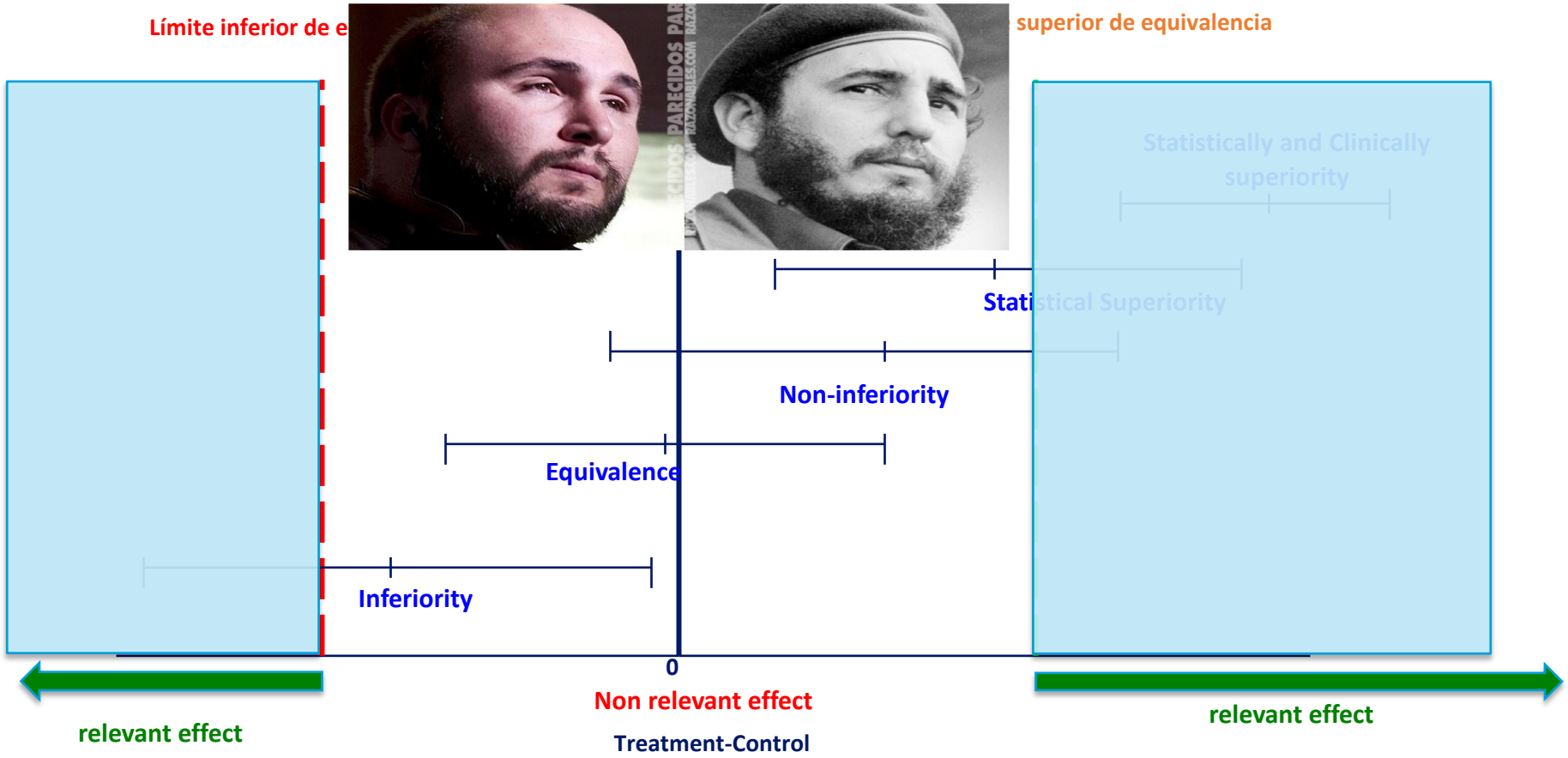
Equivalencia clínica y estadística



Equivalencia clínica y estadística



Equivalencia clínica y estadística



Equivalencia clínica y estadística

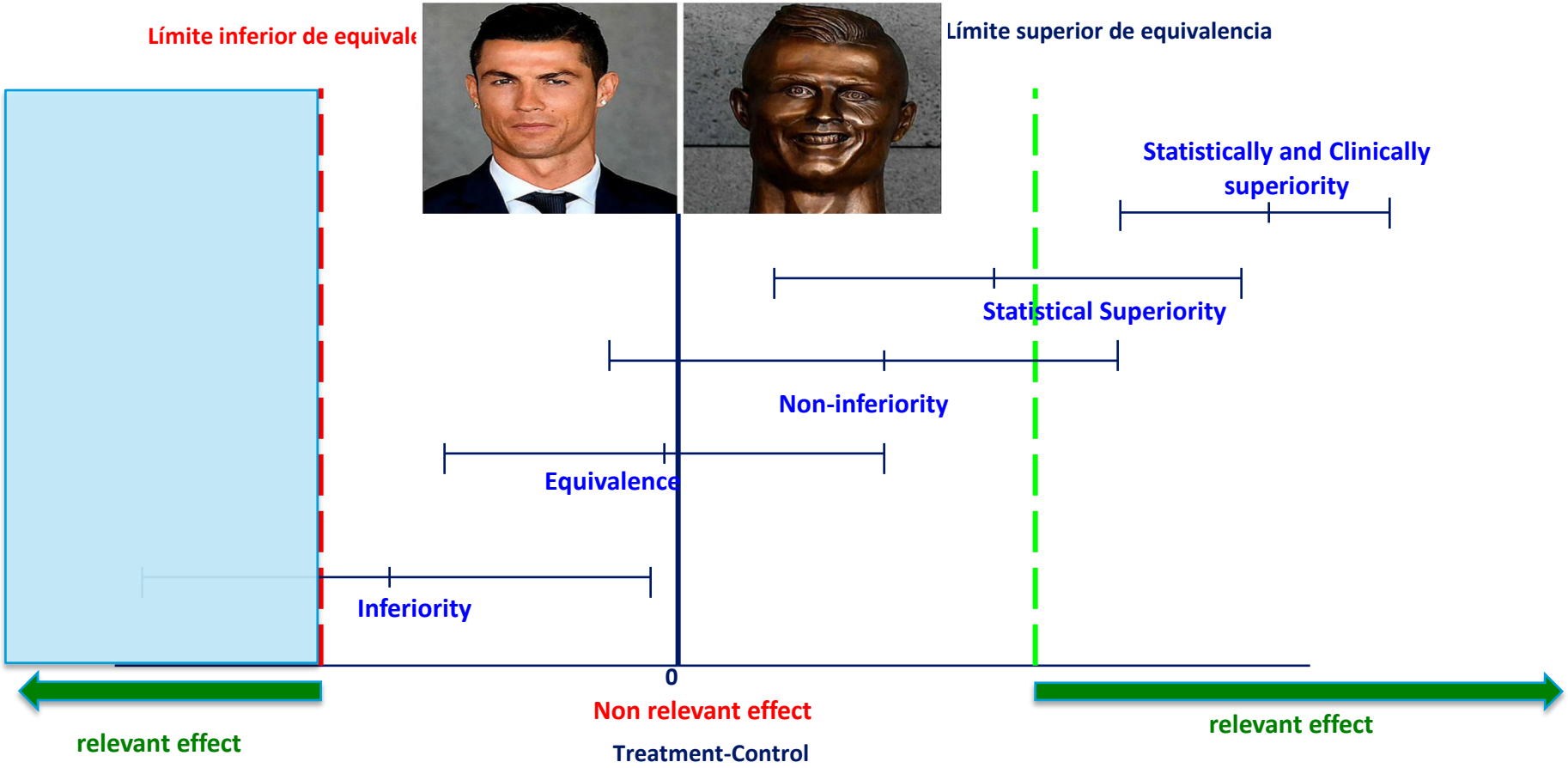
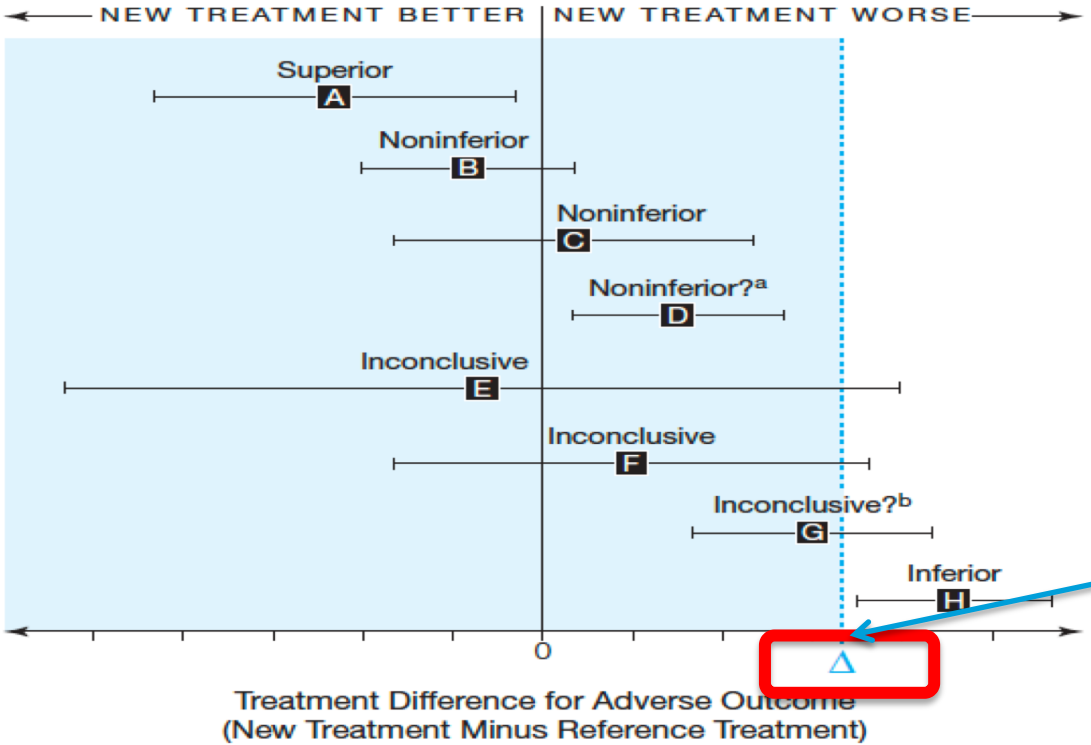


