

Artículo original

Uso del enfoque *win ratio* para evaluar los resultados del ensayo clínico DapaTAVI

Xavier Rossello^{a,b,c,◇,*}, Rafael Gonzalez-Manzanares^{d,e,f,◇}, Ignacio Amat-Santos^{f,g}, Vicente Peral Disdier^{a,b,c}, Luis Nieto Roca^{a,b,c}, Diego López Otero^{h,i,j}, Luis Nombela Franco^k, Livia Gheorge^l, Jorge Sanz-Sánchez^m, Javier Gómez Herrero^{f,g}, Rocío González Ferreiro^{j,n}, Antonio Jesús Muñoz García^o, Victoria Vilalta^p, Soledad Ojeda^{d,e,q}, Gabriela Veiga Fernández^{r,s}, Juan Gabriel Córdoba Soriano^t, Ander Regueiro^u, Miriam Sandín Rollán^v, Xacobe Flores Ríos^w, Aitor Urizarri^x, Roberto Martín Reyes^y, Rafael Romaguera^z, Pablo Avanzas^{f,aa,ab,ac}, Sergio García Blas^{f,ad,ae}, Juan A. Franco-Peláez^{af}, Javier Martín Moreiras^{ag}, José Ramón González Juanatey^{f,i,ah,ai}, Gabriela Tirado^k, Germán Calle^l, José Luis Díez^m, Sandra Santos-Martínez^{f,g}, María Melendo Viu^{j,n}, Xavier Carrillo Suarez^p, Xoan Sanmartín^{h,i,r}, Nieves Gonzalo^k, Alejandro Gutiérrez Barrios^l, Inmaculada González Bermúdez^{j,n}, Carlos Real^{c,k}, Valentín Fuster^{c,aj}, Borja Ibáñez^{c,f,af} y Sergio Raposeiras-Roubín^{c,j,n,ai,*}

^a Unidad de Cardiología, Institut d'Investigació Sanitària Illes Balears (IdISBa), Hospital Universitari Son Espases, Palma de Mallorca, Balears, España

^b Facultad de Medicina, Universitat de les Illes Balears (UIB), Palma de Mallorca, Balears, España

^c Departamento de Investigación Clínica, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, España

^d Unidad de Cardiología, Hospital Universitario Reina Sofía, Córdoba, España

^e Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), Córdoba, España

^f Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Instituto de Salud Carlos III, Madrid, España

^g Unidad de Cardiología, Hospital Clínico Universitario de Valladolid, Valladolid, España

^h Unidad de Cardiología, Hospital Universitario Montecelo, Pontevedra, España

ⁱ Unidad de Cardiología, Hospital Clínico Universitario de Santiago, Santiago de Compostela, A Coruña, España

^j Instituto de Investigación Sanitaria Galicia Sur, Vigo, Pontevedra, España

^k Unidad de Cardiología, Hospital Clínico Universitario San Carlos, Madrid, España

^l Unidad de Cardiología, Hospital Puerta del Mar, Cádiz, España

^m Unidad de Cardiología, Hospital Universitario La Fe, Valencia, España

ⁿ Unidad de Cardiología, Hospital Universitario Álvaro Cunqueiro, Vigo, Pontevedra, España

^o Unidad de Cardiología, Hospital Universitario Virgen de la Victoria, Málaga, España

^p Unidad de Cardiología, Hospital Universitario German Trias i Pujol, Badalona, Barcelona, España

^q Departamento de Ciencias Médicas y Quirúrgicas, Universidad de Córdoba, Córdoba, España

^r Unidad de Cardiología, Hospital Universitario Marqués de Valdecilla, Santander, España

^s Instituto de Investigación Valdecilla (IDIVAL), Santander, España

^t Unidad de Cardiología, Hospital Universitario de Albacete, Albacete, España

^u Unidad de Cardiología, Hospital Clínic, Barcelona, España

^v Unidad de Cardiología, Hospital General Universitario de Alicante, Alicante, España

^w Unidad de Cardiología, Hospital Universitario Juan Canalejo, A Coruña, España

^x Unidad de Cardiología, Hospital Universitario Vall d'Hebron, Barcelona, España

^y Unidad de Cardiología, Hospital Universitario QuirónSalud La Luz, Madrid, España

^z Unidad de Cardiología, Hospital Universitario de Bellvitge, L'Hospitalet de Llobregat, Barcelona, España

^{aa} Unidad de Cardiología, Hospital Universitario Central de Asturias, Oviedo, España

^{ab} Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, España

^{ac} Departamento de Medicina, Universidad de Oviedo, Oviedo, España

^{ad} Servicio de Cardiología, Hospital Clínico Universitario de Valencia, Valencia, España

^{ae} Instituto de Investigación Sanitaria INCLIVA, Valencia, España

^{af} Unidad de Cardiología, Hospital Universitario Fundación Jiménez Díaz, Madrid, España

^{ag} Unidad de Cardiología, Hospital Clínico Universitario de Salamanca, Salamanca, España

^{ah} Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, A Coruña, España

^{ai} Facultad de Medicina, Universidad de Santiago de Compostela, Santiago de Compostela, A Coruña, España

^{aj} Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, Nueva York, Estados Unidos

Historia del artículo:

Recibido el 29 de mayo de 2025

Aceptado el 12 de agosto de 2025

On-line el 17 de octubre de 2025

RESUMEN

Introducción y objetivos: El enfoque *win ratio* (WR) se utiliza para evaluar episodios compuestos de forma jerárquica. Esta novedosa metodología ofrece una excelente oportunidad para valorar la solidez de los hallazgos obtenidos en ensayos clínicos clave, como el estudio DapaTAVI.

* Autores para correspondencia.

Correos electrónicos: x.rossello@ucl.ac.uk (X. Rossello); raposeiras26@hotmail.com (S. Raposeiras-Roubín).

◇ Ambos autores contribuyeron por igual a este trabajo.

<https://doi.org/10.1016/j.recesp.2025.08.006>

0300-8932/© 2025 Sociedad Española de Cardiología. Publicado por Elsevier España, S.L.U. Se reservan todos los derechos, incluidos los de minería de texto y datos, entrenamiento de IA y tecnologías similares.

Palabras clave:

Implante percutáneo de válvula aórtica
 Inhibidores del cotransportador sodio-glucosa tipo 2
 Método win ratio

Métodos: Se aplicó el método del WR para evaluar el efecto del tratamiento con dapagliflozina sobre episodios clínicos ordenados jerárquicamente. Se analizaron diversas combinaciones de episodios, incluidas las variables de tiempo hasta el episodio, binarias y continuas.

Resultados: La WR del episodio principal original fue 1,36 (IC95%, 1,03-1,78; $p = 0,028$), comparable al recíproco de la *hazard ratio* (HR) original ($1/HR = 1,38$; IC95%, 1,06-1,81). La diferencia de *wins* (victorias) fue del 4,84% (IC95%, 0,55-9,12), lo que confirma también la consistencia de los hallazgos en términos de efecto absoluto. Las combinaciones alternativas del episodio principal, priorizando de forma distinta sus componentes, mostraron efectos del tratamiento y significación estadística similares. Ignorar el tiempo hasta el episodio e incluir episodios recurrentes no modificó sustancialmente la eficacia del tratamiento ni su significación estadística. En cambio, la inclusión en la jerarquía del total de días de hospitalización por insuficiencia cardiaca desplazó la estimación hacia el valor nulo. La incorporación de la clase funcional de la *New York Heart Association* mejoró la precisión de la estimación (WR = 1,31; IC95%, 1,09-1,56; $p = 0,003$). Por el contrario, incluir la calidad de vida mediante comparaciones del *Kansas City Cardiomyopathy Questionnaire* desplazó la estimación general hacia el valor nulo (WR = 1,10; IC95%, 0,94-1,30; $p = 0,236$).

Conclusiones: El WR es un método sólido para evaluar la eficacia de un tratamiento. Se encontraron resultados consistentes aplicando este enfoque en el ensayo DapaTAVI.

© 2025 Sociedad Española de Cardiología. Publicado por Elsevier España, S.L.U. Se reservan todos los derechos, incluidos los de minería de texto y datos, entrenamiento de IA y tecnologías similares.

Use of the win ratio approach to assess outcomes in the DapaTAVI trial

ABSTRACT

Introduction and objectives: The win ratio (WR) approach is used to assess composite endpoints in a hierarchical fashion. This novel method offers an excellent opportunity to assess the robustness of the findings yielded by landmark trials, such as the DapaTAVI trial.

Methods: We applied the WR method to evaluate the treatment effect of dapagliflozin in hierarchically ordered clinical outcomes. Several combinations of outcomes were tested, including time-to-event, binary, and continuous endpoints.

