

## Editorial

**BIOESTADÍSTICA Y ALGO MÁS**

Hace varios años, al hacerme cargo de la edición de la Revista, un poco apremiado por la imposibilidad de negarme y un poco abrumado por la responsabilidad del compromiso adquirido, lo primero que hice fue solicitarle a todos los ex editores incorporarse al Consejo Editorial. Me sentía así acompañado de un grupo de gran experiencia e intelectualizaba al mismo tiempo mis propias facencias para desempeñarme como editor.

Conversando entonces con uno de ellos, el Dr. Jorge Dagnino, con quién trabajáramos juntos durante mis primeros años de formación y con quién mantengo una duradera amistad, al verme un poco afligido me ofreció ayuda incondicional. Le tomé la palabra en forma inmediata. Me acordé de un ejemplar de la Revista dedicado a estadística que me había servido mucho en forma personal, y supongo que a muchos colegas y alumnos de postgrado de la época<sup>1</sup>. Para ser exactos aquella edición de 1997 era la recopilación de una serie de artículos que el Dr. Dagnino había publicado en la Revista a partir 1991.

Creí que era el momento de actualizarla, no porque muchos conceptos hayan cambiado (los promedios, las medianas, las desviaciones estándar, siguen siendo los mismos y  $2 + 2$  siguen siendo aproximadamente 4), sino porque muchas personas que están actualmente en formación, no tuvieron acceso a ese número de la Revista y porque era el momento de ampliar el tema y no restringirlo sólo a conceptos estadísticos duros, sino a cosas prácticas como la forma de presentación de datos, la aplicación de la computación en la investigación, los aspectos éticos de la investigación, niveles de evidencia y grados de recomendación, etc. (entonces la medicina basada en la evidencia estaba aún en ciernes). Y también era el momento de complementar los temas tratados en aquel número de la revista con temas como revisiones sistemáticas o metanálisis, que no eran muy usados en esa época.

Obtuve el compromiso del Dr. Dagnino y es así como, finalmente, damos a la luz este número especial. El material obtenido fue tan vasto, que el resultado final es la edición de 2 números de la revista dedicado a estos temas, en que los artículos originales fueron revisados, perfeccionados y definitivamente reescritos y complementados con la

colaboración de artículos de otros autores: el N°2 del Volumen 43 de la Revista Chilena de Anestesia, dedicado especialmente a estadística descriptiva y no descriptiva, (el N°3 de este año corresponderá a los Temas Libres del XLII Congreso Chileno de Anestesiología) y el N°4 dedicados a otros temas complementarios de la estadística y temas un poco más alejados de la bioestadística y más cercanos a la clínica.

El nombre de estos números especiales fue un motivo de debate. No se va a hablar sólo de estadística o sólo de epidemiología. El término bioestadística no es un neologismo, pues se trata de una rama específica de la estadística que se ocupa de los problemas planteados en el área biológica, de modo que hay que incluirlo, pues acá la mayoría de los ejemplos citados en todos los artículos, son para solucionar problemas médicos o específicamente anestesiológicos. La medicina basada en la evidencia utiliza los datos derivados de ensayos clínicos controlados, estudios de investigación secundarios, investigaciones de farmacovigilancia poscomercialización, metanálisis, revisiones sistemáticas o análisis económicos, pero no es el tema central de estos volúmenes y es un paso posterior a la bioestadística. Finalmente hemos decidido titularlos “Notiones de bioestadística y epidemiología para una lectura crítica de la literatura”, por eso he llamado más simplemente a este editorial “Bioestadística y algo más”.

Contando con la participación de destacados expertos, así como de otros autores que han colaborado desinteresadamente en esta empresa, el editor y coeditor de la revista nos sentimos orgullosos de convocar como editor invitado al Dr. Jorge Dagnino Sepúlveda, quién llegó a la estadística como un autodidacta, por la necesidad que tuvo en su tiempo de entender la literatura anestesiológica, quien ahora es un maestro en este tema como en muchos otros, además de Profesor Titular de la Pontificia Universidad Católica de Chile; él es autor de casi la totalidad de los artículos de este primer volumen y de parte del segundo. Como coeditores también, invitamos a dos personas que han colaborado enormemente en la realización de estos 2 números y aportado significativamente con artículos: el Dr. Dagoberto Ojeda Dinamarca, anestesiólogo de

Clínica Dávila con un postgrado en Estadística en la Universidad de Chile y el Dr. David Torres Pérez, anestesiólogo de Clínica Santa María con un postgrado en Epidemiología en la Universidad de la Frontera, ambos docentes de la Escuela de Medicina de la Universidad de los Andes en sus respectivos campos clínicos. Ellos son autores de un buen número de artículos, han demostrado un gran entusiasmo por esta publicación y son la fiel demostración de que la nueva generación de anestesiólogos ya no se conforma con ser autodidactas.

El principal objetivo de estos números especiales es la comprensión crítica de la literatura médica en general y anestesiológica en particular, así como detectar sus errores más frecuentes. El objetivo no es enseñar la vasta gama de técnicas estadísticas disponible para el análisis de datos, pues hay suficientes textos dedicados a ese propósito tanto de estadística general<sup>2</sup> como de bioestadística<sup>3,4,5,6</sup> y epidemiología<sup>7,8</sup> que pueden ser consultados. Es por esto que tal vez los temas no son desarrollados en toda su extensión y aunque la lectura de estos 2 números de la revista pueda ayudar al diseño de algún trabajo de investigación, tarde o temprano se necesitará el apoyo de un experto para problemas que se van haciendo cada vez más complejos.

