



SUMARIO

<i>La interpretación de los datos en Medicina</i>	13
<i>Aminoglucósidos: ¿Mayor eficacia y menor toxicidad en dosis única diaria?</i>	15

LA INTERPRETACIÓN DE LOS DATOS EN MEDICINA

La Medicina Basada en la Evidencia (MBE) establece que sólo los metaanálisis y los megaensayos clínicos pueden proporcionar suficiente evidencia para aceptar efectos moderados o pequeños de los tratamientos sobre variables de morbilidad y mortalidad. Sin embargo, la toma de decisiones clínicas, no es sólo un problema de los resultados de los ensayos clínicos o metaanálisis sino que descansa también en la interpretación que se haga de los datos. Esta interpretación es un proceso complejo que está influido por muchos factores. Por poner un ejemplo, todos sabemos que un estudio suficientemente grande podrá encontrar significación estadística en diferencias que no tienen relevancia clínica y, lo contrario, uno muy pequeño no encontrará significación estadística para diferencias relevantes. Además, esta toma de decisiones está, o debe estar, modulada por otros muchos factores que no aparecen en los ensayos clínicos y que, en muchos casos, se basan en la aplicación de criterios cercanos al sentido común, como la "cláusula de necesidad" (no incluir un medicamento, eficaz y seguro, cuando no aporta ventajas sobre el que tenemos), la experiencia de uso (que es la auténtica prueba de fuego de su seguridad en amplias poblaciones) y muchos otros. En este número del BIT incluimos el primero de una serie de artículos, pensados para ayudar al lector a interpretar y valorar, en su justa medida, los datos científicos aparecidos en la literatura

Ante los resultados de un estudio hay que hacerse varias preguntas:

¿Las diferencias encontradas son reales o son debidas al azar? Es decir, ¿son reproducibles o no?. La respuesta la tenemos en el valor "p" o grado de significación estadística, es decir, la probabilidad de que, no existiendo diferencias en la población, las diferencias encontradas en los sujetos estudiados se deba al azar. Si la "p" es pequeña (inferior al nivel de significación alfa), significa que la probabilidad de que los resultados obtenidos se deban al azar es pequeña pero **no significa nada más**. No nos da una idea de la magnitud de la diferencia, ni de la fuerza de la asociación, ni del grado de eficacia de la intervención. Tan sólo confirma que la probabilidad de que me equivoque es menor que lo establecido al diseñar el estudio y esto depende, no sólo de la magnitud de la diferencia hallada, sino también del número de pacientes analizados.

La siguiente pregunta que debemos plantearnos es: **estas diferencias, ¿tienen relevancia clínica?** Las pruebas de significación estadística no deben suplantar

La magnitud de la "p" indica probabilidad, no magnitud de la diferencia de efecto.

La mejor manera de expresar la magnitud de la diferencia de efecto es mediante el riesgo atribuible (RA) y el número necesario de tratamientos (NNT).

Una buena manera de aumentar la eficacia de los aminoglucósidos es ajustar la dosis única diaria, mediante la determinación de niveles plasmáticos.

jamás al juicio clínico, pues hemos visto que aquellas pruebas sólo responden a la pregunta: ¿cuál es la probabilidad de que la diferencia observada se deba al azar? En definitiva, no todo lo que es "estadísticamente significativo" es significativo desde el punto de vista clínico. Así, diferencias muy pequeñas, pueden alcanzar significación estadística cuando incluimos en el estudio un número muy elevado de pacientes y, al contrario: diferencias importantes desde el punto de vista clínico no darán diferencias estadísticamente significativas si el estudio es de pequeño tamaño.

Por esto, hay que ser muy crítico con este aspecto. Una estrategia muy utilizada por la industria para comercializar nuevos medicamentos, es diseñar estudios de pequeño tamaño para valorar eficacia y que, el nuevo medicamento, resulte ser "igual al que se compara" con lo cual el lector poco exigente asume que la eficacia es la misma. Lo único que demuestran estos estudios es lo obvio: con el poder (o capacidad para detectar diferencias) que tiene el estudio (muy bajo) no se detectan diferencias. Con los estudios de seguridad, no queda más remedio que la estrategia contraria: si quiero demostrar que un nuevo antiinflamatorio es menos gastrolesivo que los anteriores, aunque las diferencias

sean pequeñas, la solución está en un megaensayo que alcance significación estadística para pequeñas diferencias clínicas.