Results: The WR of the original primary endpoint was 1.36 (95%CI, 1.03-1.78; $P = .028$), comparable to the reciprocal of the original hazard ratio ($1/HR$, 1.38; 95%CI, 1.06-1.81). The win difference was 4.84% (95%CI, 0.55-9.12), confirming consistent findings in terms of absolute effect. Alternative combinations of the primary outcome with different prioritization of its components yielded similar treatment effects and statistical significance. Ignoring a time-to-event approach and including recurrent events did not substantially affect treatment efficacy and its statistical significance. In contrast, the inclusion of the total length of stay for heart failure hospitalizations in the hierarchy shifted the point estimate toward the null. Including New York Heart Association functional class improved the precision of the estimate (WR = 1.31; 95%CI, 1.09-1.56; $P = .003$). Conversely, including quality of life through Kansas City Cardiomyopathy Questionnaire comparisons shifted the overall estimate toward the null (WR = 1.10; 95%CI, 0.94-1.30; $P = .236$).

Conclusions: The WR approach is a solid method to assess treatment efficacy. We observed consistent findings using this approach in the DapaTAVI trial.

© 2025 Sociedad Española de Cardiología. Published by Elsevier España, S.L.U. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Keywords:

Transcatheter aortic valve replacement
 Sodium-glucose transporter 2 inhibitors
 Win ratio

Abreviaturas

IC: insuficiencia cardiaca
 KCCQ: *Kansas City Cardiomyopathy Questionnaire*
 NYHA: *New York Heart Association*
 TAVI: implante percutáneo de válvula aórtica
 WR: *win ratio* (cociente de victorias)

INTRODUCCIÓN

En muchos ensayos clínicos aleatorizados, se utilizan criterios de valoración combinados para reducir el tamaño de la muestra¹, pero la importancia relativa atribuida a cada criterio de valoración individual de la combinación tiende a diferir entre pacientes e

investigadores². Desde su presentación en 1992, los creadores del criterio de valoración cardiovascular combinado para los ensayos de cardiología sugirieron que los episodios se ponderaran de forma diferente o se evaluaran en un modelo jerárquico, de modo que se reflejara su importancia relativa³. Sin embargo, ninguno de estos 2 enfoques ha caracterizado los criterios de valoración compuestos utilizados en los ensayos clínicos aleatorizados cardiovasculares. En 2012, se describió el *win ratio* (WR) como un enfoque innovador para analizar criterios de valoración combinados utilizando una jerarquía que prioriza los componentes con mayor valor clínico^{4,5}. Este método también puede incorporar episodios repetidos (por ejemplo, hospitalizaciones)⁶ y resultados cuantitativos (como puntuaciones de calidad de vida)^{7,8}. El uso creciente del WR en ensayos cardiovasculares^{8,9} ha conducido a una necesidad cada vez mayor de comprender mejor su fundamento^{10,11}.

El ensayo DapaTAVI (*Dapagliflozin in patients undergoing Transcatheter Aortic Valve Implantation*) ha puesto de manifiesto recientemente que la dapagliflozina es beneficiosa en pacientes con estenosis aórtica tras un TAVI^{12,13}. El criterio de valoración

principal de muerte por cualquier causa o agravamiento de la insuficiencia cardiaca (AIC) se produjo en 91 pacientes (15,0%) del grupo de dapagliflozina y en 124 (20,1%) del de tratamiento estándar (*hazard ratio* [HR] = 0,72; intervalo de confianza del 95% [IC95%], 0,55-0,95)¹². Dada la significación estadística limítrofe ($p = 0,02$) y la fragilidad de los resultados (índice de fragilidad de 4)^{14,15}, se hace necesario evaluar su robustez con el empleo de metodologías alternativas.

El objetivo de este trabajo fue aplicar el enfoque WR en diversas formas en el ensayo DapaTAVI con objeto de: a) evaluar la robustez de los resultados; b) explorar el beneficio del tratamiento con dapagliflozina en diversas combinaciones de criterios de valoración, incluidos los de tiempo transcurrido hasta el episodio, episodios repetidos y variables continuas, y c) evaluar la jerarquía (orden) y la naturaleza de los componentes. Se ha realizado un reanálisis metodológico exploratorio *post hoc* del ensayo DapaTAVI.

MÉTODOS

Diseño del estudio y encuesta

El ensayo DapaTAVI fue un ensayo multicéntrico, aleatorizado, abierto, con evaluación ciega del resultado, controlado, iniciado por los investigadores, de la dapagliflozina en pacientes a los que se les practicó un TAVI. Se ha publicado ya una información detallada sobre el diseño del estudio (NCT04696185)¹³. El Comité de Ética de Investigación Clínica de Galicia, la Agencia Española de Medicamentos y Productos Sanitarios, y el comité de revisión ética de cada centro aprobaron el protocolo.

En este análisis *post hoc*, se evaluó el efecto del tratamiento con dapagliflozina mediante un enfoque WR utilizando varias combinaciones de criterios de valoración ordenados jerárquicamente. Paralelamente, se invitó a los investigadores principales de cada centro participante a responder a un cuestionario diseñado para captar sus conocimientos y percepciones preexistentes sobre la metodología WR. El cuestionario, elaborado por X. Rosello y R. Gonzalez-Manzanares, se distribuyó antes de hacer públicos los resultados de este análisis y tenía por objeto delimitar las lagunas existentes en el conocimiento.

Población e intervención del estudio

Se consideró aptos para la inclusión en el estudio a los pacientes con estenosis aórtica grave tratados con un TAVI que habían tenido un episodio previo de insuficiencia cardiaca (IC). Esto incluía cualquier hospitalización o visita de urgencia por IC que hubiera requerido la administración de diuréticos por vía intravenosa antes del TAVI. Además, los pacientes debían cumplir al menos 1 de los siguientes criterios de enriquecimiento de la muestra: filtración glomerular estimada (FGe) de 25-75 ml/min/1,73 m², diagnóstico de diabetes mellitus o fracción de eyección del ventrículo izquierdo $\leq 40\%$. Los criterios de exclusión incluyeron cualquier contraindicación de la dapagliflozina, el uso actual de una sulfonilurea u otro inhibidor del transportador de sodio-glucosa 2 (SGLT2), una presión arterial sistólica < 100 mmHg, una presión arterial diastólica < 50 mmHg, una FGe < 25 ml/min/1,73 m², o antecedentes de cistitis crónica o infecciones urinarias recurrentes (≥ 2 episodios en el año anterior).

Análisis win ratio

Se utilizó el enfoque WR para volver a reanalizar el criterio de valoración principal inicial del ensayo DapaTAVI y otras combinaciones adicionales de criterios de valoración clínicos ordenados de forma jerarquizada, a saber: 1: muerte por cualquier causa,

hospitalización por IC, visita de urgencia por IC; 2: muerte por cualquier causa, AIC; 3: AIC, muerte por cualquier causa; 4: muerte por cualquier causa, visita de urgencia por IC, hospitalización por IC; 5: hospitalización por IC, visita de urgencia por IC, muerte por cualquier causa; 6: visita de urgencia por IC, hospitalización por IC, muerte por cualquier causa; 7: muerte por cualquier causa, número total de episodios de AIC; 8: muerte por cualquier causa, hospitalización por IC, visita de urgencia por IC, número total de episodios de AIC; 9: muerte por cualquier causa, número total de episodios de AIC, hospitalización por IC, visita de urgencia por IC; 10: muerte por cualquier causa, hospitalización por IC, duración de la hospitalización por IC; 11: muerte por cualquier causa, número total de hospitalizaciones por IC, duración de la hospitalización por IC; 12: muerte por cualquier causa, hospitalización por IC, visita de urgencia por IC, puntuación del *Kansas City Cardiomyopathy Questionnaire* (KCCQ); 13: muerte por cualquier causa, puntuación del KCCQ; 14: puntuación del KCCQ; 15: muerte por cualquier causa, hospitalización por IC, visita de urgencia por IC, clase de la *New York Heart Association* (NYHA); 16: muerte por cualquier causa, clase de la NYHA; 17: clase de la NYHA, y 18: muerte por cualquier causa y número total de episodios de AIC en el mismo nivel de jerarquía (criterio combinado).