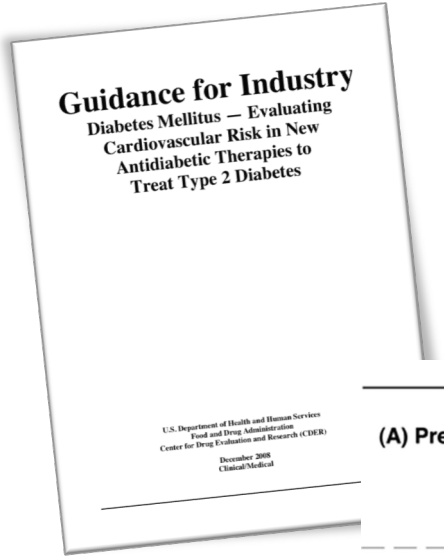
Figure 1. Possible Scenarios of Observed Treatment Differences for Adverse Outcomes (Harms) in Noninferiority Trials



En nuestros casos 'recientes', HR: 1.3

Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. **JAMA. 2012 Dec 26;308(24):2594-604**

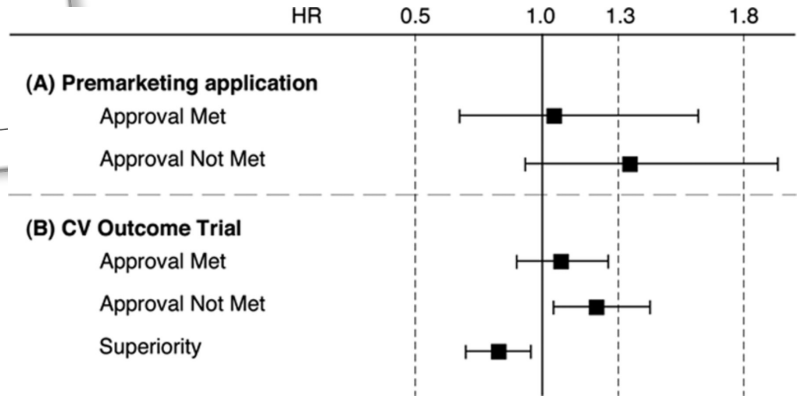
CVOTs: Opciones allá y que llegaron aquí

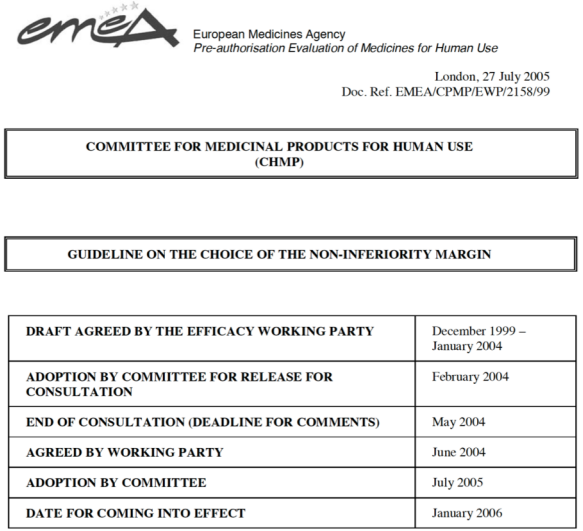

Guidance for Industry

Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

December 2008
Clinical/Medical




European Medicines Agency
Pre-authorisation Evaluation of Medicines for Human Use

London, 27 July 2005
Doc. Ref. EMEA/CPMP/EWP/2158/99

COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE
(CHMP)

GUIDELINE ON THE CHOICE OF THE NON-INFERIORITY MARGIN

DRAFT AGREED BY THE EFFICACY WORKING PARTY	December 1999 – January 2004
ADOPTION BY COMMITTEE FOR RELEASE FOR CONSULTATION	February 2004
END OF CONSULTATION (DEADLINE FOR COMMENTS)	May 2004
AGREED BY WORKING PARTY	June 2004
ADOPTION BY COMMITTEE	July 2005
DATE FOR COMING INTO EFFECT	January 2006

CVOTs, cardiovascular outcomes trials

FDA Guidance for Industry. February 2008. Available at: https://www.fda.gov/downloads/Drugs/.../Guidances/ucm071624.pdf
EMA Guideline available at https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-choice-non-inferiority-margin_en.pdf

CVOTs. El ejemplo de los arGLP-1. Post-aprobación: No inferioridad

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome

Marc A. Pfeffer, M.D., Ph.D., Brian Claggett, Ph.D., Rafael Diaz, M.D., Kenneth Dickstein, M.D., Ph.D., Hertzel C. Gerstein, M.D., Lars V. Køber, M.D., Francesca C. Lawson, M.D., Lin Ping, M.D., Xiaodan Wei, Ph.D., Eldrin F. Lewis, M.D., M.P.H., Aldo P. Maggioni, M.D., John J.V. McMurray, M.D., Ph.D., Jeffrey L. Probstfield, M.D., Matthew C. Riddle, M.D., Scott D. Solomon, M.D., and Jean-Claude Tardif, M.D., for the ELIXA Investigators*

ABSTRACT

BACKGROUND

Cardiovascular morbidity and mortality are higher among patients with type 2 diabetes, particularly those with concomitant cardiovascular diseases, than in most other populations. We assessed the effects of lixisenatide, a glucagon-like peptide 1–receptor agonist, on cardiovascular outcomes in patients with type 2 diabetes who had had a recent acute coronary event.

METHODS

We randomly assigned patients with type 2 diabetes who had had a myocardial infarction or who had been hospitalized for unstable angina within the previous 180 days to receive lixisenatide or placebo in addition to locally determined standards of care. The trial was designed with adequate statistical power to assess whether lixisenatide was noninferior as well as superior to placebo, as defined by an upper boundary of the 95% confidence interval for the hazard ratio of less than 1.3 and 1.0, respectively, for the primary composite end point of cardiovascular death, myocardial infarction, stroke, or hospitalization for unstable angina.

RESULTS

The 6068 patients who underwent randomization were followed for a median of 25 months. A primary end-point event occurred in 406 patients (13.4%) in the lixisenatide group and in 399 (13.2%) in the placebo group (hazard ratio, 1.02; 95% confidence interval [CI], 0.89 to 1.17), which showed the noninferiority of lixisenatide to placebo (P<0.001) but did not show superiority (P=0.81). There were no significant between-group differences in the rate of hospitalization for heart failure (hazard ratio in the lixisenatide group, 0.96; 95% CI, 0.75 to 1.23) or the rate of death (hazard ratio, 0.94; 95% CI, 0.78 to 1.13). Lixisenatide was not associated with a higher rate of serious adverse events or severe hypoglycemia, pancreatitis, pancreatic neoplasms, or allergic reactions than was placebo.

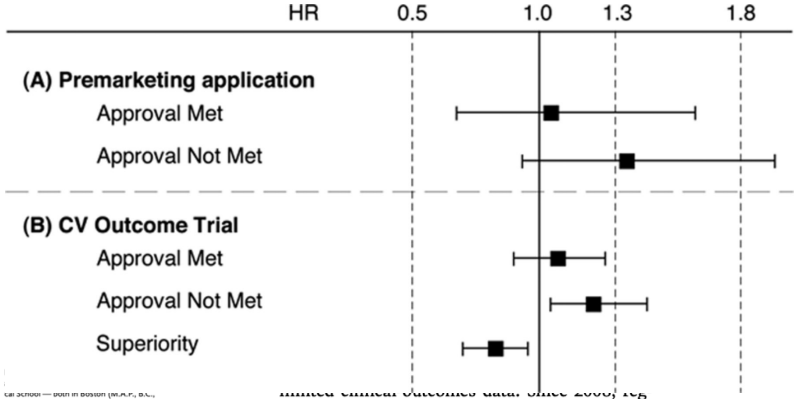
CONCLUSIONS

In patients with type 2 diabetes and a recent acute coronary syndrome, the addition of lixisenatide to usual care did not significantly alter the rate of major cardiovascular events or other serious adverse events. (Funded by Sanofi; ELIXA ClinicalTrials.gov number, NCT01147250.)

Cardiology, Brigham and Women's Hospital, Boston (M.A.P., B.C., E.F.L., S.D.S.); Estudios Clinicos Latinoamérica, Rosario, Argentina (R.D.); University of Bergen, Stavanger University Hospital, Stavanger, Norway (K.D.); McMaster University, Hamilton, ON, Canada (H.C.G.); Rigshospitalet Copenhagen University Hospital, Copenhagen (L.V.K.); Sanofi U.S., Bridgewater, NJ (F.C.L., L.P., X.W.); Research Center of the Italian Association of Hospital Cardiologists, Florence (A.P.M.); British Heart Foundation Cardiovascular Research Centre, University of Glasgow, Glasgow, United Kingdom (J.J.V.M.); University of Washington Medical Center, Seattle (J.L.P.); Oregon Health and Science University, Portland (M.C.R.); and Montreal Heart Institute, Université de Montréal, Montreal (J.C.T.). Address reprint requests to Dr. Pfeffer at the Cardiovascular Division, Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115, or at mpfeffer@rics.bwh.harvard.edu.

*A list of the investigators and committee members in the Evaluation of Lixisenatide in Acute Coronary Syndrome (ELIXA) trial is provided in the Supplementary Appendix, available at NEJM.org.

N Engl J Med 2015;373:2247-57.
DOI: 10.1056/NEJMoa1509225
Copyright © 2015 Massachusetts Medical Society.



limited clinical outcomes data. Since 2008, regulatory agencies have required robust cardiovascular-outcome data from randomized, controlled trials in order to grant and sustain approvals of drugs or biologic agents for the treatment or prevention of diabetes mellitus.[11-13]

Lixisenatide, a once-daily GLP-1–receptor agonist, is effective in reducing the glycated hemoglobin level in patients with type 2 diabetes by lowering both the fasting and the postprandial blood glucose levels.[22-25] However, the studies showing such glycemic control and other favorable metabolic effects of lixisenatide were not designed to accrue sufficient clinical events to adequately show its cardiovascular safety. Although lixisenatide and several other GLP-1–receptor agonists are approved in many countries for use as glucose-lowering agents in patients with type 2 diabetes, large cardiovascular-outcome trials with any agent in this class have been lacking.

STATISTICAL ANALYSIS

The primary analysis was conducted in the intention-to-treat population with the use of the Cox proportional-hazards model, with study group and geographic region as the covariates, to estimate the hazard ratio for the comparison of lixisenatide with placebo. The noninferiority of lixisenatide to placebo would be shown if the upper boundary of the 95% confidence interval of the hazard ratio was less than 1.3 and the superiority would be shown if the upper boundary was less than 1.0.[11] The P values for the tests of noninferiority and superiority are derived from the Cox proximal-hazards model.

