

JORGE THIERER

Ablación por catéteres en fibrilación auricular: ¿igual o mejor que el tratamiento farmacológico? El estudio CABANA y sus diferentes lecturas

Packer DL, Mark DB, Robb RA, Monahan KH, Bahnson TD, Poole JE, et al. Effect of Catheter Ablation vs Antiarrhythmic Drug Therapy on Mortality, Stroke, Bleeding, and Cardiac Arrest Among Patients With Atrial Fibrillation: The CABANA Randomized Clinical Trial. **JAMA** 2019;321:1261-74. <http://doi.org/c6c9q>

Mark DB, Anstrom KJ, Sheng S, Piccini JP, Baloch KN, Monahan KH, et al. Effect of Catheter Ablation vs Medical Therapy on Quality of Life Among Patients With Atrial Fibrillation: The CABANA Randomized Clinical Trial. **JAMA** 2019;321:1275-85. <http://doi.org/c6c9r>

Durante años el tratamiento de la fibrilación auricular (FA) consistió en el empleo de drogas dromotrópicas negativas para lograr control de la frecuencia cardíaca o drogas antiarrítmicas (DA) destinadas a control del ritmo, y anticoagulación oral (ACO). El control de ritmo demostró en diferentes estudios aleatorizados resultados a largo plazo similares a los de control de frecuencia; a eso debemos sumar la incidencia variable de efectos adversos debidos al uso de DA, y una efectividad alejada que distaba de ser óptima. En las últimas dos décadas asistimos al surgimiento y desarrollo de la terapia de ablación (Ab) por catéteres de la FA, como una alternativa de control de ritmo que empezó a disputar la primacía al tratamiento con DA. Progresivamente los pacientes tratados con esta técnica fueron más complejos, pero más allá de estudios observacionales con evidente sesgo de selección (pacientes tratados con Ab cuando las condiciones eran óptimas) que mostraban resultados prometedores, y de estudios aleatorizados, también con pacientes muy cuidadosamente elegidos, que sugerían mejores resultados con la Ab (menor recurrencia de la FA, mejor evolución de la fracción de eyección ventricular izquierda), no ha habido hasta el momento estudios de grandes dimensiones que confirmen ventaja de la Ab sobre el tratamiento farmacológico usual. Recientemente conocimos el estudio CASTLE AF, que comentamos en RAC 2018; vol 86 nro 1. Como dijimos en dicho comentario, este estudio incluyó poco más de 300 pacientes sobre más de 3.000 inicialmente considerados, lo que revela una selección muy marcada. De allí que los resultados impactantes de dicho estudio (reducción de mortalidad, el primero) no hayan modificado sustancialmente la práctica clínica.

Llegan ahora a nosotros dos publicaciones del estudio CABANA, de mucho mayores dimensiones, que comparó la Ab con el uso de DA en pacientes con FA. Fueron criterios de ingreso haber presentado en los últimos 6 meses 2 o más episodios de FA paroxística

o al menos 1 episodio de FA persistente (duración ≥ 7 días o necesidad de cardioversión eléctrica luego de al menos 48 horas de instalada); y tener edad ≥ 65 años o < 65 años con al menos 1 factor de riesgo para accidente cerebrovascular (ACV): hipertensión arterial, diabetes, ACV previo, insuficiencia cardíaca u otros problemas cardiovasculares. Los pacientes fueron adjudicados aleatoriamente en relación 1:1 a Ab (siendo obligatorio realizar aislamiento de las venas pulmonares y quedando a discreción de los médicos tratantes, que debían tener experiencia de al menos 100 procedimientos previos, cualquier técnica adicional), o el empleo de DA (con la recomendación de intentar inicialmente control de frecuencia y solo si éste fracasaba, control de ritmo). Se excluyó a aquellos que hubieran sido sometidos previamente a una Ab fallida, o en los que hubiera fracasado el tratamiento con 2 DA. Inicialmente el punto final primario del estudio fue mortalidad total. Se planteó una mortalidad esperada a 3 años de 12% en la rama DA, y un descenso del 30% con Ab. Con estos supuestos, considerando un poder de 90% y un error alfa de 5%, sería necesario incluir a 3.000 pacientes. El punto final secundario fue un compuesto de muerte, ACV incapacitante, sangrado serio o paro cardíaco.

En el curso del estudio resultó evidente que la tasa de inclusión y la incidencia de eventos eran menores que las esperadas. Ello llevó a invertir los puntos finales: el secundario pasó a ser primario, y se extendió el seguimiento a 4 años. Con ello el número de pacientes necesarios bajó a 2.200. La mortalidad total quedó como punto final secundario, al igual que la recurrencia de la FA, cambios en la calidad de vida y un compuesto de mortalidad y hospitalización de causa cardiovascular. Para evaluar la recurrencia de la FA un subgrupo de pacientes recibió un dispositivo de monitoreo para registrar cada evento sintomático, así como para hacer grabaciones de 24 y 96 horas a intervalos pre especificados. Se estableció tras la entrada en el estudio un período de 3 meses de “blanqueo” durante el cual podía realizarse, ante el fracaso de la terapéutica, un nuevo procedimiento en la rama Ab, y ensayo de drogas en la rama DA, sin considerar esas recurrencias en la comparación. Para el análisis de los datos se planteó como análisis primario uno por intención de tratar. Pero además se planteó prospectivamente un análisis por protocolo (comparando a los pacientes de la rama Ab que efectivamente fueron sometidos al procedimiento con todos los pacientes de la rama DA), y uno por tratamiento real (en el que se compararan todos los pacientes sometidos a uno u otro tratamiento, independientemente de que les hubiera sido adjudicado en la aleatorización o resultara de un cruce de rama durante el estudio).

Entre 2009 y 2016 fueron incluidos en el estudio 2.204 pacientes de 126 centros en 10 países (1.108 en la rama Ab y 1.096 en la rama DA). La edad media era 68 años, el 63% eran hombres; poco más del 80% eran hipertensos, y del 25%, diabéticos. Un 10% tenía antecedente de ACV o accidente isquémico transitorio. La mediana de puntaje CHA₂DS₂ Vasc era 3. El 43% tenía antecedente de FA paroxística, un 47% de FA persistente y el resto de FA permanente.

En la rama Ab el 90,6% fue efectivamente sometido al procedimiento en el mes siguiente a la aleatorización. En 215 pacientes (el 19,4%) fue necesaria una nueva ablación por recurrencia de la FA en el seguimiento, en 25 de los casos en el período de “blanqueo”. Casi el 45% de los pacientes de esta rama recibió DA en algún momento del seguimiento; al final del estudio el 26,5% de los pacientes de la rama Ab estaba medicado.