Para cada criterio de valoración, se compararon todos los pares de pacientes posibles de los 2 grupos de tratamiento según la jerarquía de criterios de valoración preespecificada. En cada par, el resultado se clasificó como «victoria», «derrota» o «empate» basándose en la frecuencia o la cronología de los episodios. La comparación comenzó con el criterio de valoración de mayor prioridad (por ejemplo, el tiempo transcurrido hasta la muerte). Si un paciente fallecía y el otro no, a este último se le consideraba el «ganador». Si ambos pacientes fallecían, al que sobrevivía más tiempo se le consideraba el «ganador». Si ambos pacientes fallecían al mismo tiempo o ninguno de los 2 presentaba el episodio, la comparación se consideraba un empate a ese nivel. En caso de empate en el primer criterio de valoración, se procedía a evaluar el siguiente criterio de valoración en la jerarquía (por ejemplo, la hospitalización por IC), y se repetía el proceso hasta que se determinaba una victoria, una derrota o un empate. Si en todos los criterios de valoración de la jerarquía había un empate o faltaban los datos, se excluía al par de pacientes del análisis. El WR se calculó mediante el número de victorias en el grupo de dapagliflozina dividido por el número de derrotas y el de derrotas⁹. Así pues, el WR puede interpretarse como la probabilidad de que un paciente seleccionado al azar del grupo de dapagliflozina tuviera un resultado más favorable que uno seleccionado al azar del grupo de control, en el conjunto de todos los pares en los que no hubo empate¹⁰. Un WR > 1 indica un efecto beneficioso. Como medida para cuantificar el efecto absoluto del tratamiento, se calculó la «diferencia de victorias» (diferencia entre el porcentaje de victorias y el de derrotas)¹⁶.

Para evaluar la posible heterogeneidad del efecto del tratamiento en los diversos subgrupos, se comparó el WR de los distintos subgrupos mediante un estadístico Z basado en la diferencia del WR transformado logarítmicamente y su correspondiente error estándar. El valor de p se obtuvo comparando el valor absoluto del estadístico Z con la distribución normal estándar¹⁰. Los análisis se realizaron con el programa informático R (versión 4.4.2, R Foundation for Statistical Computing, Austria).

RESULTADOS

Resultados de la encuesta

Un total de 28 investigadores completaron el cuestionario. Las respuestas a las 12 preguntas se resumen en la [figura S1](#). Menos de la mitad de los encuestados no preveían que se obtuvieran resultados incoherentes con el uso del WR en lugar de la regresión

de Cox (figura S1, Q1). Es de destacar que solamente el 25% de los participantes identificaron la fórmula correcta para calcular el WR (figura S1, Q3). La mayoría de los investigadores reconocieron que el método WR implica el análisis de criterios de valoración combinados (figura S1, Q2), pero solo el 57,1% identificaron correctamente que puede incorporar criterios de valoración de tiempo transcurrido hasta el episodio o de episodios recurrentes (figura S1, Q5 y Q6), y el 42,9% identificaron correctamente ambas características. Tras la exclusión de la Q1 debido a su formato de respuesta abierta subjetiva, ningún participante respondió correctamente a todas las preguntas restantes.

Reanálisis de los resultados principales

El criterio de valoración principal de muerte o AIC se dio en 91 de 605 pacientes (15,0%) del grupo de dapagliflozina y en 124 de 617 pacientes (20,1%) del grupo de tratamiento estándar. Para facilitar la comparación con el WR, se presenta un valor de HR de 0,72 (IC95%, 0,55-0,95; $p = 0,019$) y su recíproco ($1/HR = 1,38$; IC95%, 1,06-1,81)¹², que indica un mayor riesgo de episodios con el tratamiento estándar en comparación con el de dapagliflozina (figura 1)⁷. Se observaron resultados coherentes cuando se analizaron los 3 componentes del criterio de valoración principal

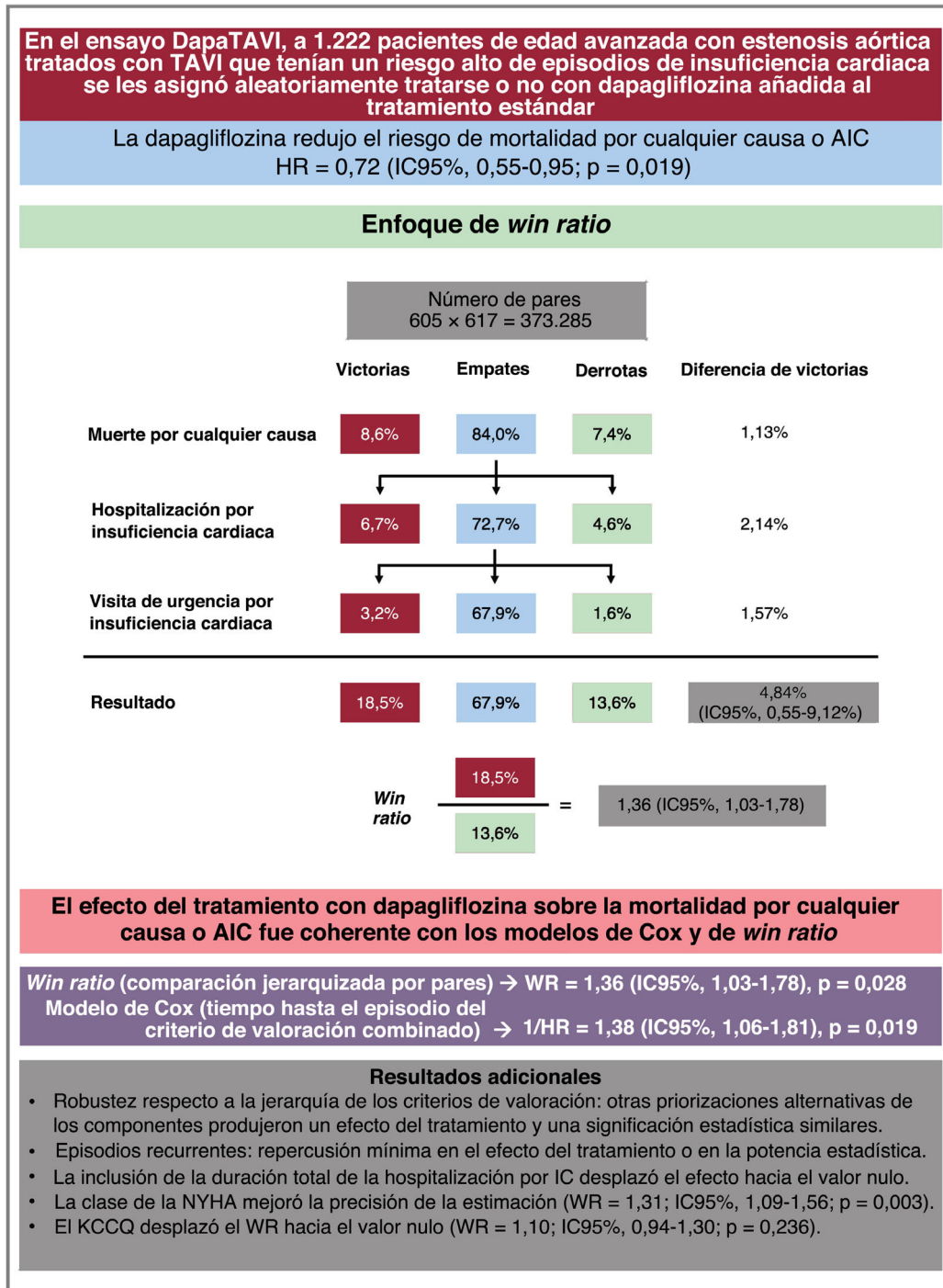


Figura 1. Figura central. Diseño del estudio, cálculo paso a paso del WR y principales conclusiones. AIC: agravamiento de la insuficiencia cardiaca; HR: hazard ratio; IC95%: intervalo de confianza del 95%; KCCQ: Kansas City Cardiomyopathy Questionnaire; NYHA: New York Heart Association; TAVI: implante percutáneo de válvula aórtica; WR: win ratio.