Ahora veamos **cómo interpretar la magnitud de las diferencias** encontradas en el estudio. En esto también hay picardía y frecuentemente se presentan los resultados de la manera más impactante o espectacular posible. Habitualmente, en un ensayo clínico, obtenemos el número de pacientes en los que ocurre el acontecimiento adverso (muerte, infarto, crisis dolorosa, etc.) en el grupo tratado (expuesto al medicamento de prueba) y en grupo control (no expuesto) y sabemos el número total de pacientes de cada grupo.

La medida de la frecuencia o del riesgo en cada grupo, la incidencia, o riesgo a secas, es el cociente entre el número de pacientes que presentan el acontecimiento y el número de pacientes en ese grupo. Así obtenemos la incidencia en el grupo de expuestos (tratado) y el de no expuestos (control). Con este valor podemos presentar las diferencia entre grupos, es decir, la eficacia del medicamento que se estudia, de varias maneras.

La medida de la fuerza de la asociación entre exposición (medicamento) y acontecimiento: El cociente o razón entre el riesgo o incidencia en los expuestos y el riesgo o incidencia en los no expuestos se ha llamado razón de riesgos o riesgo relativo (RR). El RR indicará el número de veces que tienen más riesgo de ocurrencia del evento los expuestos en relación a los no expuestos. No tiene unidades y puede tomar los valores entre 0 e infinito. Cuando RR es igual a 1 indica que la exposición no influye sobre la incidencia del evento. Cuando es mayor que 1 indica que la exposición está asociada a una mayor frecuencia de evento; es, por lo tanto, un factor de riesgo. Si RR es menor que 1 indica que la exposición se encuentra asociada a una disminución del riesgo de sufrir el fenómeno; se considera un factor protector. En muchos estudios aparece la llamada *reducción del riesgo relativo* que se halla con el cociente entre el riesgo atribuible (véase más abajo) y el riesgo en el grupo control, multiplicado por 100. El resultado, al expresarse en porcentaje, ofrece cifras impactantes, cuando en realidad es una medida que proporciona poca información.

La medida del beneficio que se obtiene con la intervención. El RR sólo informa de la fuerza de la asociación, y no valora el número de casos que se evitaría si se interviniese. Esto se puede conseguir con el riesgo atribuible (RA) que es la diferencia entre el riesgo o incidencia en los expuestos y el riesgo o incidencia en los no expuestos y por tanto también tiene en cuenta la incidencia del evento, es decir, a mayor frecuencia del evento más casos se pueden prevenir a igualdad de RR. Si la exposición no tiene efecto sobre la incidencia del evento, el RA será igual a 0. Si la exposición es un factor protector, tomará valores negativos y si se trata de un factor de riesgo tomará valores positivos.

Desde el punto de vista de la toma de decisiones es más importante el RA que el RR ya que el RA nos permite conocer el número de casos que podrían evitarse al utilizar el nuevo tratamiento o al instaurar un programa dirigido a reducir la exposición. De esto se deriva el *número necesario de tratamientos* (NNT) que es el número necesario de pacientes que es necesario tratar para prevenir un evento, que corresponde a la inversa del RA.

Pongamos un ejemplo de una intervención, con efecto muy marcado sobre la mortalidad, para ilustrar estos conceptos. El estudio MERIT-HF es un ensayo clínico donde se asignaron aleatoriamente 3991 pacientes con insuficiencia cardíaca de moderada a grave a dos