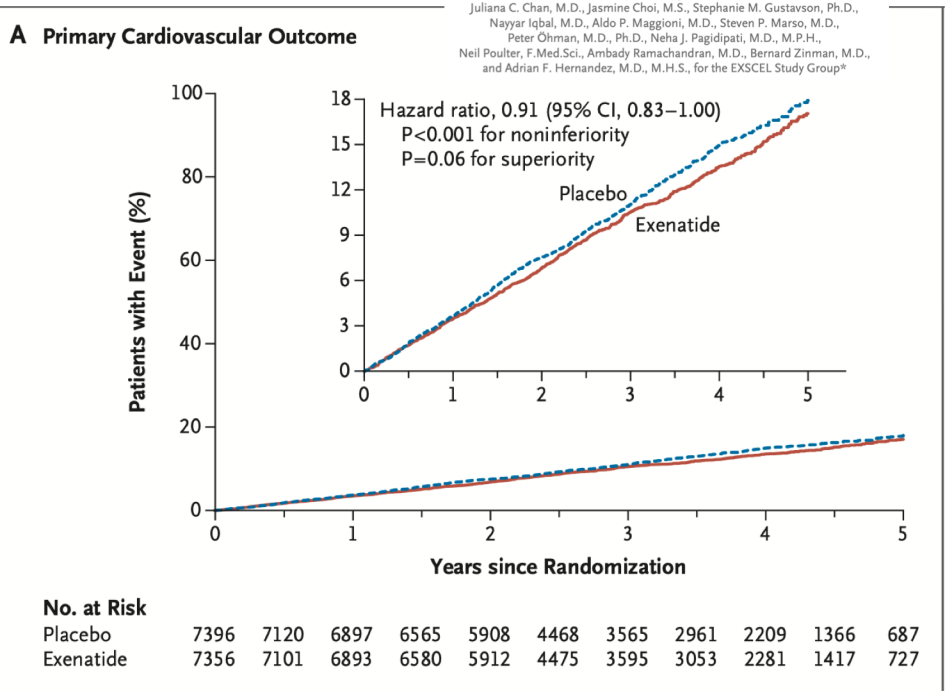
The trial was event-driven. We estimated that 6000 patients would need to be enrolled for us to observe 844 primary end-point events, which would provide the study with 96% power to show noninferiority, assuming a true hazard ratio of 1.0, and with 90% power to show superiority, assuming a true hazard ratio of 0.80. More complete details regarding sample-size assumptions have been published previously.[26]

N Engl J Med 2015;373:2247-57. DOI: 10.1056/NEJMoa1509225

CVOTs. El ejemplo de los arGLP-1. Post-aprobación. Superioridad no alcanzada

Effects of Once-Weekly Exenatide on Cardiovascular Outcomes in Type 2 Diabetes

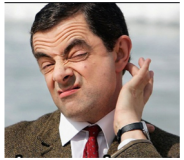
Rury R. Holman, F.Med.Sci., M. Angelyn Bethel, M.D., Robert J. Mentz, M.D., Vivian P. Thompson, M.P.H., Yuliya Lokhnygina, Ph.D., John B. Buse, M.D., Ph.D., Juliana C. Chan, M.D., Jasmine Choi, M.S., Stephanie M. Gustavson, Ph.D., Nayyar Iqbal, M.D., Aldo P. Maggioni, M.D., Steven P. Marso, M.D., Peter Öhman, M.D., Ph.D., Neha J. Pagidipati, M.D., M.P.H., Neil Poulter, F.Med.Sci., Ambady Ramachandran, M.D., Bernard Zinman, M.D., and Adrian F. Hernandez, M.D., M.H.S., for the EXSCEL Study Group*



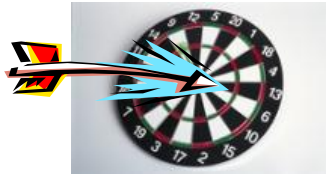
PRIMARY AND SECONDARY CARDIOVASCULAR OUTCOMES

A primary composite outcome event occurred in 839 of 7356 patients (11.4%; 3.7 events per 100 person-years) in the exenatide group and in 905 of 7396 patients (12.2%; 4.0 events per 100 person-years) in the placebo group (hazard ratio, 0.91; 95% CI, 0.83 to 1.00), with the intention-to-treat analysis indicating that exenatide, administered once weekly, was noninferior to placebo with respect to safety (P<0.001 for noninferiority) and was not superior to placebo with respect to efficacy (P=0.06 for superiority) (Table 1 and Fig. 2, and Fig. S6 in the Supplementary Appendix).

The risk of death from any cause was 6.9% in the exenatide group and 7.9% in the placebo group (hazard ratio, 0.86; 95% CI, 0.77 to 0.97); this difference was not considered to be statistically significant on the basis of the hierarchical testing plan (Table 1 and Fig. 2).



N Engl J Med 2017;377:1228-39. DOI: 10.1056/NEJMoa1612917



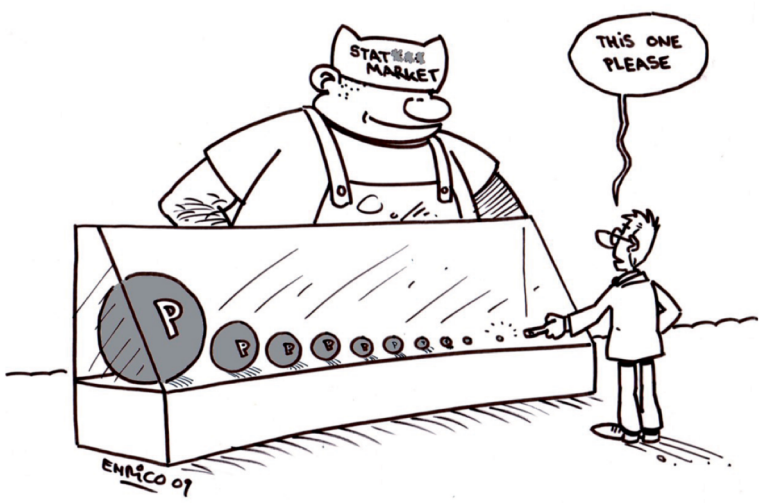
Qué es ‘a priori’?

Diseño incorrecto → cambio de planes

Problemas que genera

Diseño

Ejecución



Resultados

Multiplicidad: incremento del error de tipo I (a)

Análisis "post-hoc" guiados por resultados:
Sesgos y posible ausencia de validez externa

Fuente: imagen de internet, sin conocimientos sobre derechos de autor. Sólo empleada con fines docentes

Manejo de la multiplicidad en las **variables (o hipótesis)**

Una única variable

Identificar una variable principal y el resto son secundarias

El EC es concluyente sii la variable principal es significativa (valor de p <= 0.05, p.e.)

El cálculo del tamaño de la muestra depende sólo de la variable principal.

Digamos, el diseño ibérico

17

CVOTs. El ejemplo de los arGLP1
Post-aprobación. Ejemplo de una única hipótesis (superioridad)

Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial

Hertzel C Gerstein, Helen M Colhoun, Gilles R Dagenais, Rafael Diaz, Mark Lakshmanan, Prem Pais, Jeffrey Probstfield, Jeffrey S Riesmeyer, Matthew C Riddle, Lars Rydén, Denis Xavier, Charles Messan Atisso, Leanne Dyal, Stephanie Hall, Purnima Rao-Melacini, Gloria Wong, Alvaro Avezum, Jan Basile, Namsik Chung, Ignacio Conget, William C Cushman, Edward Franek, Nicolae Hancu, Markolf Hanefeld, Shaun Holt, Petr Jansky, Matyas Keltai, Fernando Lanas, Lawrence A Leiter, Patricio Lopez-Jaramillo, Ernesto German Cardona Munoz, Valdis Pirags, Nana Pogosova, Peter J Raubenheimer, Jonathan E Shaw, Wayne H-H Sheu, Theodora Temelkova-Kurktschiev, for the REWIND Investigators*

Statistical analysis

Assuming a primary outcome rate of 2% per year in the placebo group, a two-sided significance level of 5%, a 3-year recruitment period, and an annual dropout rate of 0·15%, follow-up of at least 9600 people for a maximum of 8 years, and accrual of 1200 unique primary outcomes was estimated to provide at least 90% power to detect a hazard ratio (HR) of 0·82 or lower for the primary outcome and 80% power to detect an HR of 0·85 or lower. The independent data monitoring committee did a formal interim analysis for superiority after 756 unique primary outcomes had occurred, after which continuation of the trial was recommended.

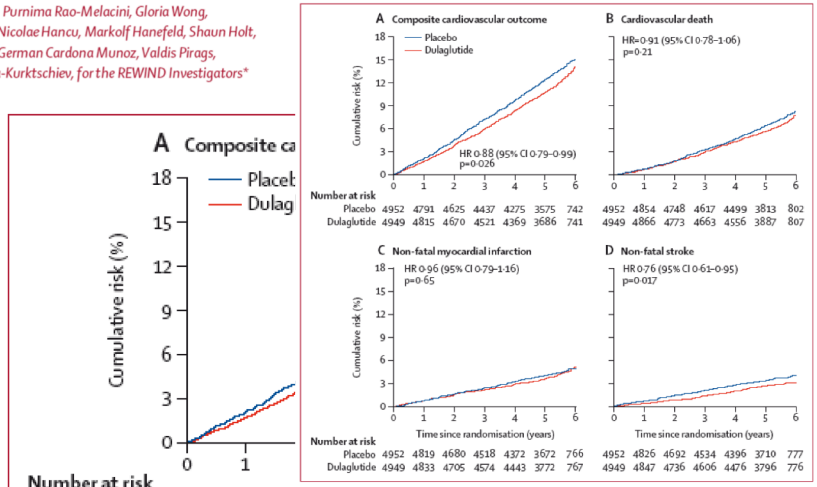


Figure 2: Cumulative incidence of cardiovascular outcomes
HR=hazard ratio. HbA_{1c}=glycated haemoglobin A_{1c}.

Gerstein HC, Colhoun HM, Dagenais GR, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. Lancet 2019;394:121-30.

Manejo de la multiplicidad en las variables (o hipótesis)

Jerárquico

Se pre-especifica la jerarquía de las n variables principales.

Se evalúa si la 1ª variables es significativa al 0.05, si lo es se comprueba la 2ª al 0.05, si lo es se comprueba la 3ª al 0.05 de significación. **Se detiene en el momento que una de las variables no resulta significativa**.

El EC es concluyente si la primera variable principal observada resulta significativa.