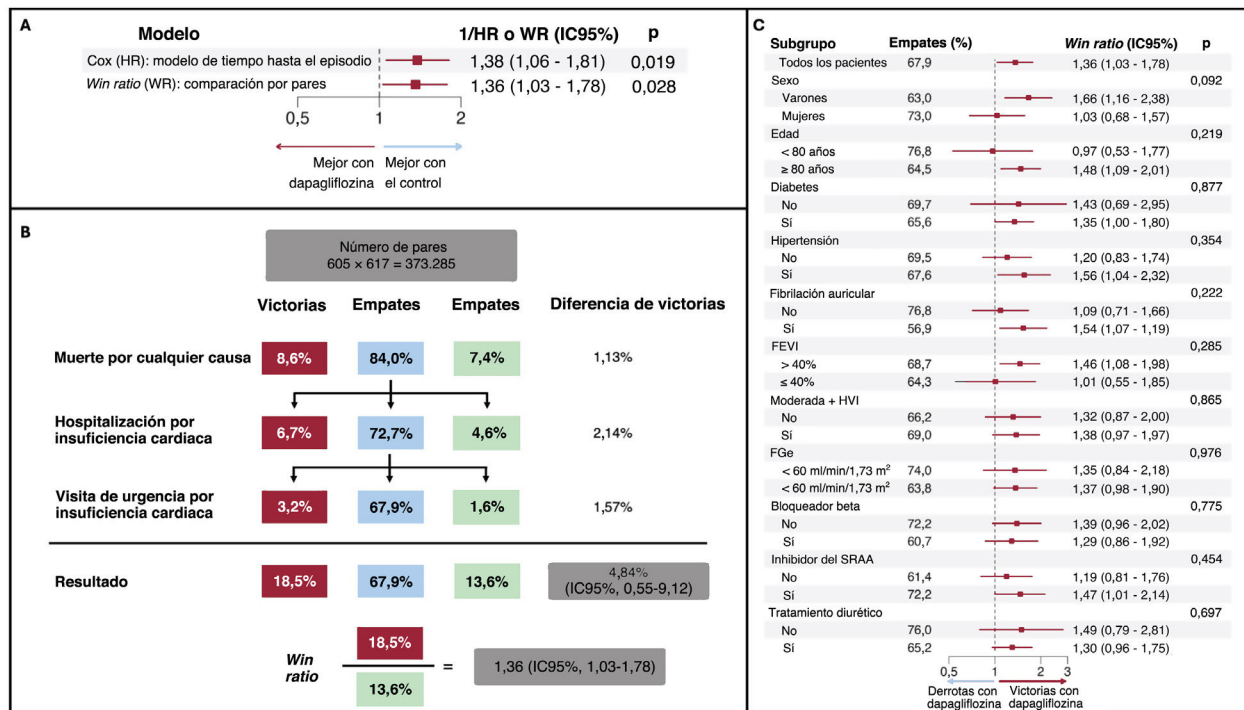


Figura 2. Reanálisis de los resultados principales y análisis de subgrupos. A: fundamento de 1/WR. B: cálculo del WR. C: análisis de subgrupos. FEVI: fracción de eyección del ventrículo izquierdo; FGe: filtración glomerular estimada; HR: hazard ratio; HVI: hipertrofia ventricular izquierda; IC95%: intervalo de confianza del 95%; SRAA: sistema renina-angiotensina-aldosterona; WR: win ratio.

utilizando el enfoque WR (WR = 1,36; IC95%, 1,03-1,78; p = 0,028) (figura 2 A). El porcentaje de victorias, derrotas y empates en cada nivel jerárquico se muestra en la figura 2 B. La diferencia de victorias fue del 4,84% (IC95%, 0,55-9,12), lo que confirma la coherencia de los resultados por lo que respecta al efecto absoluto. Al igual que ocurre en otros parámetros de medida absolutos del efecto del tratamiento, el valor nulo de la diferencia de victorias es 0.

No se observó ningún signo de heterogeneidad en la eficacia del tratamiento en los subgrupos preespecificados, incluidos los de pacientes con enfermedad renal crónica, diabetes y disfunción sistólica ventricular izquierda (figura 2 C).

Importancia de la jerarquía de criterios de valoración

Para evaluar con mayor detalle la repercusión de la selección de componentes y del orden jerárquico en la especificación del modelo, se exploraron combinaciones alternativas del criterio de valoración

principal modificando tanto la estructura como la priorización de sus componentes (figura 3). En primer lugar, la hospitalización por IC y las visitas de urgencia por IC se combinaron en una sola variable, y con ello se obtuvo un WR comparable al observado con el modelo original de 3 criterios de valoración. Posteriormente, se modificó el orden jerárquico de los criterios de valoración finales tanto en el modelo de 2 criterios como en el de 3 criterios. Tiene interés señalar que se observaron estimaciones del efecto del tratamiento progresivamente más altas a medida que se priorizaban criterios de valoración más blandos al inicio de la jerarquía.

Alternativas a los enfoques basados exclusivamente en métodos de tiempo hasta el episodio

Para evaluar la repercusión que tenía la incorporación de variables de tiempo transcurrido hasta el episodio y de episodios

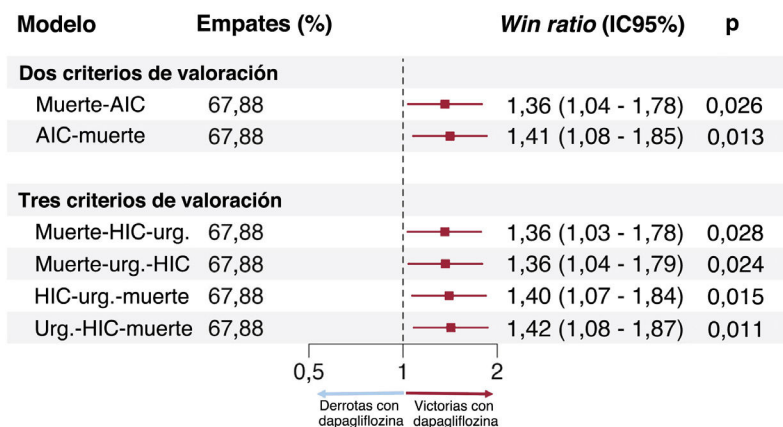


Figura 3. Importancia de la jerarquía de los criterios de valoración: 6 posibles combinaciones. AIC: agravamiento de la insuficiencia cardiaca; HIC: hospitalización por insuficiencia cardiaca; IC95%: intervalo de confianza del 95%; urg.: visita de urgencia por insuficiencia cardiaca.

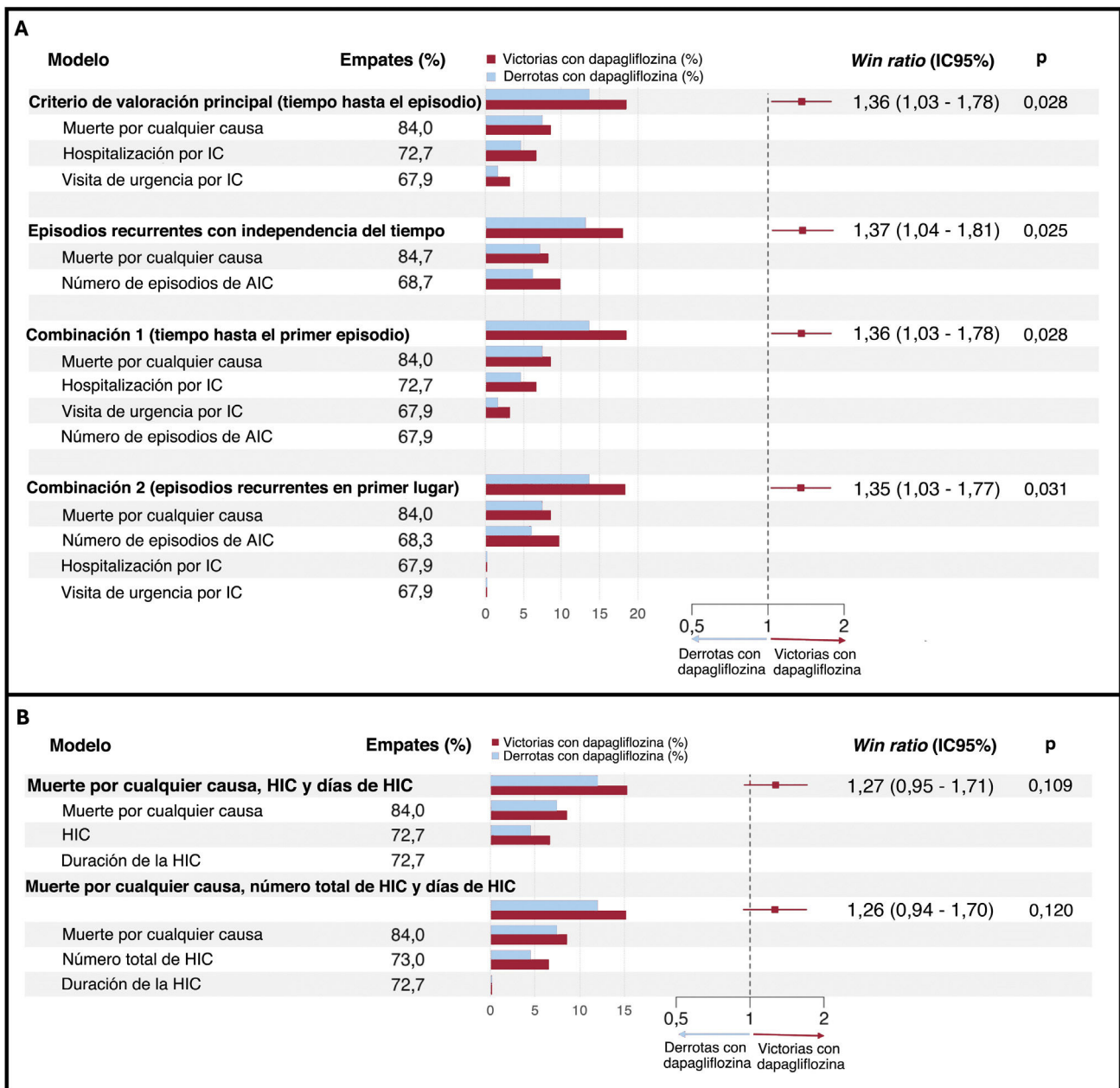


Figura 4. Alternativa a los criterios de valoración de tiempo hasta el episodio. A: episodios recurrentes. B: duración de la hospitalización. AIC: agravamiento de la insuficiencia cardiaca; HIC: hospitalización por insuficiencia cardiaca; IC: insuficiencia cardiaca; IC95%: intervalo de confianza del 95%.