En la rama DA el 99,6% recibió efectivamente tratamiento farmacológico, en la mitad de los casos una sola droga. El 88,4% recibió tratamiento dirigido al control del ritmo. A lo largo del seguimiento el 27,5% de los pacientes de esta rama fue sometido a Ab.

En el análisis por intención de tratar no hubo diferencia en la incidencia del punto final primario: 8% en la rama Ab, 9,2% en la rama DA (HR 0,86, IC 95% 0,65-1,15, $p = 0.30$). No hubo tampoco diferencia en la mortalidad total (5,2% vs. 6,1%, $p = 0.38$) pero sí en el punto final combinado de muerte u hospitalización de causa cardiovascular: 51,7% vs. 58,1% (HR 0,83, IC 95% 0,74-0,93, $p = 0.001$). En cambio, en el análisis por tratamiento recibido sí se verificó reducción significativa de riesgo con Ab respecto de DA para el punto final primario (HR 0,67, IC 95% 0,50-0,89) y para mortalidad total (HR 0,60, IC 95% 0,42-0,86). Lo mismo sucedió en el análisis por protocolo a 12 meses (HR de 0,73 y 0,68 respectivamente, ambos significativos). En lo que hace a la recurrencia de la FA el análisis por intención de tratar reveló con Ab una disminución del riesgo a menos de la mitad respecto de los resultados con DA. Si al inicio del estudio el 57% de los pacientes tenía FA persistente o permanente, la cifra había disminuido a 26% en la rama DA y 16% en la rama Ab al fin del seguimiento.

Respecto de la incidencia de eventos adversos, en la rama Ab los más frecuentes fueron hematomas menores (2,3%), pseudoaneurismas (1,1%) y taponamiento cardíaco (0,8%). En la rama DA los más incidentes fueron trastornos tiroideos (1,6%) y proarritmia (0,8%).

En forma conjunta con la publicación de los resultados clínicos principales se dio a conocer el estudio sobre calidad de vida. La misma fue evaluada con dos instrumentos específicos, los puntajes AFEQT y MAFSI. El puntaje AFEQT se obtiene considerando la respuesta a 18 de 21 preguntas específicas sobre síntomas, actividades diarias y tratamiento. El puntaje oscila entre 0 (completa incapacidad vinculada con la presencia de la FA) y 100 (ausencia total de incapacidad). Una variación en el puntaje de 5 o más puntos implica un cambio significativo en la calidad de vida. El

puntaje MAFSI considera 10 ítems, e interroga sobre frecuencia y severidad de los síntomas. Para puntajes máximos de 40 para frecuencia y 30 para severidad (los síntomas más graves en uno y otro caso) se interpretó que descensos de 1,6 y 1,3 puntos respectivamente implican mejoría significativa. Los puntajes citados fueron evaluados en forma basal, a los 3 y 12 meses el primer año y luego cada 12 meses. Por otra parte, un cuestionario que incorpora preguntas de estos 2 puntajes y de otros tradicionalmente empleados para valorar calidad de vida (como el SF 36) fue administrado a los 6 meses y luego cada 12. En total, considerando la cantidad de pacientes y un seguimiento de hasta 60 meses, se llegó a una cantidad de 20.461 cuestionarios de los cuales un 90% fue efectivamente administrado.

En lo que hace al puntaje AFEQT los pacientes de la rama Ab presentaron en forma basal un valor de 62,9, y los de la rama DA de 63,1. Al cabo de 1 año el puntaje medio subió en ambas ramas a 86,4 y 80,9 respectivamente. La diferencia de 5,3 puntos entre ambas ramas fue estadísticamente significativa e implica mejoría significativa de la calidad de vida con la Ab respecto de las DA. La mejoría fue más notable en el tercio de pacientes con peor puntaje basal: 7,7 puntos frente a 5,3 en el tercio medio y 2,7 en el tercio con puntajes más altos. A 5 años la diferencia se atenuó (3,4 puntos) pero siguió siendo significativa. Respecto del puntaje MAFSI, también fue mayor la mejoría (en este caso el descenso) en los pacientes de la rama Ab, con una diferencia en el descenso medio del puntaje de frecuencia de 1,7 puntos y de 1,5 del puntaje de severidad, en ambos casos superior a lo necesario para entender que ha habido progreso en la calidad de vida.

El estudio CABANA ilustra algunos de los puntos conflictivos cuando se trata de estudios aleatorizados. Si bien es cierto que los mismos eliminan el sesgo de selección a favor de una conducta determinada, no lo es menos que existe una selección previa, la de aquellos que ingresan al estudio. Por eso, comparados con los pacientes del mundo real, los pacientes de los ensayos clínicos son más jóvenes y con menor tasa de comorbilidades. Los médicos participantes son también seleccionados en base a su experiencia y pertenencia a centros que cumplen determinados criterios de calidad de atención. En estos estudios es más factible por tanto que las estrategias testeadas sean efectivas, y menor la incidencia de complicaciones que en el mundo real. La dificultad en la inclusión de pacientes (poco más de 300 por año en más de 120 centros implica menos de 3 pacientes por centro por año) señala la selección de la población incluida. La necesidad de cambiar de punto final en medio del estudio y prolongar el seguimiento para disminuir la cantidad de pacientes a incluir es otra prueba de lo dicho.

El análisis por intención de tratar es el único que preserva la pureza de la aleatorización, que justamente aspira a repartir por igual las características basales conocidas y desconocidas de los pacientes, permitiendo atribuir los resultados entonces solo a la intervención.

En dicho análisis la Ab no resultó superior al empleo de DA. Pero es justo también reconocer que el cruce de pacientes y la falta de cumplimiento con lo asignado en la aleatorización dificultan la interpretación de los resultados. En la rama Ab un 10% no fue sometido al procedimiento, casi la mitad recibió DA en algún momento del estudio y al final del estudio 1 de cada 4 pacientes estaba medicado con alguna de estas drogas. En la rama DA, a su vez, y como en espejo, poco más de 1 de cada 4 pacientes cruzó a la realización de Ab, diluyendo entonces en el análisis por intención de tratar cualquier ventaja que el procedimiento hubiera podido ofrecer respecto del punto final. De hecho el análisis por protocolo o tratamiento real muestra sí superioridad de la Ab sobre el uso de DA, pero la ventaja de la aleatorización ya se ha perdido: los pacientes tratados de una u otra forma ya no son similares entre sí; por algo no se han sometido al procedimiento asignado, por algo han cruzado de rama. Cuando el cruce de una rama a otra es tan pronunciado (más de un 25%) nos encontramos en un dilema: ¿cuánto de lo que informa el análisis por intención de tratar representa la realidad? Tal vez comprender que lo que se compara no son tratamientos sino estrategias pudiera reconciliar ambos puntos de vista: una estrategia inicial de uso de DA no resulta inferior a la de implementar Ab, pero es posible que por falla en el tratamiento una proporción significativa de pacientes tenga que caer en la otra alternativa y logre con eso mejor evolución. Merece destacarse sí que el empleo de Ab en el análisis por intención de tratar aseguró menor recurrencia de la FA, y menor incidencia de internación de causa cardiovascular.