grupos de tratamiento (metoprolol o placebo) y se valoraba la mortalidad (véase BIT n^o2). 1990 pacientes fueron asignados al grupo de metoprolol y 2001 al grupo placebo. Entre los del grupo de metoprolol murieron 145 pacientes ($145/1990=0.072$ (7.2%) y entre los del grupo placebo 217 ($217/2001=0.1084$ (11%)). El **riesgo relativo** o razón de riesgos es de 0.66 ($0.072/0.1084$). Este valor es menor que 1 por lo que la exposición al metoprolol es un factor protector. El **riesgo atribuible** se halla con la diferencia entre los riesgos de cada grupo ($7.2\%-11\%=-3.8\%$). Este valor indica el número de muertes que podrían evitarse al tomar medidas para tratar a todos los pacientes. Esto se puede expresar como **NNT** ($1/3.8=0.26$), es decir, sería necesario tratar a 26 pacientes para evitar una muerte. Ahora, si lo vemos expresado como **reducción del riesgo relativo** (lo que es frecuente) resultaría en 34.5% ($(3.8/11)\cdot 100$), lo que puede parecer que la intervención reduciría la mortalidad en un aparatoso 34.5%, resultado muy lejano del real, expresado mediante el riesgo atribuible que es 3.8%.

Pero lo que realmente interesa de un estudio no es el valor de estas medidas de efecto en la muestra, sino que interesa **estimar el valor de esas medidas en la población** de la que surgió la muestra. Debido a que es prácticamente imposible que la muestra reproduzca exactamente lo que ocurre en la población (es decir, que la medida de efecto sea la misma en la muestra que en la población), el valor puntual de la medida de efecto de la muestra debe ir acompañado de un intervalo, en el que, con una determinada confianza, se encuentre el verdadero valor de la medida de efecto en la población. Este se denomina *Intervalo de confianza*, y generalmente se plantea al 95%, para corresponderse con un nivel de significación de 0.05 (5%). Se recomienda que en la publicación de los estudios, los resultados se expresen en función de los intervalos de confianza, en vez de únicamente la "p". Al igual que la "p", los intervalos de confianza, nos proporcionan información sobre la significación estadística. Pero, además, nos dan una idea de la precisión del cálculo de la medida del efecto; si el IC es muy amplio, nos indicará que esa medida de efecto es muy poco precisa, por lo que hay que interpretar los resultados con cautela.

Por ejemplo en el MERIT, el riesgo relativo de la mortalidad fue de 0.66 y los intervalos de confianza al 95% entre 0.53 y 0.81, lo que nos informa que, con una confianza del 95%, al utilizar betabloqueadores en la población general con insuficiencia cardíaca, el riesgo relativo de mortalidad se encontrará entre 0.53 y 0.81. Eso, suponiendo que realizáramos el tratamiento de manera ideal, o mejor dicho, de igual manera que en el estudio MERIT. Si en la práctica, no conseguimos ser tan cuidadosos como en el contexto del ensayo, encontraremos peores resultados. Esta será la diferencia entre la **eficacia** encontrada en el ensayo clínico y la **efectividad** de nuestra intervención.

Con todo, la estadística es sólo una herramienta del investigador, que no puede corregir las incorrecciones que pueda tener el diseño del estudio. Un estudio cuyos grupos no son comparables, o cuyos datos clínicos son erróneos, o que no tienen en cuenta posibles factores de confusión, no mejora a base de pruebas estadísticas y "p" significativas. La validez interna del estudio es más importante que la precisión.

Por otra parte, queremos recordar que una búsqueda intensa de subgrupos de pacientes o de variables, acaba proporcionándonos, por una cuestión de probabilidades, alguno/a donde se ha alcanzado significación estadística. Nunca debemos aceptar resultados para los que no se haya diseñado el análisis antes de efectuar el estudio.

AMINOGLUCÓSIDOS: ¿MAYOR EFICACIA Y MENOR TOXICIDAD EN DOSIS ÚNICA DIARIA?

Los aminoglucósidos (AMG) son antibióticos, con efecto bactericida concentración-dependiente y acción sinérgica con otros antibióticos, útiles para el tratamiento de infecciones por enterococos, estafilococos, enterobacterias y pseudomonas, entre otros.

La dosis e intervalo de dosificación de este grupo de antibióticos persigue la obtención de altas concentraciones plasmáticas máximas (en el pico), imprescindible para obtener eficacia, pero evitando que dichas concentraciones se mantengan elevadas mucho tiempo para minimizar su nefro y ototoxicidad. Por ello, clásicamente se han utilizado los AMG cada 8 o 12 horas, monitorizando los niveles plasmáticos con el objetivo de conseguir concentraciones plasmáticas de 2 mg/l en el valle (concentración mínima) y 5-10 mg/l en el pico (concentración máxima) para gentamicina (G) y tobramicina (T) y de 1-8 mg/l en el valle y 20-30 mg/l en el pico para amikacina (AK).