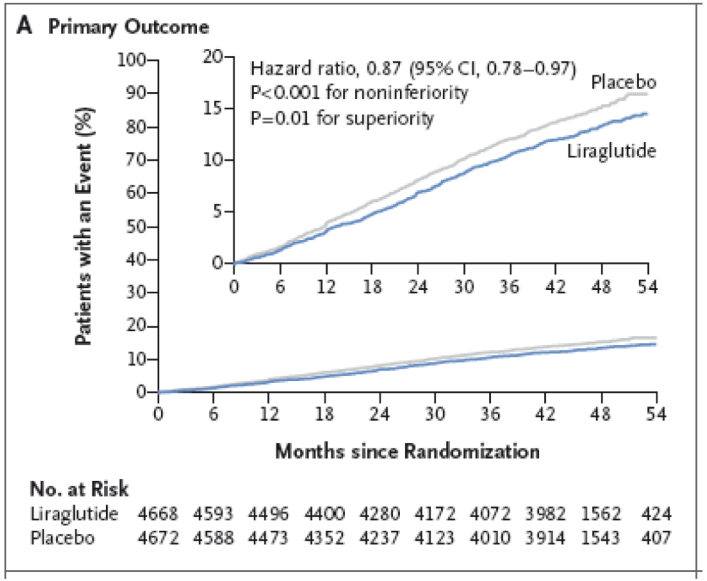
Calcular el tamaño de la muestra por separado para cada una de las variables co-primarias utilizando un Error de Tipo I de 0,05.

CVOTs. El ejemplo de los arGLP1
Post-aprobación. Ejemplo de jerarquía en las hipótesis (no-inferioridad → superioridad)

Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes

Steven P. Marso, M.D., Gilbert H. Daniels, M.D., Kirstine Brown-Frandsen, M.D., Peter Kristensen, M.D., E.M.B.A., Johannes F.E. Mann, M.D., Michael A. Nauck, M.D., Steven E. Nissen, M.D., Stuart Pocock, Ph.D., Neil R. Poulter, F.Med.Sci., Lasse S. Ravn, M.D., Ph.D., William M. Steinberg, M.D., Mette Stockner, M.D., Bernard Zinman, M.D., Richard M. Bergenstal, M.D., and John B. Buse, M.D., Ph.D., for the LEADER Steering Committee on behalf of the LEADER Trial Investigators*

hazard ratio. We used a hierarchical testing strategy for the liraglutide group versus the placebo group, first testing for noninferiority and subsequently for superiority. Noninferiority was established for the primary outcome if the upper limit of the two-sided 95% confidence interval of the hazard ratio was less than 1.30, and superiority was established if the upper limit was less than 1.00. In addition, prespecified sensitivity



¿Puedo cambiar de superioridad a no-inferioridad? ¿Y al revés? ¿Cuándo?

The European Agency for the Evaluation of Medicinal Products
Evaluation of Medicines for Human Use

London, 27 July 2000
CPMP/EWP/482/99

COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS (CPMP)

POINTS TO CONSIDER ON SWITCHING BETWEEN SUPERIORITY AND NON-INFERIORITY

DISCUSSION IN THE EFFICACY WORKING PARTY (EWP)	February 1999/ September 1999
TRANSMISSION TO THE CPMP	September 1999
RELEASE FOR CONSULTATION	September 1999
DEADLINE FOR COMMENTS	December 1999
RE-SUBMISSION TO THE EFFICACY WORKING PARTY	June 2000
APPROVAL BY THE CPMP	July 2000

A modo de resumen:

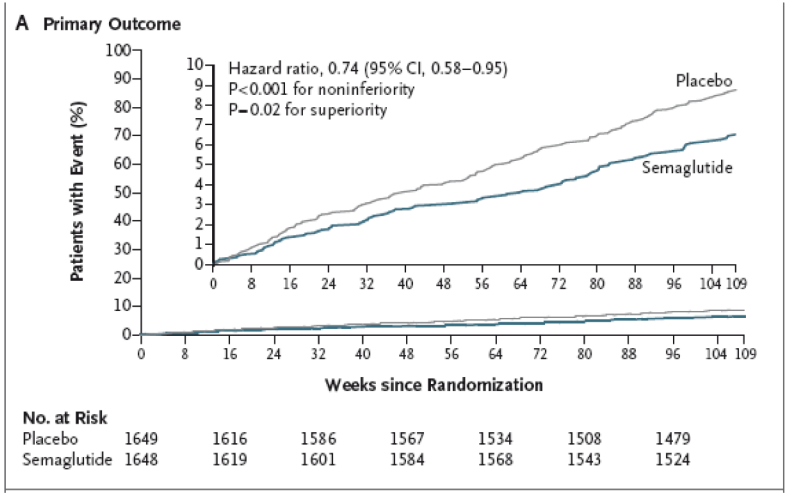
"In general, the switching between superiority hypotheses to non-inferiority is not permitted. If the trial don't confirms the hypothesis of no-inferiority or equivalence, is possible an statistical evaluation of superiority (or inferiority) in order to explain the sense of this differences."

CVOTs. El ejemplo de los arGLP1 Pre-aprobación. Ejemplo de jerarquía en las hipótesis (no-inferioridad → superioridad) No pre-especificada

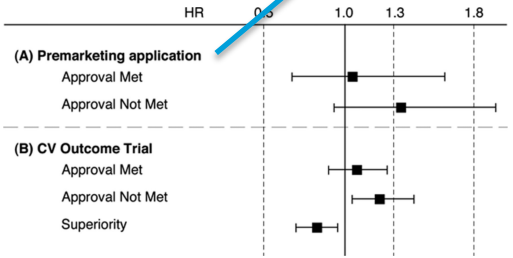
ORIGINAL ARTICLE

Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes

Steven P. Marso, M.D., Stephen C. Bain, M.D., Agostino Consoli, M.D., Freddy G. Eliaschewitz, M.D., Esteban Jódar, M.D., Lawrence A. Leiter, M.D., Ildiko Lingvay, M.D., M.P.H., M.S.C.S., Julio Rosenstock, M.D., Jochen Seufert, M.D., Ph.D., Mark L. Warren, M.D., Vincent Woo, M.D., Oluf Hansen, M.Sc., Anders G. Holst, M.D., Ph.D., Jonas Pettersson, M.D., Ph.D., and Tina Vilsbøll, M.D., D.M.Sc., for the SUSTAIN-6 Investigators*



Semaglutide, a GLP-1 analogue with an extended half-life of approximately 1 week (which permits once-weekly subcutaneous administration),[4] is currently in development but not yet approved for the treatment of type 2 diabetes. Regulatory guidance specifies the need to establish the cardiovascular safety of new therapies for type 2 diabetes in order to rule out excess cardiovascular risk.[5] The preapproval Trial to Evaluate Cardiovascular and Other Long-term Outcomes with Semaglutide in Subjects with Type 2 Diabetes (SUSTAIN-6) was designed to assess the noninferiority of semaglutide as compared with placebo in terms of cardiovascular safety in patients with type 2 diabetes.



ratio of 1.00. We determined that the enrollment of 3260 patients would be required to determine the primary outcome in at least 122 patients and provide a power of 90% to reject a hazard ratio of at least 1.80 at the 0.05 level of significance.

The prespecified analysis for the primary outcome was a Cox proportional-hazards model, with pooled treatment (semaglutide vs. placebo) as a fixed factor, and categorized according to all possible combinations of stratification factors used for randomization. The primary hypothesis was for noninferiority for the primary outcome. Such noninferiority was confirmed if the upper boundary of the two-sided 95% confidence interval of the hazard ratio was below the noninferiority margin of 1.80.[5] Testing for superiority for the primary outcome was not prespecified or adjusted for multiplicity. We conducted prespecified sensitivity analyses of the primary outcome, using alternative patient selection and data-censoring strategies for exposure to treatment, and per-protocol sensitivity analyses were performed post hoc. (Details regarding analysis sets are provided

N Engl J Med 2016;375:1834-44. DOI: 10.1056/NEJMoa1607141

Cosas del apriorismo…

• Preguntas

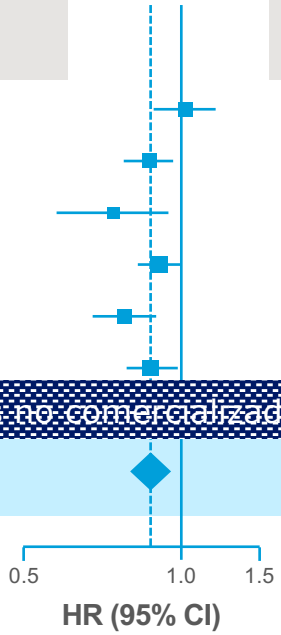
- **Poblaciones de los estudios**
  - ¿Los CVOTs incluyen poblaciones similares?
  - ¿Como impacta las diferencias en la población en el tiempo de seguimiento, tamaño de la muestra?
  - ¿En que medida los CVOTs son aplicables a la población DM2 de la practica clínica?
  - ¿Como podemos interpretar los análisis de subgrupos?

CV Outcomes in GLP-1 RA Studies

Heterogeneidad en resultados!

Meta-Analysis of MACE in GLP-1 RA CVOTs[1]

Trial	Treatment (n/N)	Placebo (n/N)	HR (95% CI)
ELIXA	400/3034	399/3034	1.02 (0.89, 1.17)
LEADER	608/4668	694/4672	0.87 (0.78, 0.97)
SUSTAIN-6	108/1648	146/1649	0.74 (0.58, 0.95)
EXSCEL	839/7356	905/7396	0.91 (0.83, 1.00)
Harmony Outcomes	338/4731	428/4732	0.78 (0.68, 0.90)
REWIND	594/4949	663/4952	0.88 (0.79, 0.99)
Resultados no mostrados por tratarse de fármacos no comercializados en España			
Overall (I^2=40.9%; P=.118)	2948/27,977	3304/28,027	0.88 (0.82, 0.94)

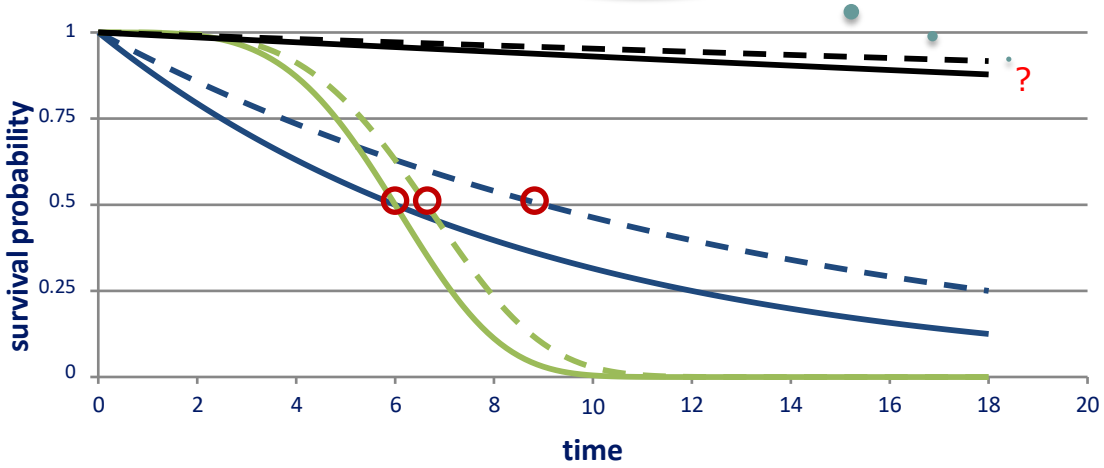


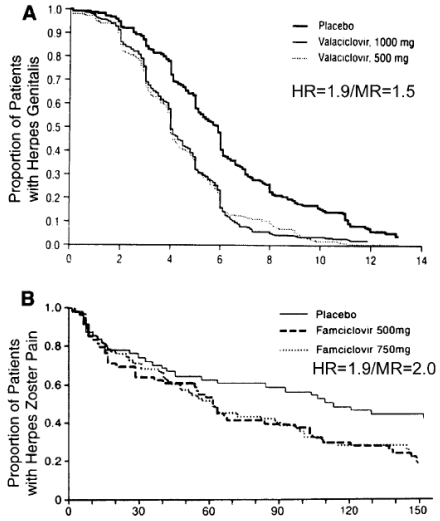
HR = hazard ratio.