En cuanto al estudio de calidad de vida, nótese que hubo mejoría de la misma en ambas ramas, con ventaja para la Ab. Por la naturaleza abierta del estudio, en que cada paciente sabe a qué rama fue asignado, es muy difícil poder excluir un efecto placebo, en que los pacientes sometidos a una intervención más compleja puedan por eso solo sentir que han mejorado sus síntomas; pero la diferencia notable en la recurrencia de FA (menos de la mitad con Ab respecto del uso de DA) da a este hallazgo un tinte de verosimilitud.

¿Son efectivas las estatinas en los mayores de 75 años? Resultados de un metaanálisis

Cholesterol Treatment Trialists C. Efficacy and safety of statin therapy in older people: a meta-analysis of individual participant data from 28 randomised controlled trials. **Lancet** 2019;393:407-15. <http://doi.org/gft6nz>

Diferentes metaanálisis han confirmado que una reducción mediante el uso de estatinas del valor de colesterol LDL de 1 mmol/l (38,67 mg/dl) se traduce en una disminución de aproximadamente un 20% en el riesgo de eventos vasculares. Ello se aplica a hombres y mujeres y en todo el espectro de riesgo cardiovascular. Sin embargo, dudas se han expresado sobre la eficacia

del tratamiento en mayores de 75 años, y se ha llegado a sostener la futilidad de la intervención por encima de dicha edad. Conocemos ahora una nueva publicación sobre el particular de la Colaboración CTT (Cholesterol Treatment Trialists¹), un grupo establecido en 1994, dedicado al metaanálisis de estudios aleatorizados de estatinas para definir su seguridad y eficacia.

Los autores han analizado 28 estudios con al menos 1.000 pacientes incluidos y un seguimiento ≥ 2 años, en 23 de los cuales las estatinas fueron comparadas con placebo o tratamiento convencional, mientras que en 5 se compararon dos esquemas de tratamiento con estatinas, más y menos intensivo. En total son considerados 186.804 pacientes, para el análisis divididos según la edad en: ≤ 55 años (21%), 56-60 años (17%), 61-65 años (20%), 66-70 años (20%), 71-75 años (15%) y el grupo de interés, los > 75 años, que representan el 8%. Los > 75 años han sido más frecuentemente representados en 4 estudios llevados a cabo en pacientes con insuficiencia cardíaca (20% vs 4% de los < 75 años) o en diálisis (3% vs 2% de los < 75 años). Debe recordarse que en ambas condiciones clínicas las estatinas no han demostrado ser efectivas.

En lo que hace al efecto de las estatinas (comparadas con control, o un tratamiento más intenso respecto de otro menos intenso) sobre la incidencia de eventos vasculares mayores, el metaanálisis demuestra una reducción del 21% por cada mmol/l de disminución del colesterol LDL (HR 0,79, IC 95% 0,77-0,81). Se verifica una tendencia a que el efecto disminuya con la edad (HR 0,75 en los ≤ 55 años, HR 0,87 en los > 75 años), que está en el límite de la significación (p para tendencia = 0,06). Al eliminar del análisis los estudios en insuficiencia cardíaca y diálisis, donde están sobre representados los > 75 años, esa tendencia desaparece, es decir que el efecto beneficioso de las estatinas no se reduce con el aumento de la edad. Al tener en cuenta la presencia o ausencia de enfermedad vascular establecida, se verifica que no hay diferencia según la edad en el efecto de las estatinas en los pacientes en prevención secundaria, mientras que en cambio el beneficio disminuye claramente en los pacientes en prevención primaria a medida que son mayores, tornándose no significativo en los > 75 años.

Al considerar el efecto sobre eventos coronarios la reducción global por cada mmol/l de disminución de colesterol LDL es de 24%. Nuevamente hay tendencia a menor efecto a medida que aumenta la edad, que se mantiene aun cuando se eliminan los estudios en insuficiencia cardíaca y diálisis; pero en los ≥ 75 años, y aunque menor, el efecto sigue siendo significativo. En lo que hace a la prevención de procedimientos de revascularización y de accidentes cerebrovasculares (25% y 16% en forma global por cada mmol/l de colesterol LDL respectivamente) no hay diferencias según la edad.

Finalmente, en lo que hace a la prevención de la muerte cardiovascular y muerte de todas las causas (12% y 9% en forma global por cada mmol/l de colesterol LDL, respectivamente), también se verifica tendencia

a pérdida del efecto con aumento de la edad. Pero al eliminar los estudios de insuficiencia cardíaca y diálisis, que concentraron el 53% de las muertes vasculares, dicha tendencia desaparece. No se verifica efecto de las estatinas sobre las muertes no vasculares o las muertes por cáncer.

La discusión sobre la utilidad de emplear estatinas en ancianos es de larga data. La baja tasa de inclusión en muchos estudios, la coexistencia de comorbilidades que conspiran contra la capacidad de estas drogas para mejorar el pronóstico (como la insuficiencia renal terminal y la insuficiencia cardíaca) son factores que pueden haber contribuido a que el efecto específico en mayores de 75 años no estuviera claro. En este sentido el metaanálisis que presentamos es de notoria importancia, porque en líneas generales confirma el beneficio de tratar a los pacientes de alto riesgo aun cuando sean más añosos. Es de remarcar sin embargo que este efecto claro se logra en pacientes sin las comorbilidades descriptas, y fundamentalmente en prevención secundaria. La disminución del beneficio en ancianos puede deberse a la naturaleza del fenómeno aterosclerótico, y la influencia creciente que la hipertensión arterial y la disfunción renal pueden tener a medida que aumenta la edad. Tal vez por eso en el caso específico de la prevención primaria la evidencia es mucho menos firme. En este sentido el estudio STAREE, actualmente en desarrollo contribuirá sin duda a darnos una respuesta.

Debe tenerse en cuenta sin embargo que aunque proporcionalmente la reducción de riesgo sea menor en ancianos que en gente joven, por ser el riesgo basal de eventos vasculares mayor en los primeros, la cantidad de eventos prevenidos cada 1.000 tratados resulta en ellos mayor. De allí que, aunque la polémica continúe, parece sin duda lo más adecuado tratar a los pacientes de riesgo independientemente de la edad, teniendo desde ya en cuenta en cada caso en cuenta el beneficio esperable, la tolerancia y el riesgo de eventos adversos.