Hace ya más de 10 años que se postuló que, administrar dosis más altas, una sola vez/día, sería más conveniente desde el punto de vista de una mayor (o como mínimo igual) eficacia, al conseguir picos más altos, con una menor nefrotoxicidad ya que se considera que ésta es función del mantenimiento de altas concentraciones a lo largo del tiempo. También tendría un menor coste al ahorrarnos una o dos administraciones al día.

De hecho, la pauta con dosis única diaria (DUD) es la más extendida en los últimos años. En este artículo, analizamos si existen suficientes evidencias que avalen esta práctica y, sobre todo, si esta pauta se utiliza de manera que se optimicen sus supuestas ventajas.

Justificación de la mayor eficacia

Se piensa que la DUD tendría mayor eficacia antibacteriana por tres razones:

1.- Los AMG presentan una actividad bactericida concentración-dependiente, es decir, cuanto mayor sea la concentración plasmática máxima respecto a la concentración mínima inhibitoria (CMI), mayor es la actividad. Esto representa que no es necesario mantener altas concentraciones a lo largo del tiempo, lo que trae aparejado una mayor nefrotoxicidad, sino conseguir un pico de concentraciones plasmáticas alto, muy por encima de la CMI. De hecho, con los AMG se ha observado que para concentraciones bajas (nivel valle/CMI < 10:1) hay relación entre el área bajo la curva (exposición a lo largo del tiempo) y la respuesta y para concentraciones altas (>10:1) hay relación entre el nivel del pico y la respuesta.

Esto es diferente en los AB tiempo-dependientes como los betalactámicos en los que, cuanto mayor es el tiempo que mantienen una concentración por encima de la CMI, mayor es su actividad bactericida. En estos un área bajo la curva elevada es imperativo.

2.- Los AMG tienen un amplio efecto postantibiótico (PAE) es decir, la inhibición del crecimiento bacteriano dura más tiempo del que se podría esperar en función del tiempo que permanece el AMG en el foco. Para concentraciones entre 2-10 veces la CMI se ha observado un PAE de 1-8 horas. Se piensa que un mayor pico de

concentración plasmática originaría un mayor PAE por lo que, en teoría, no parece probable que hubiera excesivo crecimiento bacteriano en los momentos con concentraciones bajas si usamos la pauta DUD.

3.- Los AMG generan resistencia adaptativa en la bacterias (sobre todo en las Gram -); es decir, la presencia del AMG induce cambios en las bacterias que limitan la capacidad de entrada de los AMG en la célula. Se piensa que ésta es la razón por la que se ha observado fracaso cuando se emplean en infusión continua. Esta resistencia desaparece a partir de las cuatro horas y, como es un fenómeno similar al de tolerancia (adaptación), se piensa que cuanto mayor es el tiempo entre dos exposiciones, mayor será la respuesta. Así la DUD induciría una menor resistencia adaptativa en las bacterias que las pautas cada 8-12 horas

¿Menor nefrotoxicidad?

La nefrotoxicidad está en función de la cantidad de AMG que se acumula en las células del córtex renal. Los AMG penetran en las células del túbulo proximal renal y, una vez dentro, disminuyen la actividad de la fosfolipasa y otras enzimas celulares lo que lleva, a la larga, a la destrucción celular. El marcador más precoz es la fosfolipiduria que además es mayor cuanto mayor es la lesión. Después aparecen proteínas, células epiteliales, leucocitos, electrolitos y cilindros en orina.

También se produce lesión del endotelio glomerular lo que disminuye el filtrado y el flujo renal. A los 5-7 días de comenzar el tratamiento se puede detectar fallo renal poliúrico o no oligúrico (incapacidad de concentrar la orina) y aumento plasmático de creatinina y urea. Posteriormente se recupera al suspender el AMG.