recurrentes en las estimaciones, se realizaron modelos adicionales de WR (figura 4). En primer lugar, se incluyó un modelo simplificado de 2 criterios de valoración que no tenía en cuenta el tiempo transcurrido hasta el episodio e incluía el AIC mediante el número total de episodios. El WR resultante (1,37; IC95%, 1,04-1,81) fue similar al del modelo de tiempo transcurrido hasta el episodio de 3 niveles. Cuando se añadió el número total de episodios de AIC a la jerarquía como criterio de valoración de nivel más bajo, no se observaron victorias o derrotas adicionales (figura 4 A, combinación 1). Por el contrario, cuando los episodios recurrentes se incorporaron antes en la jerarquía (antes de la hospitalización o de las visitas de urgencia), los episodios recurrentes dieron lugar a un desempate en un 15,7% de los pares de pacientes, a la vez que neutralizaban sustancialmente la contribución del tiempo hasta la hospitalización (el 0,4% adicional de los pares desempatados) y del tiempo hasta las visitas de urgencia (< 0,1% adicional de los pares desempatados) (figura 4 A, combinación 2).

En el criterio de valoración combinado formado por el tiempo hasta la muerte por cualquier causa, la hospitalización por IC y la visita de urgencia por IC, el reemplazo de la visita de urgencia por la duración total de la hospitalización redujo la eficiencia del modelo. Esto se debió principalmente al aumento de la proporción de empates (que pasó del 67,9 al 72,7%), que desplazó la estimación puntual del WR hacia el valor nulo y amplió el intervalo de confianza (figura 4 B).

Calidad de vida y clase de la NYHA

Se exploró la repercusión que tenía la incorporación del estado funcional, evaluado mediante la clase de la NYHA, y la calidad de vida referida por el paciente, evaluada mediante la puntuación KCCQ, en los modelos jerárquicos. Ambas variables redujeron de manera apreciable el porcentaje de empates entre pares de

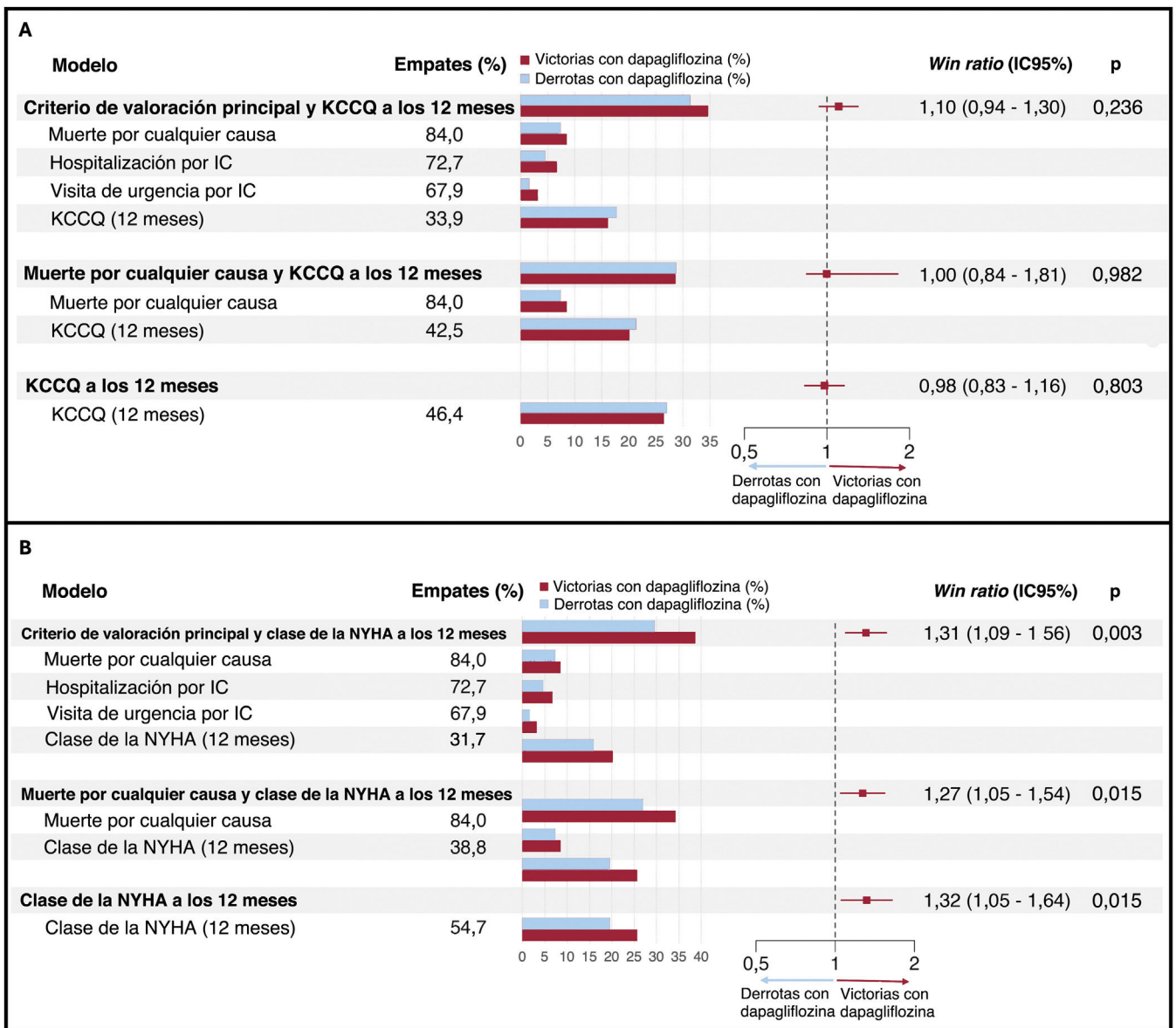


Figura 5. Inclusión de criterios de valoración blandos. A: calidad de vida, y B: clase de la NYHA. IC: insuficiencia cardiaca; IC95%: intervalo de confianza del 95%; KCCQ: Kansas City Cardiomyopathy Questionnaire; NYHA: New York Heart Association.

pacientes (figura 5). La inclusión de la clase funcional de la NYHA dio lugar a un mayor número de victorias a favor del grupo de dapagliflozina, lo que mejoró la precisión de la estimación del WR (WR = 1,31; IC95%, 1,09-1,56; p = 0,003). En cambio, las comparaciones del KCCQ dieron lugar a una distribución más equilibrada de las victorias entre los grupos, desplazando la estimación global hacia el valor nulo (WR = 1,10; IC95%, 0,94-1,30; p = 0,236). En consonancia con estos resultados, la dapagliflozina mostró un efecto beneficioso cuando se analizó por separado la clase de la NYHA (p = 0,015), pero no hubo un efecto beneficioso significativo con las puntuaciones del KCCQ (p = 0,803). Cuando se incorporaron a la jerarquía tanto el KCCQ como la clase de la NYHA junto con la muerte, el WR aumentó ligeramente en comparación con los análisis que trataban cada criterio de valoración por separado. Esto pone de relieve la posible utilidad del enfoque WR para abordar adecuadamente el truncamiento por muerte, en el que es característico que los datos no disponibles sean más desfavorables. La inclusión del KCCQ mediante el cambio respecto al valor inicial (en lugar del valor final) arrojó prácticamente las mismas estimaciones, tanto al usarlo como criterio de valoración continuo

(WR = 1,10; IC95%, 0,95-1,28; p = 0,207) como al emplear un margen de 5 puntos (WR = 1,11; IC95%, 0,95-1,30; p = 0,201).