El sedentarismo, la actividad física y su interacción a la hora de definir pronóstico. Un estudio observacional en 150.000 personas

Stamatakis E, Gale J, Bauman A, Ekelund U, Hamer M, Ding D. Sitting Time, Physical Activity, and Risk of Mortality in Adults. **J Am Coll Cardiol** 2019;73:2062-72. <http://doi.org/c6ct>

Es claro que la actividad física llevada a cabo en forma regular mejora el pronóstico cardiovascular y que el comportamiento sedentario, definido habitualmente como un gasto energético inferior a 1,5 METs (equivalentes metabólicos de consumo de oxígeno) en posición sentada o reclinada durante horas, entraña un pronóstico adverso. Pero es cierto que no estamos todo el tiempo en actividad ni todo el tiempo sentados (y si no estamos sentados puede deberse a que estamos parados o haciendo actividad física, o, por el contrario, a que estamos durmiendo). Hay información publicada acerca de que el comportamiento sedentario es predictor de

mal pronóstico cuando supera las 10 horas diarias; y también de que la actividad física moderada a intensa durante > 60 a 75 minutos diarios nulifica el mal pronóstico que acarrea el sedentarismo. El estudio que presentamos intentó justamente valorar el pronóstico conjunto de la actividad física y el sedentarismo, entendiendo que cada uno de ellos cuando está presente desplaza al otro a lo largo de las 24 horas del día.

Se trata de un estudio prospectivo de cohorte que se llevó a cabo en Nueva Gales del Sur (Australia) en hombres y mujeres ≥ 45 años. Se los sometió a un cuestionario estructurado en el que se los interrogó sobre el tiempo que diariamente estaban parados, sentados o durmiendo, así como el tiempo semanalmente consagrado a caminata, y actividad física moderada o intensa. La actividad física semanal fue categorizada como ausente (0 minutos), activa insuficiente (1-149 minutos), activa suficiente en rango inferior (150-299 minutos), activa suficiente en rango superior (300-419 minutos), y recomendada (≥ 420 minutos). El tiempo diario de sueño fue dicotomizado para el análisis en ≤ 7 horas y > 7 horas. Se recurrió a las estadísticas vitales del estado, y los datos fueron ajustados por edad, sexo, dieta, tabaquismo, diabetes, índice de masa corporal auto reportado, nivel socioeconómico, estado marital, residencia urbana o rural y valoración personal del estado de salud.

Fueron analizados los registros de 149.077 participantes de la encuesta, y su relación con la evolución. Para mortalidad total el tiempo medio de seguimiento fue de 8,9 años y para mortalidad cardiovascular de 7,4 años. Para cada uno de los grupos de actividad física citados se evaluó la mortalidad total y cardiovascular según el tiempo diario sentado (< 4 , $4 < 6$, $6 < 8$ y ≥ 8 horas). Hubo clara interacción entre el tiempo sentado y el tiempo dedicado a actividad física moderada a intensa para la mortalidad de todas las causas. El tiempo de sedentarismo fue predictor independiente de mortalidad total en las personas inactivas o insuficientemente activas. Entre aquellos con actividad suficiente en rango inferior, solo un tiempo de sedentarismo diario ≥ 8 horas se asoció a peor evolución. En aquellos con mayor actividad física el tiempo de sedentarismo ya no fue predictor de mortalidad. En lo que hace a mortalidad cardiovascular también se verificó más riesgo en aquellos con mayor tiempo de sedentarismo en los grupos con menor actividad semanal, aunque sin evidencia de una relación dosis respuesta clara.

Al analizar el efecto que tendría sobre la mortalidad total reemplazar 1 hora diaria de estar sentado por 1 de estar parado, se verificó una disminución del riesgo de un 3% solo entre aquellos que están sentados ≤ 6 horas diarias, pero no entre aquellos que están sentados más tiempo. En cambio, reemplazar 1 hora diaria de estar sentado por 1 de caminar o hacer actividad física intensa se asoció a reducción del riesgo sobre todo en los que están sentados > 6 horas diarias. Respecto de la muerte cardiovascular, cada hora adicional de estar sentado se asoció a exceso de riesgo (del 7%) solo

entre aquellos que están sentados > 6 horas diarias; el reemplazo de una hora sentado por una hora parado, a disminución del riesgo solo en aquellos que están sentados $\leq$ 6 horas diarias; pero el reemplazo de una hora sentado por una hora de actividad moderada a intensa, a reducción del riesgo sin importar la cantidad de horas que se está sentado.

La asociación del sedentarismo con una peor evolución tiene base fisiopatológica. Mayor activación de fenómenos inflamatorios, obesidad, disminución de la capacidad vasodilatadora, disfunción endotelial son algunos de los fenómenos que la explican. La actividad física revierte la situación, y mejora el pronóstico. Este análisis tiene de llamativo el hecho de considerar en forma conjunta el efecto del sedentarismo y el de la actividad física sobre el pronóstico vital. Habitualmente leemos publicaciones que toman en cuenta uno u otra. Pero como muy bien señalan los autores, existe entre ambos un fenómeno de interacción: una variable influye de diversa manera en la evolución según los diferentes estratos de la otra. Si la actividad física es suficiente, el pronóstico adverso que impone el sedentarismo se diluye, hasta desaparecer en aquellos que más tiempo están activos. Si el tiempo de sedentarismo es menor, alcanza con estar más tiempo parado para disminuir el riesgo de eventos; si es mayor será necesario hacer actividad física. Es cierto que aquellos más sedentarios pueden ser personas más enfermas, y que puede ser esto lo que explique el peor pronóstico; el hecho de ajustar por la valoración del estado de salud auto percibido corrige al menos parcialmente esta cuestión. El mensaje de esta publicación es claro: aquellos que por sus condiciones laborales o inclinación personal están gran parte del día sentados deben esforzarse en compensar esta condición con actividad física regular. Como comentario adicional quisiéramos volver a expresar la admiración por sociedades en las que, como en este caso, se puede describir la evolución de casi 150.000 personas que responden a un cuestionario y son seguidas durante casi 10 años, con la posibilidad de ajustar las observaciones por sus características basales. ¿Seremos capaces alguna vez de reproducir la experiencia?