Los AMG penetran en las células del túbulo proximal renal por un mecanismo saturable, es decir, a partir de una concentración límite, no aumenta el paso con concentraciones mayores, excepto para la tobramicina que es lineal (aumenta en función de la concentración) y para la amikacina que es saturable a dosis bajas y lineal a dosis altas. Por esto, se supone que las dosis muy altas no serán más tóxicas que las dosis bajas. Al contrario, concentraciones altas y no mantenidas serían menos tóxicas que concentraciones más bajas pero mantenidas. Pero si esto es así, la toxicidad de tobramicina y amikacina sería diferente y relacionada con la dosis cuando esto no ha podido ser demostrado de manera concluyente. Veamos los estudios realizados sobre este tema.

Hay 7 estudios en animales que comparan la pauta DUD con la pauta de 2-3 veces/día analizando nefrotoxicidad: 6 de ellos encuentran menos tóxica la pauta de DUD y 1 no encuentra diferencias. Cuando el AMG es tobramicina hay 2 sin diferencias y 3 que encuentran menos tóxica la pauta DUD.

En dos estudios en seres humanos que iban a ser nefrectomizados se administró a éstos, gentamicina, amikacina o netilmicina y se observó que había mucho menos AMG en córtex renal con pauta DUD que con 2-3 veces/día y que con infusión continua, pero con tobramicina no hubo diferencias lo que confirma el mecanismo fisiopatológico antes mencionado.

Ensayos clínicos

Hemos detectado un total de 31 ensayos clínicos controlados (ECC) y aleatorizados que comparan pautas de DUD con pautas de 2-3 administraciones/día de diferentes AMG con similar dosis total por día y analizan la diferencia en eficacia, nefrotoxicidad u ototoxicidad. También hemos encontrado 9 metaanálisis de esos ECC.

De los 31 ECC, 24 no encuentran diferencias en ninguna de las variables. De los otros 7, todos favorecen la pauta DUD, y dos de los 31 encuentran mejor eficacia, 5/31 menos nefrotoxicidad y 1/31 menos ototoxicidad. Es decir, datos no muy favorables.

Pero todavía hay otras limitaciones de estos datos. El cálculo del número de pacientes necesario para detectar diferencias con las variables que se utilizan en la mayoría de los estudios (mortalidad, curación clínica y bacteriológica –eficacia-, creatinina y aclaramiento de creatinina -nefrotoxicidad- y audiometría de alta frecuencia -ototoxicidad-) nos da una aproximación de 105 pacientes necesarios para detectar curación (eficacia), 630 para nefrotoxicidad y 1170 para ototoxicidad. Tan sólo 11 ECC tenían poder suficiente para detectar diferencias en eficacia, 1 para nefrotoxicidad y ninguno para detectar la diferencia de ototoxicidad. Los dos ECC positivos en mayor eficacia tenían 23 y 60 pacientes y los 5 que detectaron diferencias en nefrotoxicidad tenían 60, 78, 141, 123 y 677 pacientes. En dos estudios se emplearon, además, variables más sensibles de la nefrotoxicidad como la fosfolipiduria y en las dos ocasiones el resultado fue favorable para la pauta DUD.

Cuatro de los 9 metaanálisis encontraron mejor eficacia con la pauta DUD, tres de 9 menor nefrotoxicidad y ninguno menor ototoxicidad

Conclusión

No hay suficientes evidencias que demuestren de manera concluyente una mayor eficacia o una menor toxicidad con la pauta DUD. Los metaanálisis dan resultados contradictorios y el único ensayo de tamaño adecuado, con amikacina en neutropénicos, encuentra una eficacia similar y una menor nefrotoxicidad. Cuando se emplean variables muy sensibles se confirma la menor nefrotoxicidad. Aunque aceptáramos esta menor nefrotoxicidad, las diferencias no serían de gran cuantía. A favor de la pauta DUD también está que es más cómoda y barata pero hay que tener en cuenta que hay muchas indicaciones donde no ha sido evaluada como la endocarditis, las infecciones respiratorias, niños, etc.