Criterios de valoración secundarios

Se evaluó la coherencia de las estimaciones del WR para los criterios de valoración secundarios clave especificados *a priori*. Como en el caso del criterio de valoración principal, en la figura 6 se presentan gráficamente los recíprocos de los valores de *subhazard ratio* y los cocientes de tasas, junto con el WR correspondiente. Los resultados fueron en la misma dirección que los resultados principales publicados, de tal manera que el WR mostraba una ausencia de efecto del tratamiento sobre la muerte de causa cardiovascular y un efecto beneficioso tanto en el tiempo hasta la hospitalización por IC o muerte cardiovascular como en el número total de hospitalizaciones por IC o de muertes de causa cardiovascular. Aunque las estimaciones del WR fueron coherentes con las de los modelos convencionales (riesgo competidor y regresión binomial negativa), la magnitud del efecto se atenuó, lo

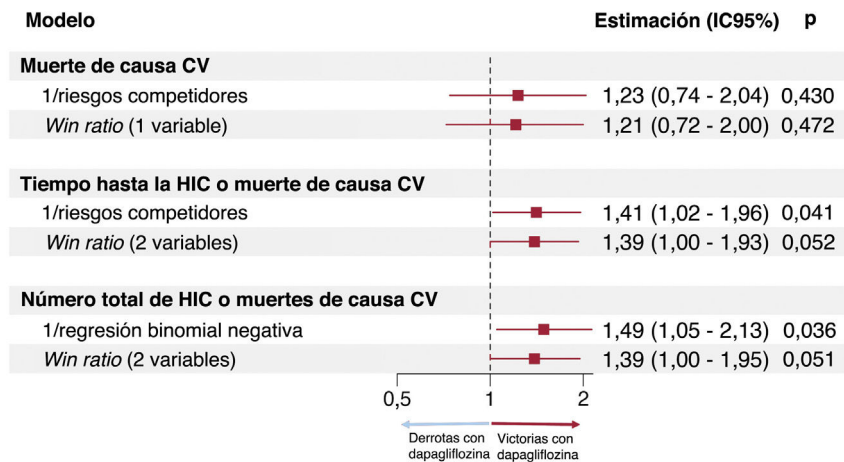


Figura 6. Criterios de valoración secundarios clave. CV: cardiovascular; HIC: hospitalización por insuficiencia cardiaca; IC95%: intervalo de confianza del 95%.

que dio lugar a valores de p que superaban ligeramente el umbral de valor p convencional de 0,05.

DISCUSIÓN

Con el empleo de la metodología WR en el ensayo DapaTAVI¹³, un ensayo controlado aleatorizado, pragmático, abierto, con evaluación ciega de los criterios de valoración, se ha confirmado la solidez de los resultados de dicho estudio. Además, se ha demostrado que el orden de los diversos componentes del criterio de valoración principal tuvo una repercusión mínima en la estimación y los resultados de las pruebas de hipótesis, y que la adición de la IC recurrente no tuvo ningún valor aditivo fuera cual fuera su posición en el orden de la jerarquía. Mientras que la inclusión de la clase funcional de la NYHA mejoró la precisión y significación de la estimación, la inclusión de la calidad de vida evaluada mediante comparaciones del KCCQ desplazó la estimación global hacia el valor nulo. En general, el enfoque WR proporciona unos resultados sólidos en diversos criterios de valoración siempre que la selección de estos sea razonable. En consonancia con los resultados obtenidos en la encuesta interna realizada entre los investigadores del ensayo DapaTAVI, se espera clarificar que los resultados de este ensayo son sólidos y coherentes independientemente de la metodología utilizada. Del mismo modo, se espera que los lectores consideren el enfoque WR una herramienta fiable y comprendan mejor su interpretación cuando utilicen diferentes criterios de valoración.

En los pacientes de edad avanzada con estenosis aórtica y antecedentes de IC a los que se les practica un TAVI, se ha demostrado que la dapagliflozina reduce la incidencia de muerte o AIC¹². En este artículo, se han presentado unos resultados coincidentes en términos relativos y absolutos utilizando un enfoque WR en lugar de un enfoque tradicional de tiempo hasta el primer episodio¹⁶. En este sentido, la estimación del WR fue similar al recíproco de la HR (1,36 frente a 1,38), con prácticamente la misma significación estadística. Del mismo modo, la diferencia de victorias fue del 4,8%, mientras que la diferencia en los criterios de valoración principales fue del 5,1%. Esto conduce a 2 conclusiones principales: Tanto el ensayo DapaTAVI como el método del WR son sólidos. Algunos autores sugieren que se informe también de las probabilidades de victoria, un parámetro de medida relacionado que asigna la mitad de los empates a victorias de cada grupo, lo que proporciona una estimación más conservadora¹⁷.

Aunque tanto el enfoque de tiempo transcurrido hasta el episodio como el de WR utilizan criterios de valoración combi-

nados, el método de WR tiene 2 ventajas principales: a) da prioridad a los criterios de valoración de mayor importancia clínica que forman parte de la combinación, y b) aumenta la flexibilidad en el uso de distintos tipos de variables de valoración.

Por lo que respecta a la priorización de los criterios de valoración de mayor importancia clínica que forman parte de la combinación, debe señalarse que cada victoria tiene el mismo peso en el recuento total de victorias, sea cual sea el criterio de valoración en el que se produzca. Así, la contribución de cada criterio de valoración a la estimación global viene determinada no solo por su prioridad clínica, sino también por su frecuencia. En el ensayo DapaTAVI, el criterio de valoración de mayor prioridad, que fue la muerte por cualquier causa, contribuyó a una mayor proporción de desempates (16,0%) que la hospitalización por IC (11,3%) o la visita de urgencia por IC (4,8%). Esta característica fue fundamental para obtener resultados coherentes. Si el criterio de valoración de menor prioridad fuera más frecuente que el episodio de la máxima prioridad, el WR habría sido menos fácil de predecir y muy probablemente mayor que el recíproco del valor de HR. La conclusión es que el WR podría ser más eficiente para mostrar la eficacia si se prevé una repercusión en los criterios de valoración más blandos, pero no hay diferencia si la mayor parte de la eficacia se concentra en el criterio de valoración más duro de todos. La presentación de las victorias, derrotas y empates específicos de cada nivel, e incluso de las estimaciones de nivel, aumenta la transparencia y facilita la interpretación de la trascendencia clínica de la estimación global del WR. En este sentido, se observa un beneficio con la dapagliflozina en los 3 niveles de criterios de valoración, aunque fue más pronunciado para los criterios de valoración de IC. Hasta cierto punto, esto podía esperarse, ya que los iSGLT2 han mostrado una repercusión ligeramente superior en los criterios de valoración de IC en comparación con la que tienen en la mortalidad en todo el espectro de los ensayos cardiovasculares^{18–21}. Junto con la posibilidad de priorizar los criterios de valoración, la flexibilidad para incluir distintos tipos de variables es la otra ventaja del WR. Esta flexibilidad se traduce en la posibilidad de incluir episodios repetidos (por ejemplo, el número de hospitalizaciones por IC), parámetros funcionales (como la clase de la NYHA) y medidas de resultados percibidos por los pacientes (RPP).

Los episodios repetidos pueden incorporarse fácilmente al WR, al tiempo que proporcionan un enfoque más fiable y no requieren de supuestos previos, en comparación con otros métodos. Sin embargo, es posible que no mejore necesariamente la potencia estadística en comparación con el uso exclusivo del primer episodio. En el ensayo DapaTAVI, el análisis WR de un criterio

de valoración jerárquico que incluía la muerte y el número total de episodios de AIC proporcionó una estimación coherente pero ligeramente menos eficiente que la obtenida a partir del análisis principal de tiempo transcurrido hasta el episodio (por ejemplo, el intervalo de confianza más amplio y el valor p más alto en la comparación entre modelos). Lo más probable es que esto se debiera a una mayor proporción de empates en el modelo de episodios repetidos en comparación con el análisis del tiempo transcurrido hasta el episodio (el 68,7 frente al 67,9%). Cuando se combinaron en un único modelo jerárquico los componentes de tiempo transcurrido hasta el episodio y de episodios repetidos, la adición de estos últimos como criterio de valoración de menor prioridad no dio lugar a derrotas ni victorias adicionales, ya que la probabilidad de que 2 pacientes experimentaran episodios de AIC exactamente en el mismo tiempo de seguimiento era extremadamente baja.

No obstante, la futilidad de utilizar episodios repetidos en este estudio exige una interpretación matizada. En primer lugar, por coherencia con el análisis del criterio de valoración principal, la unidad de tiempo para los criterios de valoración de tiempo fueron los días hasta el episodio. Si el tiempo se hubiera medido en unidades más amplias (como los meses), los episodios recurrentes podrían haber contribuido a generar victorias o derrotas adicionales. En segundo lugar, el número total de episodios repetidos de IC fue relativamente bajo en el ensayo DapaTAVI. En los ensayos en los que se prevé un número mayor de episodios recurrentes (por ejemplo, IC crónica) o un seguimiento más largo, la incorporación de criterios de valoración de episodios repetidos en un análisis de WR puede dar lugar o no a un menor número de empates y a una mayor repercusión en la estimación global. Un nuevo análisis del ensayo CHARM-preserved utilizando los episodios repetidos (en lugar del tiempo transcurrido hasta el primer episodio) mostró una mayor eficiencia con el modelo Lin Wei Yang Ying y con el modelo binomial negativo, pero no con el método WR. Esto se debió probablemente al efecto limitado del candesartán sobre el criterio de valoración de mayor prioridad (muerte de causa cardiovascular), lo que limitó la capacidad de los episodios recurrentes de influir en la estimación global⁶. En el ensayo ATTR-ACT, los episodios repetidos sí marcaron la diferencia. El WR para el tafamidis, en un criterio de valoración jerárquico de muerte por cualquier causa y número de hospitalizaciones cardiovasculares, mostró un efecto del tratamiento mayor que el de otros métodos más convencionales^{8,22}. Los episodios recurrentes tienden a marcar la diferencia en los trastornos crónicos como la IC. Sin embargo, en un análisis *post hoc* de los ensayos EMPEROR-Reduced y EMPEROR-Preserved, el uso de los episodios recurrentes redujo la potencia estadística en comparación con el empleo del tiempo transcurrido hasta la primera hospitalización. Pocock et al.¹⁰ plantearon la hipótesis de que esto pudiera deberse a un mayor efecto del tratamiento en una fase temprana o a la agrupación de episodios recurrentes en una pequeña cohorte de pacientes.