1. Kirstenen S, et al. *Lancet*. 2019;7:776-785.

CV Outcomes in GLP-1 RA Studies

Los HR entonces también son diferentes…





Y ojo… ni que fuesen homogéneos

Time (days)

Diferencias en las características de los pacientes incluidos en estudios previos al REWIND

Heterogeneidad población incluida !

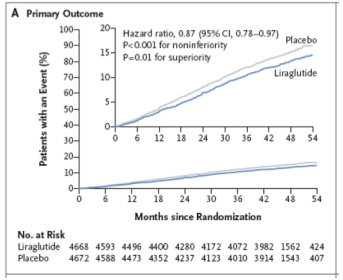
	Semaglutide[1]	Liraglutide[2]	Lixisenatide[3]	Exenatide QW[4]	Albiglutide[5]
Nº de Pacientes	3.297	9.340	6.068	14.752	9.463
Edad	65	64	60	62 (mediana)	64
Sexo (mujeres)	39%	36%	31%	38%	31%
ECV establecida	83%	81%	100%	73%	100%
IMC	33	33	30	32 (mediana)	32
HbA1c	8.7%	8.7%	7.7%	8.0% (mediana)	8.7%
Nº de eventos	254	1302	805	1744	766
Seguimiento	2.1 años	3.8 años	2.1 años	3.2 años	1.6 años
MACE-3	0.74 (0.58-0.95)	0.87 (0.78-0.97)	1.02 (0.89-1.17)	0.91 (0.83-1.00)	0.78 (0.68, 0.90)

1. Pfeffer MA *et al. N Engl J Med* 2015;373:2247; 2. Marso SP *et al. N Engl J Med* 2016;375:311; 3. Marso SP *et al. N Engl J Med* 2016;375:1834; 4. Holman RR *et al. N Engl J Med* 2017;377:1228 5. Hernandez AF et al. Lancet 2018;392(10157):1519-1529. 6. Gerstein HC et al. Diabetes Obes Metab 2018;20(1):42-49.

¿Y en el tiempo de seguimiento?

Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes

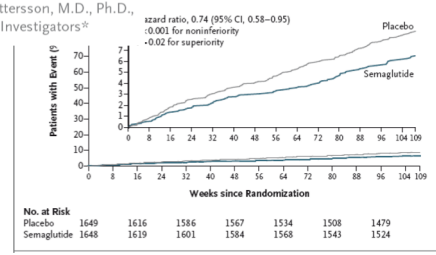
Steven P. Marso, M.D., Gilbert H. Daniels, M.D., Kirstine Brown-Frandsen, M.D., Peter Kristensen, M.D., E.M.B.A., Johannes F.E. Mann, M.D., Michael A. Nauck, M.D., Steven E. Nissen, M.D., Stuart Pocock, Ph.D., Neil R. Poulter, F.Med.Sci., Lasse S. Ravn, M.D., Ph.D., William M. Steinberg, M.D., Mette Stockner, M.D., Bernard Zinman, M.D., Richard M. Bergenstal, M.D., and John B. Buse, M.D., Ph.D., for the LEADER Steering Committee on behalf of the LEADER Trial Investigators*



ORIGINAL ARTICLE

Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes

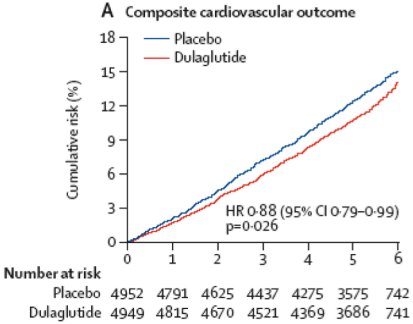
Steven P. Marso, M.D., Stephen C. Bain, M.D., Agostino Consoli, M.D., Freddy G. Eliaschewitz, M.D., Esteban Jódar, M.D., Lawrence A. Leiter, M.D., Ildiko Lingvay, M.D., M.P.H., M.S.C.S., Julio Rosenstock, M.D., Jochen Seufert, M.D., Ph.D., Mark L. Warren, M.D., Vincent Woo, M.D., Oluf Hansen, M.Sc., Anders G. Holst, M.D., Ph.D., Jonas Pettersson, M.D., Ph.D., and Tina Vilsbøll, M.D., D.M.Sc., for the SUSTAIN-6 Investigators*



También!

Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial

Hertzel C Gerstein, Helen M Colhoun, Gilles R Dagenais, Rafael Diaz, Mark Lakshmanan, Prem Pais, Jeffrey Probstfield, Jeffrey S Riesmeyer, Matthew C Riddle, Lars Rydén, Denis Xavier, Charles Messan Atisso, Leanne Dyal, Stephanie Hall, Purnima Rao-Melacini, Gloria Wong, Alvaro Avezum, Jan Basile, Namsik Chung, Ignacio Conget, William C Cushman, Edward Franek, Nicolae Hancu, Markolf Hanefeld, Shaun Holt, Petr Jansky, Matyas Keltai, Fernando Lanas, Lawrence A Leiter, Patricio Lopez-Jaramillo, Ernesto German Cardona Munoz, Valdis Pirags, Nana Pogosova, Peter J Raubenheimer, Jonathan E Shaw, Wayne H-H Sheu, Theodora Temelkova-Kurktschiev, for the REWIND Investigators*



MOTIVO DE LA HETEROGENEIDAD:

En estos diseños se ha de buscar el equilibrio en la planificación

- Tamaño de la muestra en ‘estudios de tiempo hasta evento’:
  - Magnitud del efecto mínimamente relevante: HR
  - Precisión necesaria: Amplitud del 95%CI

Número de eventos que necesito observar

Riesgo basal de presentar el evento

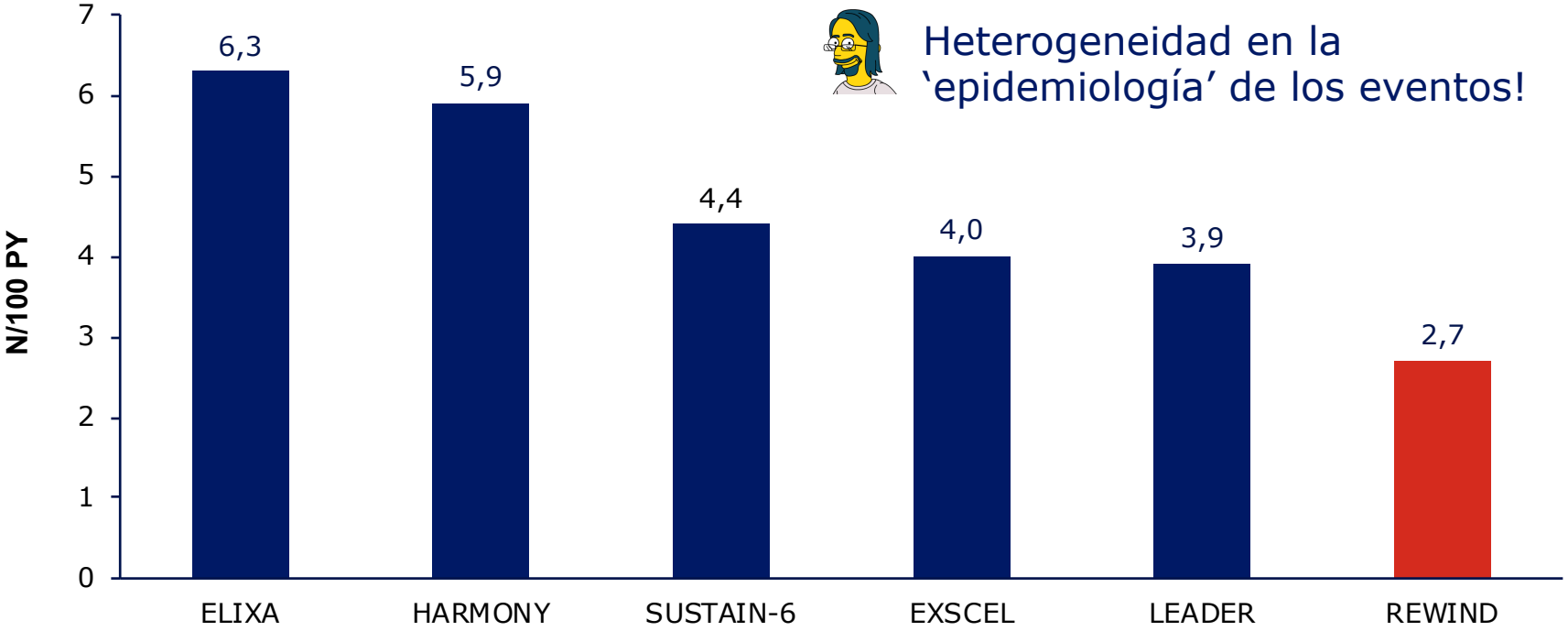
Tiempo de seguimiento

!!!

N ?