La recalibración de los scores más empleados de riesgo cardiovascular mejora su capacidad predictiva e iguala su prestación

Pennells L, Kaptoge S, Wood A, Sweeting M, Zhao X, White I, et al. Equalization of four cardiovascular risk algorithms after systematic recalibration: individual-participant meta-analysis of 86 prospective studies. *Eur Heart J* 2019;40:621-31. <http://doi.org/c6cv>

Un score, puntaje, o regla de predicción clínica es una herramienta que permite aumentar la certeza en el diagnóstico, en el pronóstico o en la predicción de respuesta a una terapéutica tomando en cuenta en cada individuo datos de la historia, el examen y exámenes complementarios. Su origen radica, en general, en la

determinación del mejor modelo multivariado que exprese la asociación de las variables predictoras con la respuesta. En la valoración de un score se consideran dos propiedades: calibración y discriminación. La calibración del modelo tiene que ver con la capacidad de predecir en cada caso una probabilidad (de acuerdo con el valor de las covariables) que sea fiel reflejo de la probabilidad observada. Si la probabilidad estimada (número de observados sobre el total) que presentarán el punto final de interés) coincide con la observada, el score calibra adecuadamente. La discriminación se refiere a la capacidad de asignar a los individuos que efectivamente presentan la condición o evento de interés mayor probabilidad de ocurrencia de la misma que a los que carecen de ella, y se evalúa mediante el índice C, que corresponde el área bajo la curva ROC. Un modelo pronóstico que evalúe adecuadamente el riesgo de eventos ayuda a tomar decisiones terapéuticas. De hecho, la indicación de estatinas y la elección de un tratamiento de alta o moderada intensidad se basan en la utilización de diferentes modelos.

Hasta el año 2014, más de 100 modelos de riesgo cardiovascular habían sido publicados. Menos de la tercera parte de ellos contaban con validación externa. Y cuando hubo dicha validación, el desempeño no fue similar al del modelo original. Algunos ejemplos son por todos conocidos: el score de Framingham, por citar uno de los casos más señalados, sobreestima el riesgo de eventos cardiovasculares en población británica o española (mala calibración). Es cierto que en ese caso volver a calibrar el score en la nueva población (con las características propias de la misma) puede llevar a una modificación del test que le devuelva su utilidad (por ejemplo, el score REGICOR en España). Diferentes scores recomendados fuertemente por las guías de práctica pueden diferir sutilmente en el evento primario que pronostican, en las variables predictoras o en la forma en que las mismas han sido modeladas. La comparación entre diferentes scores aplicados a la misma población para valorar la indicación primaria de uno u otro es muy poco frecuentemente llevada a cabo. Por eso cobra valor la publicación que presentamos.

Fueron considerados 4 puntajes de uso habitual para valorar riesgo cardiovascular: las Ecuaciones de Cohorte Agrupadas, recomendado por American College of Cardiology y American Heart Association; el puntaje SCORE, recomendado por la Sociedad Europea de Cardiología; el puntaje de Reynolds y el tradicional score de Framingham. Se aplicaron los 4 modelos a 360.737 participantes sin enfermedad cardiovascular previa, de 86 estudios prospectivos de cohorte llevados a cabo entre 1963 y 2003 y se definió la capacidad de calibración y discriminación. Ninguna de esas cohortes había sido utilizada para la confección de los puntajes evaluados. El punto final primario fue el riesgo a 10 años de los eventos cardiovasculares específicamente considerados por cada score: infarto agudo de miocardio (IAM) no fatal, evento coronario fatal o cualquier accidente cerebrovascular (ACV) en el caso del score de

Framingham y el de ACC/AHA; lo mismo más revascularización coronaria o cualquier muerte cardiovascular para el score de Reynolds; muerte cardiovascular para el SCORE. Como punto final secundario para todos los puntajes se consideró un evento cardiovascular común a los cuatro, el ya referido punto final primario para los puntajes de Framingham y ACC/AHA. Cada score fue recalibrado modificando el algoritmo pronóstico de manera que tuviera en cuenta el perfil de factores de riesgo y la incidencia observada de eventos en cada cohorte, ajustando de esta manera el riesgo predicho al observado. Por otra parte, para permitir la comparación cabeza a cabeza entre los cuatro, se recalibraron los puntajes SCORE y Reynolds para que predijeran el evento cardiovascular común.

La edad media de los participantes fue 59 años; el 53% eran hombres. El 69% era europeo, el 18% norteamericano, el resto de Japón y Australia. La mediana de seguimiento fue de 10,2 años. La mediana de riesgo cardiovascular estimado a 10 años fue 5,5% con el score de Framingham, 2,5% con SCORE y 6,4% con el score ACC/AHA. La capacidad de discriminación de los diferentes puntajes, y para los diversos puntos finales considerados no difirió a grandes rasgos, con un índice C que osciló entre 0,70 y 0,76; hubo cierta superioridad (área bajo la curva ROC entre 0,003 y 0,013 mayor) de los puntajes ACC/AHA y SCORE sobre el score de Framingham. Ninguno de los 4 puntajes calibró adecuadamente al estimar el riesgo del evento específicamente predicho por cada modelo, respecto de lo que verdaderamente ocurrió: en forma global, considerando todas las cohortes, el score de Framingham sobreestimó el riesgo en un 10%, el de ACC/AHA en un 41%, el SCORE en un 52%, mientras que el de Reynolds lo infraestimó en un 10%. Los puntajes SCORE y ACC/AHA cometieron el mismo error de calibración en hombres y mujeres y en todo el rango etario; el de Framingham sobreestimó el riesgo en hombres y mujeres jóvenes, y lo infraestimó en mujeres mayores. Al considerar cada cohorte de las 86 por separado, hubo casos de sobreestimación de riesgo de hasta un 400%, y de infraestimación del mismo de hasta un 50%.

Ahora bien, tras procederse a la recalibración de los 4 scores ajustando los algoritmos pronósticos a la distribución de los factores de riesgo en cada cohorte, y tomando como objetivo el punto final común propuesto (IAM no fatal, evento coronario fatal o cualquier ACV), el desempeño de los 4 puntajes fue similar. Antes de la recalibración habían sido considerados como de alto riesgo, y por lo tanto con indicación de iniciar estatinas, el 32% de los participantes con el puntaje de Framingham, el 29% con SCORE, el 39% con el puntaje ACC/AHA y el 32% con el puntaje de Reynolds. Tras la recalibración las cifras respectivas descendieron al 22%, 22%, 24% y 23% respectivamente. No hubo diferencias significativas en la capacidad de discriminación.