En nuestra opinión, la mayor ventaja de la pauta DUD es que nos permitiría aumentar aún más la dosis total diaria en los pacientes que eliminan adecuadamente el AMG sin aumentar la toxicidad. Esto puede ser importante en infecciones con mala penetración del AMG o gérmenes de sensibilidad intermedia. Para ello, sería útil averiguar cómo elimina el AMG cada paciente, mediante

la determinación de niveles plasmáticos (ver punto siguiente), y se forzaría la dosis en los pacientes en que es necesario, en función de mantener aquellos niveles en el valle que garanticen ausencia de nefrotoxicidad.

Monitorización de niveles plasmáticos (MNP)

Los niveles plasmáticos (NP) de AMG deben ser siempre valorados en función de las características del paciente, del germen y su sensibilidad y de la localización de la infección. Los rangos terapéuticos establecidos para las pautas habituales (2-3 veces/día) ya no son válidos para la pauta DUD porque los niveles en el pico son más elevados (es lo que se pretende) y los del valle son indetectables al declinar el nivel a lo largo de 24 horas con una función renal normal. Con esta pauta, niveles en el valle, antes adecuados, de 2 o 1 mg/l son ahora inaceptables pues reflejan que el AMG permanece demasiado tiempo por encima de estos niveles a lo largo del día y por tanto existe un elevado riesgo de toxicidad. De hecho, dos ECC han demostrado que con la pauta DUD el mantener el nivel en el valle <2 o incluso <1 mg/l se acompaña de excesiva toxicidad.

Se han propuesto dos soluciones: un criterio es la medida del área bajo la curva concentración plasmática-tiempo (AUC) que mide el grado de exposición total al AMG y que sustituiría a la falta de esa concentración plasmática de referencia. Un criterio más clínico es utilizar la relación entre AUC y CMI. Se requerirían valores de AUC/CMI por encima de 100 para garantizar eficacia antibacteriana.

Las ventajas de la MNP de AMG con la pauta DUD serían:

- 1) No está bien definida cuál es la dosis ideal máxima en la pauta DUD. Sólo mediante la MNP es posible individualizar esta dosis. Podríamos dar una dosis estándar, de 3-7 mg/kg según función renal y otros factores, determinar el área bajo la curva y comparando con nomogramas basados en las pautas de 2-3 veces/día, ajustar la dosis al máximo posible, obteniendo así un mayor efecto antibacteriano cuando fuera necesario.
- 2) La MNP nos permite conocer cómo elimina el riñón del paciente el AMG, nos permite detectar posible fallo renal por el AMG y ajustar la dosis en función de ésta.
- 3) En el mejor de los casos, la disminución de la nefrotoxicidad utilizando la pauta DUD es pequeña. Es mucho mayor lo que se puede hacer ajustando la dosis mediante MNP.

Por tanto, lo más adecuado, en nuestra opinión, sería la extracción de una muestra para MNP, 30 minutos tras la infusión de 30 minutos de duración (pico o concentración máxima) y otra muestra a las 6-14 horas de la infusión. Con ambos puntos podemos calcular el AUC en ese paciente y realizar el oportuno ajuste de dosis.

Referencias bibliográficas

Primer artículo:

-Br. Med. J. 2001; 322: 1317-18. -Br. Med. J. 2000; 320: 1720-3. -Br. Med. J. 2001; 322: 226-31. -Med. Clin. (Barc)1988; 90: 463-468. -Med. Clin. (Barc)1988; 90: 122-126. -Med. Clin. (Barc)1989; 92: 223-228. -Med. Clin. (Barc)1987; 89: 296-301. -BIT de AP del Sergas 1997; 1 (extraordinario n°3): 1-6

Segundo artículo:

-Ren Fail 1992; 14 (1): 1-9. -Clin Pharmacokinet 1999; 36: 89-98. -Drugs 1996; 52: 344-70. -Am J Hosp Pharm 1994; 51: 2016-21. -Clin Pharmacokinet 1994; 27: 32-48 -Drugs 1994; 47: 902-913. -J Clin Pharmacol 1992; 32: 686-91. -DICP 1991; 25: 153-63. -Rev Infect.Dis 1981; 3: 74-83. -Clin. Pharmacokinet. 2001; 40: 803-14