Fijados los errores estadísticos de Tipo I Y Tipo II

Daños colaterales para la comparabilidad: Annual Placebo MACE Rates by Trial[1-7]

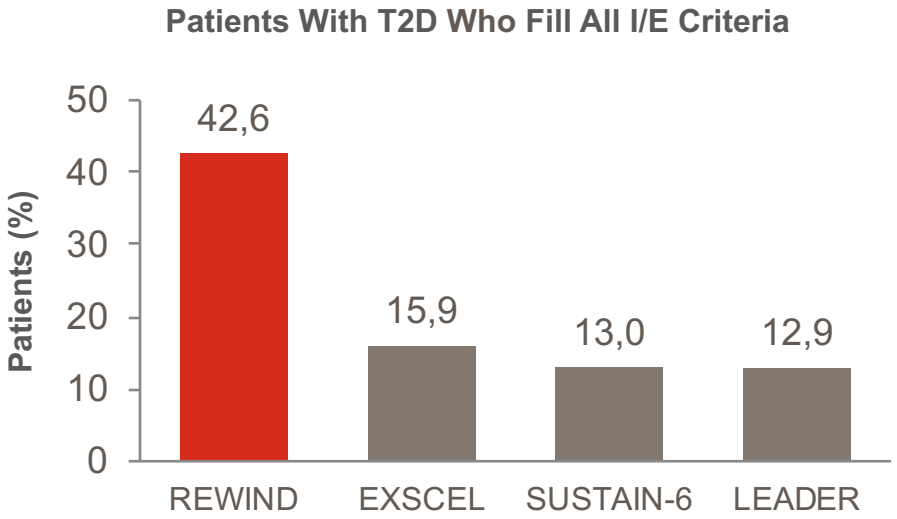


Heterogeneidad en la 'epidemiología' de los eventos!

Caution should be used when comparing GLP-1 RA CVOTs due to differences in study design, population, and key inclusion/exclusion criteria.

1. Zelniker TA, et al. *Circulation.* 2019;139(17):2022-2031. 2. Marso SP, et al. *N Engl J Med*. 2016;375(19):1834-1844. 3. Marso SP, et al. *N Engl J Med*. 2016;375(4):311-322. 4. Pfeffer MA, et al. *N Engl J Med*. 2015;373(23):2247-2257. 5. Holman RR, et al. *N Engl J Med*. 2017;377(13):1228-1239. 6. Gerstein HC, et al. *Diabetes Obes Metab*. 2018;20(1):42-49. 7. Hernandez AF, et al. *Lancet*. 2018;392(10157):1519-1529.

Daños colaterales para la comparabilidad: Representatividad de la población incluida



US Population with T2D:
Applying the I/E criteria to the population of adults with T2D, REWIND would have included over 2.5 times more patients than EXSCEL and over 3 times more patients than LEADER and SUSTAIN-6.

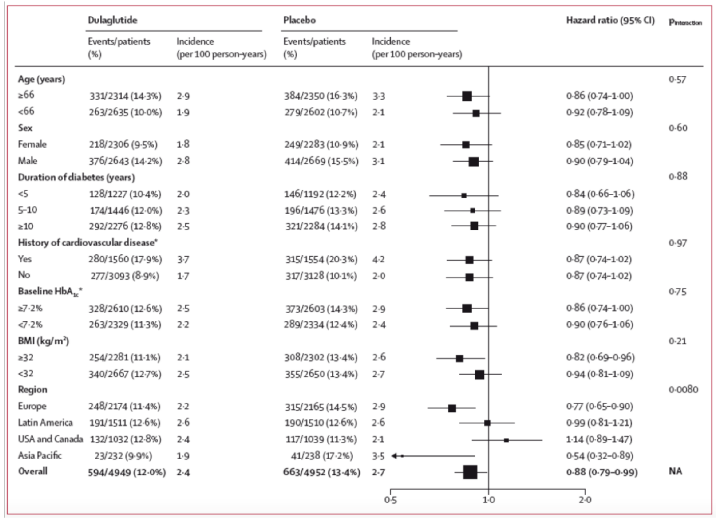
REWIND participants were similar to patients with T2D and CV risk factors who are routinely seen in clinical practice

I/E = inclusion/exclusion.
Boye KS, et al. *Diabetes Obes Metab*. 2019;21(6):1299-1304.

Consecuencia de todo ello:
Interferencia en la validez externa del resultado

Richard, Mr Chenolock, the insurance man, is here to determine your life expectancy

¿Qué es esto?

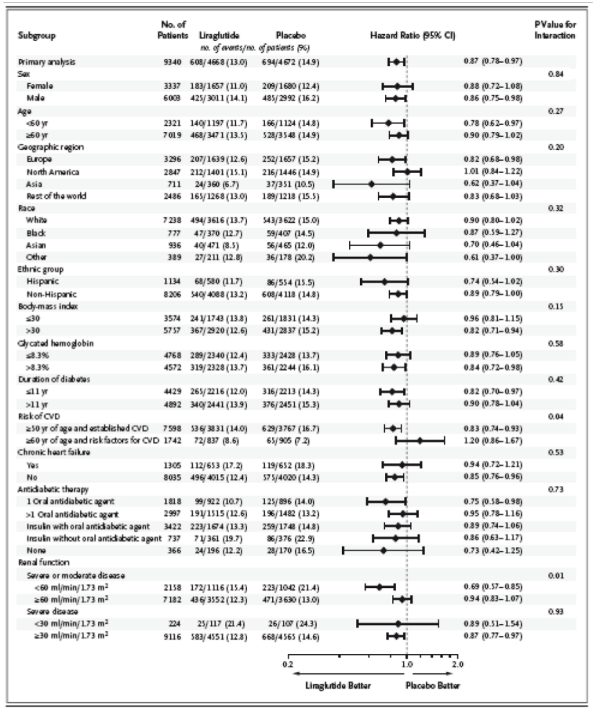


Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial

Hertzel C Gerstein, Helen M Colhoun, Gilles R Dagenais, Rafael Diaz, Mark Lakshmanan, Prem Pais, Jeffrey Probstfield, Jeffrey S Riesmeyer, Matthew C Riddle, Lars Rydén, Denis Xavier, Charles Messan Atisso, Leanne Dyal, Stephanie Hall, Purnima Rao-Melacini, Gloria Wong, Alvaro Avezum, Jan Basile, Namsik Chung, Ignacio Conget, William C Cushman, Edward Franek, Nicolae Hancu, Markolf Hanefeld, Shaun Holt, Petr Jansky, Matyas Keltai, Fernando Lanas, Lawrence A Leiter, Patricio Lopez-Jaramillo, Ernesto German Cardona Munoz, Valdis Pirags, Nana Pogosova, Peter J Raubenheimer, Jonathan E Shaw, Wayne H-H Sheu, Theodora Temelkova-Kurktschiev, for the REWIND Investigators*

Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes

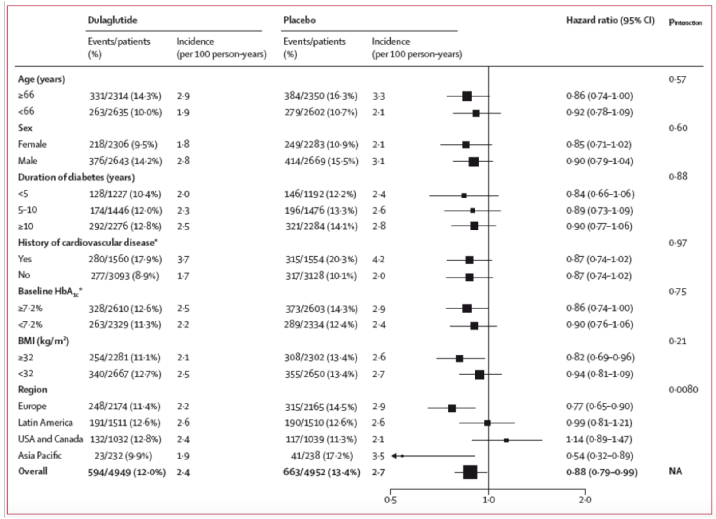
Steven P. Marso, M.D., Gilbert H. Daniels, M.D., Kirstine Brown-Frandsen, M.D., Peter Kristensen, M.D., E.M.B.A., Johannes F.E. Mann, M.D., Michael A. Nauck, M.D., Steven E. Nissen, M.D., Stuart Pocock, Ph.D., Neil R. Poulter, F.Med.Sci., Lasse S. Ravn, M.D., Ph.D., William M. Steinberg, M.D., Mette Stockner, M.D., Bernard Zinman, M.D., Richard M. Bergenstal, M.D., and John B. Buse, M.D., Ph.D., for the LEADER Steering Committee on behalf of the LEADER Trial Investigators*



Marso SP, et al. "Liraglutide and cardiovascular outcomes in type 2 diabetes". The New England Journal of Medicine. 2016. 375(4):311-322.

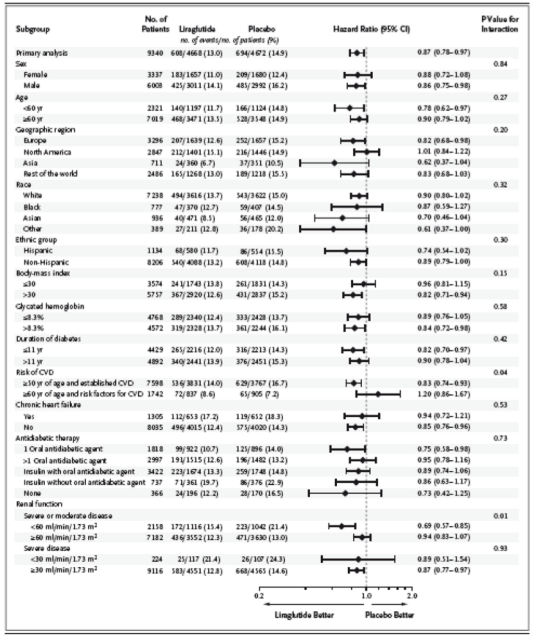
Gerstein HC, Colhoun HM, Dagenais GR, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. Lancet 2019;394:121-30.

¿Qué utilidad tiene?



Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes

Steven P. Marso, M.D., Gilbert H. Daniels, M.D., Kirstine Brown-Frandsen, M.D., Peter Kristensen, M.D., E.M.B.A., Johannes F.E. Mann, M.D., Michael A. Nauck, M.D., Steven E. Nissen, M.D., Stuart Pocock, Ph.D., Neil R. Poulter, F.Med.Sci., Lasse S. Ravn, M.D., Ph.D., William M. Steinberg, M.D., Mette Stockner, M.D., Bernard Zinman, M.D., Richard M. Bergenstal, M.D., and John B. Buse, M.D., Ph.D., for the LEADER Steering Committee on behalf of the LEADER Trial Investigators*



Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial

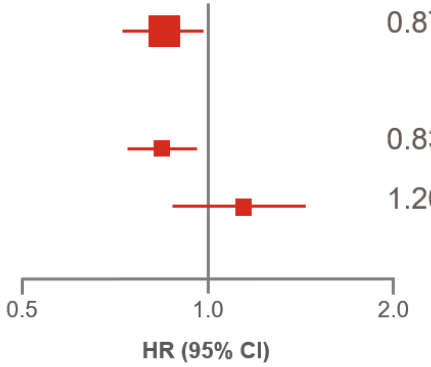
Hertzel C Gerstein, Helen M Colhoun, Gilles R Dagenais, Rafael Diaz, Mark Lakshmanan, Prem Pais, Jeffrey Probstfield, Jeffrey S Riesmeyer, Matthew C Riddle, Lars Rydén, Denis Xavier, Charles Messan Atisso, Leanne Dyal, Stephanie Hall, Purnima Rao-Melacini, Gloria Wong, Alvaro Avezum, Jan Basile, Namsik Chung, Ignacio Conget, William C Cushman, Edward Franek, Nicolae Hancu, Markolf Hanefeld, Shaun Holt, Petr Jansky, Matyas Keltai, Fernando Lanas, Lawrence A Leiter, Patricio Lopez-Jaramillo, Ernesto German Cardona Munoz, Valdis Pirags, Nana Pogosova, Peter J Raubenheimer, Jonathan E Shaw, Wayne H-H Sheu, Theodora Temelkova-Kurktschiev, for the REWIND Investigators*

Marso SP, et al. "Liraglutide and cardiovascular outcomes in type 2 diabetes". The New England Journal of Medicine. 2016. 375(4):311-322.

Gerstein HC, Colhoun HM, Dagenais GR, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. Lancet 2019;394:121-30.

Criterio de valoración del MACE-3 en pacientes con y sin enfermedad CV establecida, estudio LEADER

Subgrupos	Liraglutida Eventos/Total (%)	Placebo Eventos/Total (%)	HR (95% CI)	*P*-Value de Interacción
LEADER	608/4668 (13.0)	694/4672 (14.8)	0.87 (0.78, 0.97)	
Historial de ECV				
≥ 50 años con ECV establecida	536/3831 (14.0)	629/3767 (16.7)	0.83 (0.74, 0.93)	
≥ 60 años con FRCV	72/837 (8.6)	65/905 (7.2)	1.20 (0.86, 1.67)	0.04

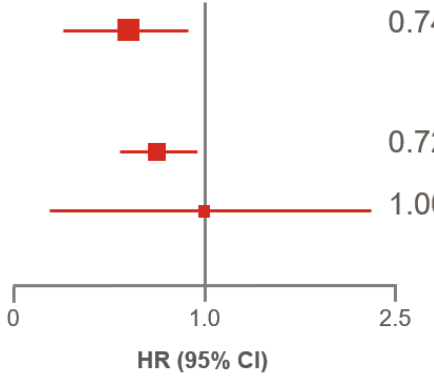


¿En pacientes con FRCV **no se observaron beneficios CV** tras el tratamiento con liraglutida?
O Bien…
¿podría haber **heterogeneidad en el efecto** del fármaco en función de la historia de ECV?

Marso et al. N Engl J Med. 2016 Jul 28;375(4):311-22

Criterio de valoración del MACE-3 en pacientes con y sin enfermedad CV establecida, estudio SUSTAIN-6

	Semaglutida	Placebo			
Subgrupos	**Eventos/Total (%)**	**Eventos/Total (%)**		**HR (95% CI)**	***P*-Value de interacción**
SUSTAIN-6	108/1648 (7.0)	146/1649 (9.0)		0.74 (0.58, 0.95)	
Historial de ECV					
ECV establecida	98/1353 (5.4)	137/1382 (9.9)		0.72 (0.55, 0.93)	
FRCV	10/295 (3.4)	9/267 (3.4)		1.00 (0.41, 2.46)	0.49

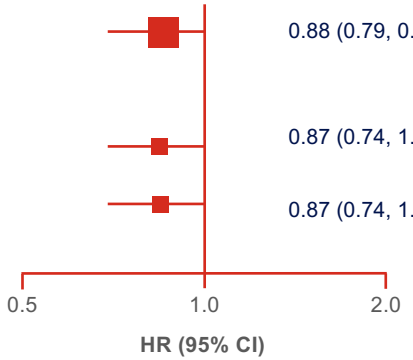


¿Podemos sacar conclusiones con únicamente **19 eventos CV** en el subgrupo de FRCV?

Marso et al. N Engl J Med. 2016 Jul 28;375(4):311-22

Criterio de valoración del MACE-3 en pacientes con y sin enfermedad CV establecida, estudio REWIND

	Dulaglutide		Placebo			
Subgroups	**Events/Total (%)**	**Rate/100 PY**	**Events/Total (%)**	**Rate/100 PY**	**HR (95% CI)**	***P*-Value for Interaction**
Overall	594/4949 (12.0)	2.4	663/4952 (13.4)	2.7	0.88 (0.79, 0.99)	
History of CVD						
Prior CVD	280/1560 (17.9)	3.7	315/1554 (20.3)	4.2	0.87 (0.74, 1.02)	
No prior CVD	277/3093 (8.9)	1.7	317/3128 (10.1)	2.0	0.87 (0.74, 1.02)	.970



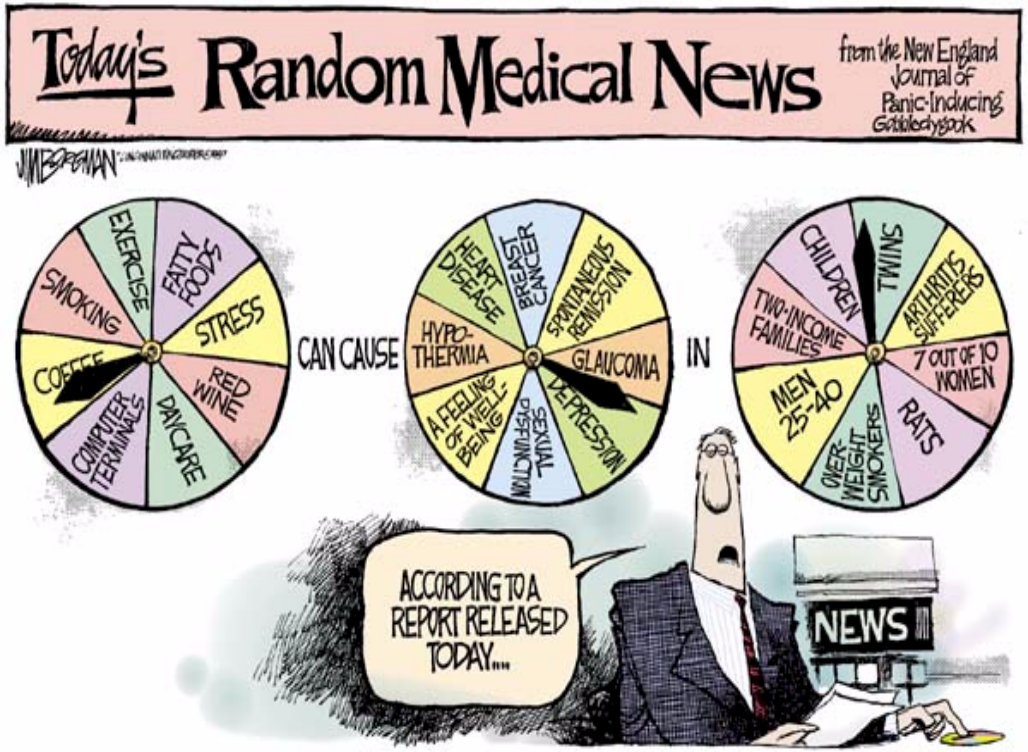
¿Se aprecia un **efecto homogéneo** independientemente de la historia previa de CVD?

Gerstein HC, et al. *Lancet.* 2019;394(10193):121-130.

Consideraciones

- Buscar validación interna en grupos de interés (Fase III)
  - Consistencia en la estimación global
- ¿Me los creo?
  - Plausibilidad biológica
  - Validación con datos externos
- Problemas de multiplicidad
  - Aumento de falsos positivos
  - En caso de Fase III negativos pueden crear falsas expectativas

Ojo con los análisis de subgrupos



• Preguntas

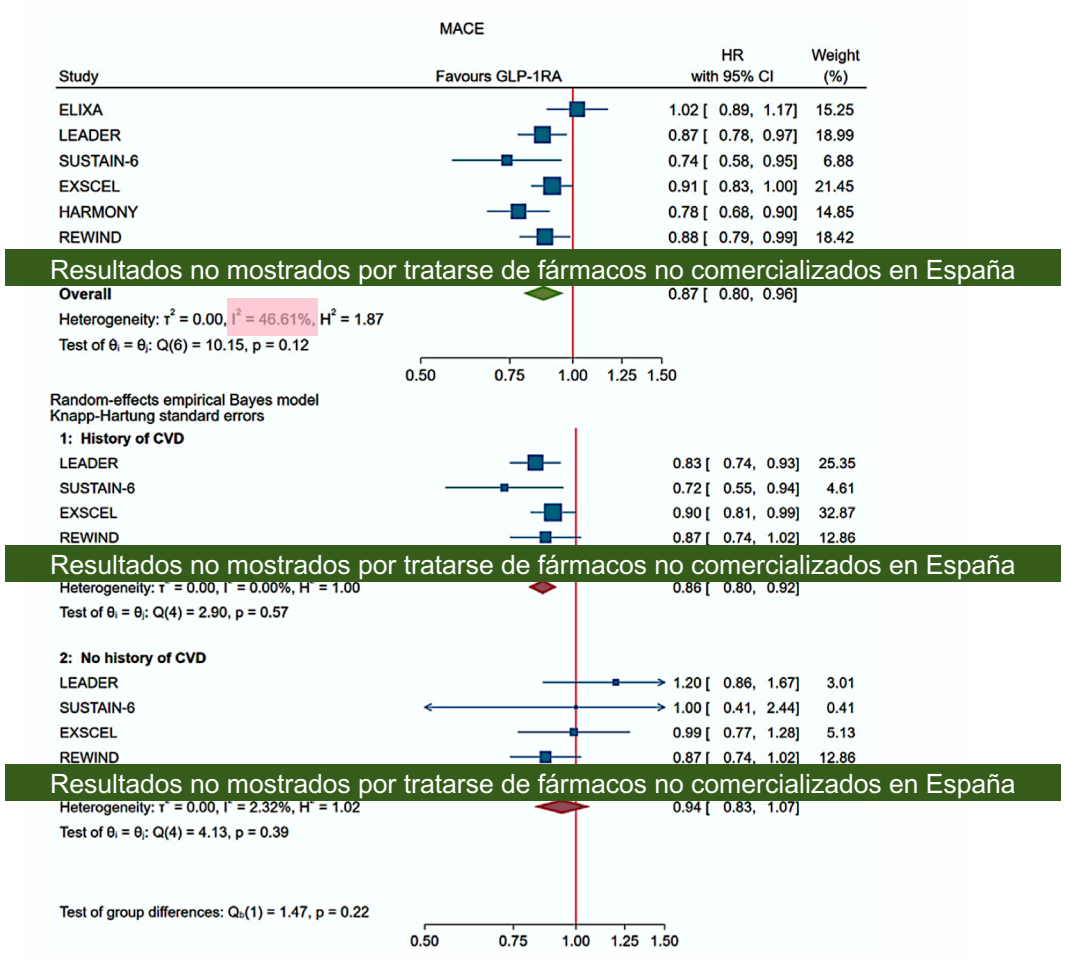
- **Metaanálisis:**
  - ¿Por qué los hacemos?
  - ¿Que limitaciones y riesgos tienen?
  - ¿Que precauciones tengo que tener para su correcta interpretación?
  - Teniendo en cuenta los diferentes meta-analisis que se están publicando con arGLP-1s. ¿Pensáis realmente que hay un efecto de clase?