La construcción de un score tiene 2 etapas básicas: la derivación a partir de la cohorte original, y la validación en otras cohortes. La mayor parte de las veces

los autores se conforman con publicar el score edificado a partir de la cohorte de derivación, y no validan el mismo. Ello sucede, como dijimos, aproximadamente con 2 de cada 3 modelos pronósticos que se exponen a la consideración de la comunidad científica. Cuando se procede a la validación, muchas veces la misma se lleva a cabo en otra cohorte de la misma población de la que procede el score, con lo cual los resultados no son muy distintos a los originales. Pero un score que aspire a ser de uso clínico habitual requiere haber sido validado en otras poblaciones. Allí surge el problema de la mala calibración. La misma se explica por el hecho de emplear el modelo en una población con condiciones basales diferentes (en cuanto a la distribución de los factores de riesgo, condiciones socioeconómicas, acceso a tratamiento, etc.) y diferente incidencia de eventos. Los 4 scores considerados en esta publicación han sido ampliamente validados; nótese sin embargo como, aplicados a las mismas poblaciones difieren en su capacidad predictiva. Ello sucede porque sus algoritmos son diferentes, porque las variables pronósticas no son exactamente las mismas, porque su origen no es similar, y porque el punto final predicho no es exactamente el mismo. Cuando se procede a la recalibración tomando en cuenta las características de la población a la que se aplican, y se modifican para que predigan exactamente el mismo punto final, las diferencias desaparecen; y la proporción de pacientes de alto riesgo es menor, ya que en general sobreestiman el riesgo.

Los hallazgos son interesantes desde el punto de vista metodológico, pero al mismo tiempo ilustran las flaquezas de los modelos pronósticos: las condiciones cambiantes de las poblaciones a las que se aplican hacen ilusoria la certeza predictiva cuando se predice en base a condiciones que ya no son las mismas. El trabajo de recalibrar los scores es una tarea que difícilmente se lleve a cabo periódicamente e implica importante uso de recursos, pero es un intento de lograr mayor cercanía con la verdad. Solo conoce la realidad aquel que es capaz de predecirla...acertadamente.

Asociación entre consumo de bebidas azucaradas y dietéticas con la mortalidad en dos grandes estudios de cohorte

Malik VS, Li Y, Pan A et al. Long-Term Consumption of Sugar-Sweetened and Artificially Sweetened Beverages and Risk of Mortality in US Adults. **Circulation** 2019;139:2113-25. <http://doi.org/c6cx>

El consumo de bebidas azucaradas (BA) es la fuente más importante de azúcar en la dieta. Representa en adultos en los Estados Unidos un 6% de la ingesta calórica total. En estudios epidemiológicos se ha vinculado dicho consumo con mayor riesgo de obesidad, diabetes tipo 2, enfermedad coronaria y accidente cerebrovascular. La relación con la mortalidad de todas las causas o con mortalidad cardiovascular es menos clara, con algunos estudios que sostienen la existencia de dicha asociación y otros que la ponen en duda. En

el caso de las bebidas dietéticas (BD) la información es más escasa. Para aclarar este punto se llevó a cabo un análisis de 2 grandes cohortes que a lo largo de las últimas décadas han brindado información relevante: el Estudio de las Enfermeras, que se inició en 1980, e incluyó 121.700 mujeres de entre 30 y 55 años, y el Estudio de los Profesionales de la Salud, que comenzó en 1986 e incluyó 51.529 hombres de entre 40 y 75 años. En ambos estudios, entre otras determinaciones basales, se administró un cuestionario que definió las características de la alimentación diaria en cada participante, incluyendo preguntas específicamente dirigidas al consumo de BA y BD. Se excluyó en este análisis a los participantes que presentaran al momento de ser incluidos diabetes, enfermedad cardiovascular o cáncer, quedando entonces 80.647 mujeres y 37.716 hombres. Para cada bebida se establecieron las siguientes categorías de consumo: < 1 vez por mes, 1-4 veces por mes, 2 a 6 veces por semana, 1 vez por día o 2 a más veces diarias.

A lo largo del seguimiento (34 años entre las mujeres, 28 años entre los hombres) se verificó un descenso progresivo del consumo de BA, y un incremento seguido por un descenso en el consumo de BD. Aquellos con mayor consumo de BA fueron más jóvenes, menos activos físicamente y más fumadores. En ellos fue mayor el consumo de carne roja y carga glucémica, y menor el consumo de granos enteros y vegetales. Aquellos con mayor consumo de BD fueron también más jóvenes, con mayor índice de masa corporal y menor carga glucémica. El consumo creciente de BA, ajustado por edad y consumo de BD, se asoció a la mortalidad de todas las causas. Respecto de la categoría más baja, la más alta presentó un HR para mortalidad total de 1,52 (IC 95% 1,43-1,61). El riesgo fue mayor en las mujeres que en los hombres (HR de 1,63 y 1,29 respectivamente). Tras ajustar por historia familiar, etnia, factores de riesgo coronario, alimentación e índice de masa corporal, la relación se atenuó pero siguió siendo significativa (HR 1,21, IC 95% 1,13-1,28). El exceso de riesgo asociado al incremento de una porción adicional (lata, vaso o botella según la bebida) fue del 7%, y tras ajustar por los factores citados, del 5%. Se verificó interacción con el sexo, con una influencia mayor en mujeres que en hombres. El consumo creciente de BA se asoció también a mayor mortalidad cardiovascular (HR ajustado 1,31, IC 95% 1,15-1,50) y a mayor riesgo de muerte por cáncer (HR ajustado 1,16, IC 95% 1,04-1,29), sobre todo con riesgo de muerte por cáncer de mama.

El consumo de BD mostró una asociación de menor cuantía con la mortalidad total y cardiovascular en las mujeres: HR de 1,1 y 1 respectivamente para la categoría más alta de consumo en relación con la más baja, que se atenuaron tras el ajuste multivariado. Cuando se consideró una categoría de al menos 4 veces por día la relación se hizo más fuerte, con exceso de riesgo de 30% para la mortalidad total y 43% para la cardiovascular. No se verificó asociación con la mortalidad total o cardiovascular en los hombres. El análisis permitió

estimar que reemplazar una porción de BA por una de BD se traduciría en una reducción del 4% en la mortalidad total, 5% en la cardiovascular y 4% en la mortalidad por cáncer.

La asociación entre el consumo de BA y la mortalidad total, evidenciada en este análisis de 2 grandes estudios de cohorte confirma hallazgos previos y es biológicamente plausible, habida cuenta de los efectos metabólicos adversos que la alta carga de carbohidratos puede producir, con mayor riesgo de obesidad, diabetes, activación de fenómenos inflamatorios y enfermedad vascular, amén de la relación demostrada con ciertos tipos de cáncer. La asociación del consumo de BD con la mortalidad es menos convincente. En principio porque aún no tenemos una explicación fisiopatológica clara. Hay quienes sostienen que el sabor dulce de estas bebidas estimularía el consumo de alimentos también dulces pero no dietéticos. Sin embargo, no hay evidencia de que consumir BD lleve per se a aumento de peso. La causalidad reversa (consumen BD personas hipertensas, obesas y por ello con resistencia a la insulina incrementada, y a la larga con mayor riesgo de diabetes) no puede ser totalmente excluida. Por otra parte llama la atención que dicha relación se haya demostrado solo en las mujeres, y solo con consumo elevado, por lo que la asociación luce menos firme, y debiera ser confirmada en nuevos estudios.