La adición de un criterio de valoración cuantitativo (por ejemplo, la puntuación de la calidad de vida) a la jerarquía puede ser útil en ensayos que carecen de suficiente potencia estadística para evaluar los efectos sobre los episodios clínicos por sí solos¹⁰. Aunque este enfoque podría ser atractivo para diseñar ensayos en enfermedades raras o poblaciones con un riesgo bajo de episodios, hay que tener en cuenta ciertas consideraciones. En primer lugar, la inclusión de un criterio de valoración cuantitativo no aumenta necesariamente la eficacia del método WR. Si no hay plausibilidad biológica de que el criterio de valoración se modifique con la intervención, se prevé que el WR total se desplace hacia el valor nulo. En este análisis, la adición de la puntuación KCCQ a los episodios clínicos redujo la eficiencia del modelo, diluyendo cualquier posible efecto observado sobre estos. Aunque la adición de este criterio de valoración era biológicamente plausible, dado el

impacto conocido de los iSGLT2 sobre el KCCQ en los pacientes con IC, se tiene la impresión de que este efecto se vio atenuado debido a la mejora sustancial tras el TAVI en ambos grupos de tratamiento (figura S2), es decir, no había margen de mejora más allá de la proporcionada por el TAVI inicial que se aplicó a todos los pacientes²³. Por el contrario, la adición de la clase NYHA a la jerarquía, aunque solo redujo la proporción de empates en un 2,2% adicional en comparación con el KCCQ, mejoró notablemente la eficiencia del modelo (WR = 1,31; IC95%, 1,09–1,56; p = 0,003). Es importante destacar que este beneficio, que según la hipótesis manejada se debe al alivio de la congestión residual, fue coherente con los criterios de valoración de mayor prioridad. Este hecho lleva a una segunda consideración: la trascendencia clínica de cualquier criterio de valoración cuantitativo debe sopesarse cuidadosamente tanto en la fase de diseño como a la hora de interpretar los resultados. Si el efecto de la dapagliflozina sobre la clase de la NYHA hubiera sido neutro, la inclusión de este criterio de valoración habría reducido la potencia estadística al desplazar el WR hacia el valor nulo. Es de destacar que se trataba de un ensayo abierto, y que cualquier conclusión sobre los resultados de la NYHA (evaluada por el médico) y del KCCQ (evaluado por el paciente) podría verse afectada por un sesgo de observador. Una evaluación ciega de estos episodios habría reforzado los resultados.

Un uso válido del WR requiere una jerarquía de criterios de valoración que tenga sentido clínico, preferiblemente con una coherencia de efectos entre los componentes. Los usos incorrectos han consistido en una jerarquía mixta excesivamente elaborada, como la descrita en el ensayo DAPA-MI⁹. En este ensayo, la mayor parte del beneficio «cardiometabólico» de la dapagliflozina se debió a una mayor proporción de pacientes que alcanzaron una disminución de peso $\geq 5\%$, lo que redujo la proporción de empates en los 6 primeros componentes del criterio de valoración jerárquico, que pasó del 62,8 al 42,5%. Por el contrario, la proporción de empates en las comparaciones de la muerte y la hospitalización por IC fue del 93,7%, lo que supuso un WR de 1,01 (IC95%, 0,68–1,49). Desde una perspectiva clínica, los resultados se debieron principalmente a 2 factores «metabólicos» (diabetes y menor reducción del peso), mientras que el efecto de la dapagliflozina sobre los factores «cardíacos» fue insignificante²⁴. No está claro si un beneficio en estos 2 criterios de valoración metabólicos se traduciría en una mejora relevante en los criterios de valoración clínicos en los pacientes tras un IAM. En una visión conjunta, el uso de criterios de valoración cuantitativos es tal vez la característica que plantea más dificultades en el enfoque WR.

Limitaciones

Este estudio tiene varias limitaciones. En primer lugar, dado el carácter *post hoc* del estudio, los análisis solo pueden considerarse exploratorios. En favor del estudio, se utilizaron criterios de valoración preespecificados y evaluados de forma centralizada para mostrar las cuestiones metodológicas sin explorar nuevas vías por lo que respecta a la eficacia del tratamiento. En segundo lugar, no se realizó ningún ajuste formal para comparaciones múltiples a pesar de ensayar múltiples combinaciones de criterios de valoración²⁵. Así pues, en este estudio *post hoc*, los resultados de las combinaciones alternativas deben interpretarse como generadores de hipótesis. En tercer lugar, la encuesta pretendía poner de relieve las lagunas existentes en los conocimientos. No puede extraerse ninguna conclusión causal de esta pequeña encuesta, que probablemente no sea representativa. Por último, a pesar de la inclusión de criterios de valoración adicionales de nivel inferior, el número relativamente elevado de empates en algunos modelos probablemente contribuyó a que los valores de p fueran dudosos.

CONCLUSIONES

El WR es un método sólido y flexible para analizar criterios de valoración combinados y jerarquizados. En este análisis *post hoc* del ensayo DapaTAVI, se confirmó la uniformidad y robustez de los resultados originales, y se exploró el beneficio del tratamiento con dapagliflozina en varias combinaciones de criterios de valoración, incluido el tiempo transcurrido hasta el episodio, los episodios repetidos y las variables de valoración continuas. Es esencial considerar cuidadosamente la jerarquía y naturaleza de los componentes, su trascendencia clínica y la interpretabilidad de la estimación para optimizar la utilidad de este enfoque en futuros ensayos cardiovasculares.

¿QUÉ SE SABE DEL TEMA?

- El WR es un método novedoso para analizar criterios de valoración combinados, mediante la priorización de las variables de valoración clínicamente significativas.
- Permite la incorporación de diversos tipos de criterios de valoración, incluidas variables de tiempo transcurrido hasta el episodio, episodios recurrentes y variables continuas.
- Aunque su uso cada vez es más frecuente en los ensayos cardiovasculares, su interpretación y utilidad siguen sin comprenderse completamente por parte de muchos clínicos y autores de ensayos clínicos.
- El ensayo DapaTAVI demostró que la dapagliflozina redujo la muerte o el AIC en personas de edad avanzada con estenosis aórtica a los que se les practicó un TAVI y que presentaban un riesgo alto de episodios de IC.

¿QUÉ APORTA DE NUEVO?

- Este estudio confirma la robustez de los resultados del ensayo DapaTAVI utilizando el enfoque WR mediante múltiples combinaciones de criterios de valoración jerarquizados.
- Demuestra la flexibilidad del método para incorporar el tiempo transcurrido hasta el episodio, los episodios recurrentes y variables continuas como las puntuaciones de calidad de vida.
- El WR y la diferencia de victorias arrojaron estimaciones coherentes con la HR original, lo que refuerza el beneficio de la dapagliflozina.
- Estos resultados respaldan el WR como un enfoque estadístico fiable y ofrecen perspectivas prácticas para su aplicación en futuros ensayos cardiovasculares.

FINANCIACIÓN

Con el apoyo de una subvención concedida a X. Rossello (FIS PI24/01893 del Instituto de Salud Carlos III, Madrid, España) y de una subvención (FIS PI19/01882, concedida a S. Raposeiras-Roubín) del Instituto de Salud Carlos III; una subvención (GRS 2459/A/21, 2022, concedida a I. Amat-Santos) de la Gerencia Regional de Salud de la Junta de Castilla y León y Fondos; una subvención (SEC/FEC-INV-CLI 21/004, concedida a S. Raposeiras-Roubín) de la Sociedad Española de Cardiología; una subvención (SOGACAR 2022, concedida a S. Raposeiras-Roubín) de la Sociedad Española de

Cardiología de Galicia; una subvención (SGLT2-HYPE 101156555, concedida a B. Ibáñez) de la Comisión Europea y el Centro Nacional de Investigaciones Cardiovasculares (CNIC) de España, que cuenta con el apoyo del Instituto de Salud Carlos III, el Ministerio de Ciencia, Innovación y Universidades, la Fundación Pro CNIC, y una subvención (CEX2020-001041-S financiada por MICIU/AEI/10.13039/501100011033) del Centro de Excelencia Severo Ochoa. R. Gonzalez-Manzanares fue beneficiario de contratos de investigación (CM22/00259, JR24/00064) del Instituto de Salud Carlos III.