Meta-analisis

Forest plots depicting results of meta-analysis of

- CVOTs with GLP-1RAs concernng MACE (top)
- CVOTs with GLP-1RAs concerning MACE in patients with or without history of CVD at baseline (bottom).

Results are expressed as HR (hazard ratio)

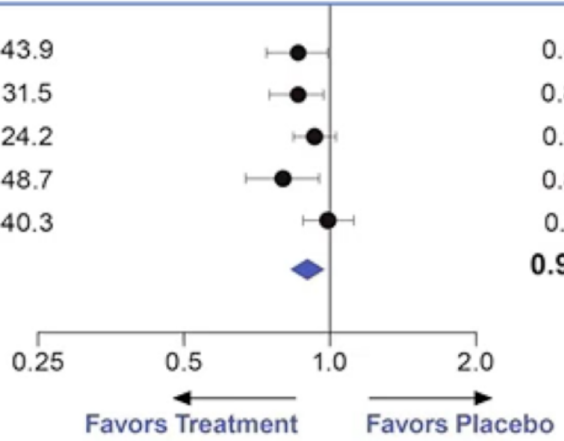


Giugliano D DOM 2019; DOI: 10.1111/dom.13847

Time to first MACE

MACE

	Treatment Rate/1000 patient-years	Placebo Rate/1000 patient-years	Hazard ratio (95% CI)
EMPA-REG OUTCOME	37.4	43.9	0.86 (0.74-0.99)
CANVAS Program	26.9	31.5	0.86 (0.75-0.97)
DECLARE-TIMI 58	22.6	24.2	0.93 (0.84-1.03)
CREDENCE	38.7	48.7	0.80 (0.67-0.95)
VERTIS CV*	40.0	40.3	0.99 (0.88-1.12)
Pooled estimate			**0.90 (0.85-0.95)**
(Q statistic $P = 0.27$; $I^2 = 23.4\%$)			

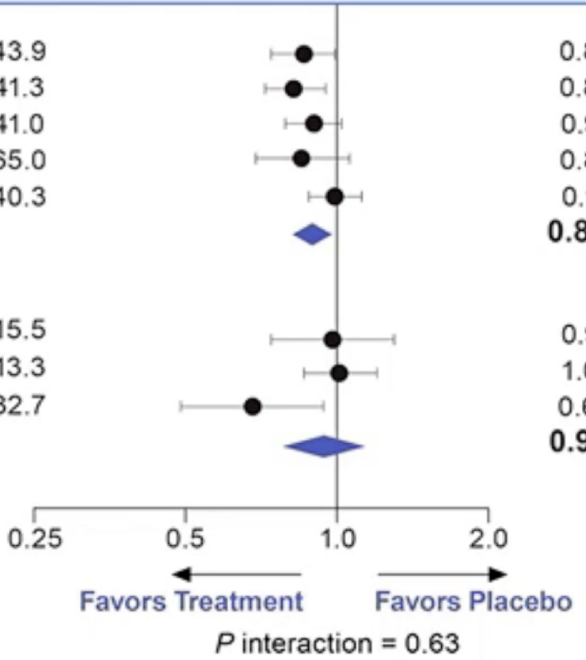


*Intention-to-treat population was used for consistency with other trials.
CI, confidence interval; MACE, major adverse cardiovascular events.

Time to first MACE – subgroup analysis by ASCVD

MACE

		Treatment Rate/1000 patient-years	Placebo Rate/1000 patient-years	Hazard ratio (95% CI)
ASCVD	EMPA-REG OUTCOME	37.4	43.9	0.86 (0.74-0.99)
	CANVAS Program	34.1	41.3	0.82 (0.72-0.95)
	DECLARE-TIMI 58	36.8	41.0	0.90 (0.79-1.02)
	CREDENCE	55.6	65.0	0.85 (0.69-1.06)
	VERTIS CV	40.0	40.3	0.99 (0.88-1.12)
	Pooled estimate (Q statistic $P = 0.34$; $I^2 = 11.8\%$)			**0.89 (0.84-0.95)**
No ASCVD	CANVAS Program	15.8	15.5	0.98 (0.74-1.30)
	DECLARE-TIMI 58	13.4	13.3	1.01 (0.86-1.20)
	CREDENCE	22.0	32.7	0.68 (0.49-0.94)
	Pooled estimate (Q statistic $P = 0.10$; $I^2 = 56.5\%$)			**0.94 (0.83-1.07)**



ASCVD, atherosclerotic cardiovascular disease; CI, confidence interval; MACE, major adverse cardiovascular events.

Meta-analisis

Friky ALERT!

Meta-analisis:
Necesidad de clasificar y comparar resultados

Escala de evidencia en estudios clínicos para el manejo del paciente diabético según guías publicadas por Glucodosky et al 2018	
1++	Siempre se ha hecho así
2	A mi me gusta/me va bien así
2A 2B	Mi experiencia dice que así suele ir bien De momento no me ha dado problemas
3A 3B	Bah! Pónselo que no pasa nada Prueba a ver!
4	No tienes valor para hacerlo*

* Adaptación políticamente correcta de la fuente

Meta-analisis:
Necesidad de clasificar y comparar resultados

- Son estudios comparables en **resultados**?

Advantages of I^2

- Focuses attention on the effect of any heterogeneity on the meta-analysis
- Interpretation is intuitive—the percentage of total variation across studies due to heterogeneity
- Can be accompanied by an uncertainty interval
- Simple to calculate and can usually be derived from published meta-analyses
- Does not inherently depend on the number of studies in the meta-analysis
- May be interpreted similarly irrespective of the type of outcome data (eg dichotomous, quantitative, or time to event) and choice of effect measure (eg odds ratio or hazard ratio)
- Wide range of applications

How much is too much heterogeneity?

A naive categorisation of values for I^2 would not be appropriate for all circumstances, although we would tentatively assign adjectives of low, moderate, and high to I^2 values of 25%, 50%, and 75%. Figure 2 shows that about a quarter of meta-analyses have I^2 values over 50%. Quantification of heterogeneity is only one component of a wider investigation of variability across studies, the most important being diversity in clinical and methodological aspects. Meta-analysts must also

Summary points

Inconsistency of studies' results in a meta-analysis reduces the confidence of recommendations about treatment

Inconsistency is usually assessed with a test for heterogeneity, but problems of power can give misleading results

A new quantity I^2, ranging from 0-100%, is described that measures the degree of inconsistency across studies in a meta-analysis

I^2 can be directly compared between meta-analyses with different numbers of studies and different types of outcome data

I^2 is preferable to a test for heterogeneity in judging consistency of evidence

Higgins J, Thompson SG, Deeks JJ, Altman DG. Measuring inconsis- tency in meta-analyses. BMJ. 2003;327:557–560. https://doi.org/10.1136/bmj.327.7414.557

Meta-analisis:
Necesidad de clasificar y comparar resultados

- Son estudios comparables en **diseño**?
- Son estudios comparables en **variable principal**?
- Son estudios comparables en **población diana**?

How much is too much heterogeneity?

A naive categorisation of values for I^2 would not be appropriate for all circumstances, although we would tentatively assign adjectives of low, moderate, and high to I^2 values of 25%, 50%, and 75%. Figure 2 shows that about a quarter of meta-analyses have I^2 values over 50%. Quantification of heterogeneity is only one component of a wider investigation of variability across studies, the most important being diversity in clinical and methodological aspects. Meta-analysts must also

Higgins J, Thompson SG, Deeks JJ, Altman DG. Measuring inconsis- tency in meta-analyses. BMJ. 2003;327:557–560. https://doi.org/10.1136/bmj.327.7414.557

Meta-analisis: el debate

Fuente: Imagen facilitada por un oncólogo colega, o un colega oncólogo….

Conclusiones desde el punto de vista estadístico

Estudios de prevención de CVOT en pacientes diabéticos tipo II

De la duda de un perjuicio a establecer un beneficio (adicional)

Motivo (pre post comercialización) infiere en el diseño, hipótesis y tamaño de la muestra

… y en la extrapolación de resultados

Subgrupos o el arte de publicar

Valorar la homogeneidad entre factores pronósticos importantes

Generar hipótesis de trabajo (no para concluir evidencias)

Metaanálisis o el arte de agrupar

Comparaciones indirectas entre diseños diferentes en poblaciones distintas

Conclusiones desde el punto de vista clínico

- Existen diferencias relevantes entre los CVOT:
  - Diseño
  - Poblaciones estudiadas
  - Aplicabilidad de los resultados…..

- AR de GLP1 con beneficio en MACE-3 demostrado: Liraglutida, semaglutida (test de superioridad no estaba pre-definido) y dulaglutida.

- Otros estudios con arGLP1 se diseñaron para demostrar no inferioridad vs. placebo, sin embargo el REWIND testó la hipótesis de superioridad de dulaglutida vs. Placebo

- Dulaglutida es el único arGLP-1 que ha demostrado una reducción de eventos cardiovasculares graves (MACE-3*) de forma consistente en pacientes DM2 con y sin enfermedad cardiovascular establecida

- La población del REWIND representa en mayor medida a los pacientes de la práctica clínica real

Gracias!

Preguntas?

Espero que nos volvamos a ver

José Ríos Guillermo
jose.rios @uab.cat
Jose.rios@idibaps.org