Semejanzas y diferencias de los efectos de los inhibidores de SGLT2 y los análogos GLP1 en la evolución de los pacientes diabéticos. Un metaanálisis de estudio aleatorizados

Zelniker TA, Wiviott SD, Raz I et al. Comparison of the Effects of Glucagon-Like Peptide Receptor Agonists and Sodium-Glucose Cotransporter 2 Inhibitors for Prevention of Major Adverse Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus. **Circulation** 2019;139:2022-31. <http://doi.org/c6cw>

En los últimos 4 años hemos asistido a una revolución en el campo de la diabetes: tras décadas en las que el éxito del tratamiento hipoglucemiante parecía centrado en la reducción de la glucemia y de los eventos microvasculares, hemos conocido drogas (los inhibidores de SGLT2 o gliflozinas y los análogos GLP 1) capaces de modificar el pronóstico vital y específicamente disminuir la incidencia de eventos adversos cardiovasculares mayores (EACM): mortalidad cardiovascular, infarto de miocardio (IAM) no fatal y accidente cerebrovascular (ACV) no fatal. Incluso la empagliflozina y el liraglutide han demostrado reducir la mortalidad de todas las causas. Se han postulado diferentes mecanismos por los que estas drogas ejercen beneficio: natriuresis y protección renal en el caso de las gliflozinas (un efecto fundamentalmente hemodinámico) y un efecto predominantemente metabólico y anti aterosclerótico en el caso de los agonistas GLP 1. ¿Son finalmente los resultados de emplear uno u otro tipo de drogas similares en lo que hace al pronóstico de los pacientes? Un

metaanálisis de estudios aleatorizados viene a responder esta pregunta. Considera 8 estudios aleatorizados. En 5 de ellos se testearon agonistas GLP 1 (ELIXA con lixisenatide, EXSCEL con exenatide, SUSTAIN 6 con semaglutide, HARMONY con albiglutide y LEADER con liraglutide), y en 3, gliflozinas (EMPA-REG Outcome con empagliflozina, CANVAS con canagliflozina y DECLARE con dapagliflozina).

En total fueron incluidos 77.242 pacientes, el 55,6% de ellos en los estudios con agonistas GLP 1. La edad media osciló entre 60 y 65 años, la proporción de mujeres entre 28% y 40%. La proporción de pacientes con enfermedad aterosclerótica establecida fue 73,1%, pero varió entre el 41% y el 100% según los criterios de inclusión. La prevalencia de insuficiencia cardíaca fue 16,3%, (entre 10 y 24 % según los estudios). La prevalencia de disfunción renal definida como filtrado glomerular < 60 ml/min/1,73 m² varió entre 20 y 29% excepto en el estudio DECLARE, donde fue de solo el 7,4%.

El efecto de ambos tipos de drogas sobre la incidencia de EACM fue similar: HR 0,88, IC 95% 0,84-0,94 para los agonistas GLP 1 y HR 0,89, IC 95% 0,83-0,96 para las gliflozinas. En realidad, el efecto se concentró en los pacientes con enfermedad vascular establecida, en los que la reducción alcanzó el 14% y fue similar con ambas familias de drogas, mientras que en aquellos pacientes con solo la presencia de factores de riesgo el efecto fue nulo (hallazgo también común a ambos tipos de agentes).

Tanto los agonistas GLP 1 como las gliflozinas redujeron significativamente el riesgo de IAM: un 9% los primeros, un 11% las gliflozinas, sin evidencia de heterogeneidad entre ambas drogas. De igual modo ambas redujeron el riesgo de muerte cardiovascular: un 12% los agonistas GLP 1, un 16% las gliflozinas. Solo los agonistas GLP 1 disminuyeron el riesgo de ACV, un 14%, mientras que no hubo reducción significativa con las gliflozinas. En cambio, solo estas últimas disminuyeron significativamente el riesgo de hospitalización por insuficiencia cardíaca, un 31%. Si bien ambas familias de drogas disminuyeron la incidencia de un punto final compuesto de eventos renales, el efecto fue más marcado con las gliflozinas, con una reducción del 38% frente a un 18% con los agonistas GLP 1. Además, hubo diferencia en el efecto específico: el de los GLP 1 se centró en una reducción de la incidencia de macroalbuminuria; el de las gliflozinas en la duplicación de las cifras de creatinina, la incidencia de insuficiencia renal terminal y la muerte renal.

Este metaanálisis de grandes estudios aleatorizados confirma lo ya sabido acerca del efecto favorable de ambas familias de drogas sobre los EACM. Confirma que el mismo se concentra en los pacientes que ya tienen enfermedad vascular instalada, y no se verifica en los que solo tienen factores de riesgo. Sin embargo, el estudio REWIND con dulaglutide, aún no publicado, demostró reducción de los EACM en pacientes con factores de riesgo libres de enfermedad vascular demostrada, pero en un seguimiento de 8 años, sugiriendo que tal vez

hace falta más tiempo para lograr efectos significativos en etapas más precoces. El efecto beneficioso de las gliflozinas sobre la incidencia de insuficiencia cardíaca y enfermedad renal ya ha sido suficientemente evidenciado, así como el hecho de que la reducción del ACV es esperable con los análogos GLP 1 y no con las primeras. Pero lo que llama poderosamente la atención es que gliflozinas y análogos GLP 1 disminuyen por igual la incidencia de IAM. Si hasta ahora se postulaba que el efecto anti aterosclerótico es específico de los análogos GLP 1, y los efectos de las gliflozinas son fundamentalmente hemodinámicos, ¿cómo explicar este punto? Era esperable que la reducción de la tasa de IAM fuera un aspecto en que los análogos GLP 1 sacaran ventaja. Pero no debiera olvidarse que por su acción glucosúrica los inhibidores de SGLT2 ejercen también acción antiinflamatoria, y por ende anti aterosclerótica, que podría estar involucrada en los hallazgos descriptos. Y que los efectos de ambas familias de drogas son tan variados que una explicación unívoca es sin duda ilusoria. No será la primera vez que la clínica ayude a interpretar la fisiopatología.

Nivel de evidencia de las recomendaciones de las guías de práctica. ¿Ha cambiado algo en los últimos 10 años?