CONSIDERACIONES ÉTICAS

El Comité de Ética de Investigación Clínica de Galicia, la Agencia Española de Medicamentos y Productos Sanitarios y el comité de revisión ética de cada uno de los centros participantes aprobaron el protocolo. El ensayo se atuvo a los principios descritos en la Declaración de Helsinki y las directrices de Buenas Prácticas Clínicas del Consejo Internacional de Armonización, y todos los pacientes firmaron el consentimiento informado. No se aplicaron las directrices SAGER en este manuscrito metodológico; no se evaluaron variables relacionadas con el sexo o el género.

DECLARACIÓN SOBRE EL USO DE INTELIGENCIA ARTIFICIAL

Este trabajo ha sido realizado por los autores sin utilizar inteligencia artificial.

CONTRIBUCIÓN DE LOS AUTORES

X. Rossello y R. Gonzalez-Manzanares contribuyeron por igual a este trabajo. X. Rossello supervisó la conceptualización, la metodología, el análisis formal, la redacción de la propuesta inicial del manuscrito, y la revisión y corrección del manuscrito final. R. Gonzalez-Manzanares se encargó de la metodología, el análisis formal, la visualización, la redacción de la propuesta inicial del manuscrito, y la revisión y corrección del manuscrito final. B. Ibáñez se encargó de la investigación, la vigilancia, la administración del proyecto, la obtención de fondos y la edición de la revisión. S. Raposeiras-Roubín supervisó la investigación, la vigilancia, la administración del proyecto, la obtención de fondos y la edición de la revisión. Todos los demás autores contribuyeron al proceso de investigación y aportaron comentarios conceptuales. Todos los autores han aprobado la versión final del manuscrito antes de presentarlo para su publicación.

CONFLICTO DE INTERESES

P. Avanzas es editor asociado de *Revista Española de Cardiología*; se ha seguido el procedimiento editorial establecido por la revista para garantizar un tratamiento imparcial del manuscrito. X. Rossello ha formado parte del comité directivo de los ensayos EMPA-KIDNEY y EMPACT-MI a través de la CVCT Future Trialist Fellowship, pero no ha recibido remuneración alguna por estas tareas. R. Gonzalez-Manzanares ha recibido honorarios por conferencias de AstraZeneca. I. Amat-Santos declara haber recibido subvenciones o ser beneficiario de contratos de investigación de Boston Scientific. L. Nombela-Franco ha sido mentor de Abbott y Boston Scientific, y consultor de Edwards Lifesciences. S. García Blas ha participado en consejos asesores científicos de Boston Scientific y Medtronic, y ha recibido subvenciones o contratos de Edwards Lifesciences. A. Regueiro ha recibido pagos como mentor y asesor de Abbott, Edwards, iVascular y Meril. S. Raposeiras-Roubín ha recibido pagos u honorarios por conferencias y presentaciones

de Ferrer, Novo Nordisk, Daiichi Sankyo, Amarin y Boehringer Ingelheim, y apoyo para asistir a congresos de Daiichi Sankyo, Bayer y Amarin. Los demás autores no tienen conflictos de intereses que declarar.

ANEXO. MATERIAL ADICIONAL

Se puede consultar material adicional a este artículo en su versión electrónica disponible en <https://doi.org/10.1016/j.recesp.2025.08.006>.

BIBLIOGRAFÍA

- Gonzalez-Del-Hoyo M, Mas-Llado C, Siquier-Padilla J, et al. A systematic assessment of the characteristics of randomized controlled trials cited by acute coronary syndrome clinical practice guidelines. *Eur Heart J Qual Care Clin Outcomes*. 2024;10:176–188.
- Stolker JM, Spertus JA, Cohen DJ, et al. Rethinking composite end points in clinical trials insights from patients and trialists. *Circulation*. 2014;130:1254–1261.
- Braunwald E, Cannon CP, McCabe CH. Use of composite endpoints in thrombolysis trials of acute myocardial infarction. *Am J Cardiol*. 1993;72:3G–12G.
- Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33:176–182.
- Redfors B, Gregson J, Crowley A, et al. The win ratio approach for composite endpoints: practical guidance based on previous experience. *Eur Heart J*. 2020;41:4391–4399.
- Gregson J, Stone GW, Bhatt DL, et al. Recurrent Events in Cardiovascular Trials: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2023;82:1445–1463.
- Ferreira JP, Jhund PS, Duarte K, et al. Use of the Win Ratio in Cardiovascular Trials. *JACC Heart Fail*. 2020;8:441–450.
- Maurer MS, Schwartz JH, Gundapaneni B, et al. Tafamidis Treatment for Patients with Transthyretin Amyloid Cardiomyopathy. *N Engl J Med*. 2018;379:1007–1016.
- James S, Erlinge D, Storey RF, et al. Dapagliflozin in Myocardial Infarction without Diabetes or Heart Failure. *NEJM evidence*. 2024;3:EVIDoa2300286.
- Pocock SJ, Gregson J, Collier TJ, Ferreira JP, Stone GW. The win ratio in cardiology trials: lessons learnt, new developments, and wise future use. *Eur Heart J*. 2024;45:4684–4699.
- Gaudino M, Braunwald E, Stone GW. Beyond the classic major cardiovascular event outcome for cardiovascular trials. *Eur Heart J*. 2024;45:4700–4703.
- Raposeiras-Roubin S, Amat-Santos IJ, Rossello X, et al. Dapagliflozin in Patients Undergoing Transcatheter Aortic-Valve Implantation. *N Engl J Med*. 2025;392:1396–1405.
- Amat-Santos IJ, Sánchez-Luna JP, Abu-Assi E, et al. Rationale and design of the Dapagliflozin after Transcatheter Aortic Valve Implantation (DapaTAVI) randomized trial. *Eur J Heart Fail*. 2022;24:581–588.
- Gonzalez-Del-Hoyo M, Mas-Llado C, Blaya-Peña L, Siquier-Padilla J, Peral V, Rossello X. The Fragility Index in randomised clinical trials supporting clinical practice guidelines for acute coronary syndrome: measuring robustness from a different perspective. *Eur Heart J Acute Cardiovasc Care*. 2023;6:386–390.
- Gonzalez-Del-Hoyo M, Rossello X. Fragility index: tips and tricks for interpretation. *Eur Heart J Acute Cardiovasc Care*. 2024;13:446–448.
- Buyse M. Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. *Stat Med*. 2010;29:3245–3257.
- Ajufo E, Nayak A, Mehra MR. Fallacies of Using the Win Ratio in Cardiovascular Trials: Challenges and Solutions. *JACC Basic Transl Sci*. 2023;8:720–727.
- Butler J, Jones WS, Udell JA, et al. Empagliflozin after Acute Myocardial Infarction. *N Engl J Med*. 2024;390:1455–1466.
- Hernandez AF, Udell JA, Jones WS, et al. Effect of Empagliflozin on Heart Failure Outcomes After Acute Myocardial Infarction: Insights from the EMPACT-MI Trial. *Circulation*. 2024;1627–1638.
- Udell JA, Petrie MC, Jones WS, et al. Left Ventricular Function, Congestion, and Effect of Empagliflozin on Heart Failure Risk After Myocardial Infarction. *J Am Coll Cardiol*. 2024;2233–2246.
- Packer M, Anker SD, Butler J, et al. Cardiovascular and Renal Outcomes with Empagliflozin in Heart Failure. *N Engl J Med*. 2020;383:1413–1424.
- Pocock SJ, Collier TJ. Statistical Appraisal of 6 Recent Clinical Trials in Cardiology. *J Am Coll Cardiol*. 2019;73:2740–2755.
- Bhatt AS, Kosiborod MN, Vaduganathan M, et al. Effect of dapagliflozin on health status and quality of life across the spectrum of ejection fraction: Participant-level pooled analysis from the DAPA-HF and DELIVER trials. *Eur J Heart Fail*. 2023;25:981–988.
- Rossello X, Rubini Gimenez M. The dapagliflozin in patients with myocardial infarction (DAPA-MI) trial in perspective. *Eur Heart J Acute Cardiovasc Care*. 2023;12:862–863.
- Pocock SJ, Rossello X, Owen R, Collier TJ, Stone GW, Rockhold FW. Primary and Secondary Outcome Reporting in Randomized Trials: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2021;78:827–839.