Fanaroff AC, Califf RM, Windecker S, Smith SC, Jr., Lopes RD. Levels of Evidence Supporting American College of Cardiology/American Heart Association and European Society of Cardiology Guidelines, 2008-2018. **JAMA 2019;321:1069-80.** <http://doi.org/c6cz>

Las guías de práctica clínica (GPC) se han convertido en parte ineludible del proceso diagnóstico y terapéutico de la medicina contemporánea. Hay suficiente prueba de que el seguimiento de las mismas se traduce en mejores resultados en el cuidado de los pacientes. Sabemos que cada recomendación tiene 2 componentes: la fuerza de recomendación y el nivel de evidencia para sustentarla.

Respecto de la fuerza de recomendación, la misma puede ser I, II o III. Para la I se entiende que hay evidencia y/o acuerdo general en que un determinado tratamiento o procedimiento es beneficioso, útil y efectivo, de modo que diremos que la recomendación debe llevarse a cabo. Para la II hay evidencia conflictiva y/o divergencia de opinión acerca de la utilidad o eficacia del tratamiento o procedimiento; puede ser II a, si el peso de la evidencia u opinión está a favor del procedimiento o tratamiento o II b si la utilidad está menos establecida; de modo que una II a debería ser considerada y una II b, puede ser considerada. Para la III hay evidencia y/o acuerdo general en que un determinado tratamiento o procedimiento no es útil o efectivo, y en algunos casos puede llegar a ser perjudicial. Por eso es que no debe llevarse a cabo.

Respecto del nivel de evidencia, el mismo puede ser A (la evidencia proviene de 2 o más estudios aleatorizados, o de metaanálisis), B (la evidencia proviene de

1 solo estudio aleatorizado o de grandes estudios no aleatorizados), o C (la recomendación surge de consenso de opinión de expertos, y/o de evidencia de estudios pequeños, retrospectivos o registros).

En el año 2009 una publicación entonces muy comentada hizo notar que en las GPC de ACC/AHA solo el 11% de las recomendaciones se basaba en evidencia nivel A. Abundaron entonces las apelaciones a que esa situación mejorara. ¿Qué sucedió desde entonces? Un grupo de investigadores consideraron todas las GPC cardiológicas emitidas por ACC/AHA y la Sociedad Europea de Cardiología entre 2008 y 2018 y tabularon las recomendaciones según fuerza y calidad de evidencia.

En el período citado fueron publicadas 26 GPC de ACC/AHA sobre distintos tópicos, con un total de 2.930 recomendaciones. El 43,4 % fueron clase I, el 45,7% clase II y el 10,9 % clase III. A su vez, se basaron en evidencia A el 8,5%; en evidencia B el 50% y en evidencia C el 41,5%. De las recomendaciones más firmes, las I y las III, solo se basaron en evidencia A el 12,9%; en evidencia B el 48,9% y en evidencia C el 38,2%.

En el mismo período fueron publicadas 25 GPC de la Sociedad Europea de Cardiología, con un total de 3.399 recomendaciones. El 47,7 % fueron clase I, el 44,6% clase II y el 7,7% clase III. Se basaron en evidencia A el 14,2%; en evidencia B el 31 % y en evidencia C el 54,8%. De las recomendaciones más firmes, las I y las III, se basaron en evidencia A el 21,3%; en evidencia B el 29,1% y en evidencia C el 49,6%.

En 16 casos de GPC de ACC/AHA se encontró una guía publicada antes de 2008, referida al mismo tópico. La mediana de recomendaciones basadas en evidencia A no varió al comparar el período 2008-2018 respecto del anterior: 9% vs. 11,7%. En cambio, se notó un aumento de las recomendaciones basadas en evidencia B (51% vs 41,9%), y por ende un decremento de las basadas en evidencia C (36,7% vs 51,9%).

En el caso de 16 guías europeas publicadas entre 2014 y 2018, se encontró una guía sobre el mismo tópico publicada entre 2004 y 2014. En este caso no hubo diferencia significativa entre los períodos considerados en la proporción de evidencia A, B o C.

A la hora de seguir las recomendaciones de las guías esperamos que las más taxativas (lo que se debe y lo que no se debe) se basen en la evidencia más firme y menos discutible, de mayor calidad según los lineamientos de la Medicina Basada en la Evidencia. Es interesante entonces que en las guías americanas poco más de la

mitad de las recomendaciones sean I o III, pero solo el 13% se basen en evidencia A (y aun cuando la misma no estuviera disponible, y se buscara entonces evidencia B, resulta que casi un 40% descansa en consenso de expertos o pequeños estudios o registros). Y en las guías europeas la situación es apenas mejor, con una proporción similar de recomendaciones I o III, con más evidencia A (21%) pero también más evidencia C (casi el 50%).

¿Cómo podemos leer estos datos? En principio, parece claro que no hay tanta evidencia “de máxima pureza” cuando se trata de formular recomendaciones. De hecho, y yendo a las acciones más elementales, ¿conocemos ensayos aleatorizados sobre la necesidad de tomar la tensión arterial o realizar un ECG o un ecocardiograma? La inmensa mayoría de las acciones cotidianas encuentran su respaldo en la empiria, no en el ensayo clínico. ¿Y alguien discutiría la indicación inexcusable de llevar a cabo las anteriormente citadas? Parece por lo tanto irremediable que hay y seguirá habiendo muchas indicaciones I con nivel de evidencia C. Ahora bien, algunas acciones que en algún momento fueron clase I o III indiscutibles han caído en descrédito fruto a veces de la observación, a veces de algún ensayo clínico que se atrevió a desafiar la norma. ¿Dónde ha quedado la indicación de lidocaína profiláctica en el infarto, la contraindicación para los betabloqueantes en la insuficiencia cardíaca? Esto significa que en muchos casos la evidencia C podría ser removida y reemplazada por otra de mayor grado si hubiera atrevimiento y duda más frecuentes. En conclusión, podría concluirse que hay poca evidencia A en las guías.....porque efectivamente hay poca evidencia A. Los estudios aleatorizados son mayoritariamente impulsados por la industria, y se refieren por tanto en general al empleo de drogas y dispositivos. Estudios que intenten responder grandes preguntas acerca de estrategias de manejo clínico más allá de algún agente terapéutico en particular, son bienes más escasos. Pero, para el final, una pregunta difícil de responder: ¿por qué las guías americanas encuentran para sus recomendaciones evidencia A en el 8% de los casos, y las europeas casi en el doble, 14%? ¿Consideran fuentes diferentes, difieren en los ensayos o valoran la misma evidencia de forma disímil? Incluso lo que parece indiscutible (una evidencia determinada es A, o B, o C) quizás no lo sea tanto. Quizás, y bajo su apariencia de mandamiento bíblico, las recomendaciones de las guías descansan también en la interpretación.