Los números y las fórmulas, aunque han sido minimizadas, no pueden estar ausentes en una ciencia exacta como la estadística, si es que así puede llamarse: La UNESCO la clasifica dentro de las ciencias sociales, pero FONDECYT la incluye dentro de la matemática y a ésta en las ciencias exactas y naturales<sup>9</sup>. Afortunadamente ahora hay entre nosotros, los anestesiólogos, muchos más autodidactas como el Dr. Dagnino y muchos más expertos como el Dr. Ojeda o el Dr. Torres que pueden ayudarnos, porque a veces nuestro lenguaje y el de los técnicos o ingenieros estadísticos es muy diferente y puede llegar a ser frustrante.

La gran mayoría de los problemas de los anestesiólogos, tanto de los autores de estudios clínicos como de sus lectores, convergen en la interpretación del análisis estadístico<sup>10</sup>. En la década de los 80, dos prestigiosas revistas de anestesia reconocían que podía haber serios errores estadísticos hasta en el 50% de sus artículos<sup>11</sup>. El asunto se había tratado en realidad mucho antes, en 1976, en una revista médica no anestésica, como el *British Medical Journal*<sup>12</sup> y revisitado en 1994 en un editorial de la misma prestigiosa revista hecho por el estadístico Douglas Altman<sup>13</sup>, bajo el indecoroso título de "*The scandal of poor medical research*". Todos hubiéramos podido pensar que los errores se hubieran superado o al menos minimizado, pero un reciente blog del BMJ<sup>14</sup> parece insistir en que el

problema continúa y que las palabras de hace 20 años de Altman sigue siendo pertinentes: "Necesitamos menos investigación, mejor investigación e investigación hecha con buenas razones. Sería un buen comienzo abandonar el uso del número de publicaciones como una medida de la capacidad".

Haciendo una pequeña revisión, he encontrado que algunos de los errores más frecuentes de los autores son:

- **El poder de un estudio.** Muchas veces no se obtiene diferencia estadística significativa porque el grupo estudiado es muy pequeño. Una muestra pequeña puede llevar a grandes errores en la interpretación de los resultados. El no conseguir encontrar esta diferencia en estos casos se denomina error tipo II. El poder depende de la variabilidad de los grupos estudiados, del presumible efecto verdadero de lo que se quiere demostrar, del tamaño de la muestra y del nivel de significación elegido. Un buen ensayo clínico debe reportar el poder calculado, de modo que al leerlo se pueda evaluar un resultado cuando no es significativamente estadístico.
- **El valor de p.** Uno de los errores más comunes y a la vez más serios es interpretar el valor de  $p > 0,05$  para determinar diferencia entre los grupos. La tendencia es usar un valor de significancia estadística de  $p$  de 0,05 en investigaciones habituales sociológicas, mientras que valores  $p$  de 0,01 se utilizan en investigaciones médicas, en las que cometer un error puede acarrear consecuencias más graves. Independientemente de eso, dos estudios pueden demostrar distintos resultados de  $p$  dependiendo del número de casos estudiados. Las pruebas de hipótesis, como el valor de  $p$ , en ningún caso sirven para probar la veracidad de la hipótesis de trabajo o de la hipótesis alternativa, sino que se limitan a indicar que no hay evidencia suficiente para rechazar la hipótesis nula (en base a un nivel preestablecido de probabilidades). Es así que el valor de  $p$  no es un número mágico; más aun, no es ni confiable ni efectivo. Los resultados pueden ser estadísticamente significativos si se estudia un número de casos suficientes, y por otra parte, lo que no es estadísticamente significativo puede ser clínicamente significativo y viceversa. El análisis de un estudio no puede reducirse sólo a saber si un resultado fue significativo o no significativo. A pesar de todo, muchos anestesiólogos siguen creyendo que la mitad de la bioestadística consiste en demostrar si una  $p$  es o no significativa. Finalmente, pareciera más exacto dar el valor exacto de  $p$  y no limitarse a decir si es  $< 0,05$  o  $< 0,01$ .

- **Intervalo de confianza.** Con el intervalo de confianza ocurre casi lo mismo. Un intervalo de confianza del 95% es un rango razonable de valores consistente con los datos. Intervalos de confianza más amplios son reflejo de una mayor incertidumbre. El ancho del intervalo de confianza, expresada habitualmente como desviación estándar y del tamaño de la muestra, de modo que poblaciones con gran variabilidad o muestras pequeñas pueden conducir a intervalos de confianza mayores. Salvo en circunstancias especiales, en la mayoría de los casos el 95% de confianza corresponde a 1,96 veces el error estándar, de modo que se puede calcular el intervalo de confianza aun cuando el autor de un estudio no lo haya hecho. Este valor es mucho mejor que el valor de  $p$  al tratar de determinar si hay diferencia estadísticamente significativa entre dos grupos.
- **Ajuste para comparaciones múltiples.** Si se hace una investigación y se comparan 20 características diferentes de un paciente, las comparaciones múltiples van produciendo un sesgo que debe ser corregido. Muchos investigadores no realizan esta corrección, obteniendo valores de  $p$  equivocados. Uno de estos ajustes es el de Bonferroni, que también es fácil de hacer; basta con dividir 0,05 por el número de comparaciones que se efectúan. Por ejemplo, si se hacen 4 comparaciones:  $0,05 / 4 = 0,0125$ , este valor será el nivel de significación estadística. Otros ajustes, como el de Newman-Keuls o el método de Scheffé son similares.
- **Combinación de la evidencia de varios artículos.** Es habitual que múltiples artículos sobre un tema lleguen a conclusiones diferentes e incluso opuestas. Es por eso que rara vez la lectura de un artículo cambia una conducta clínica. No se puede combinar los resultados de diferentes estudios sumando casos para llegar a una evidencia definitiva. Para eso existe la técnica del metanálisis que pareciera combinar estudios en la misma forma como si se hiciera un estudio multicéntrico. Probablemente un metanálisis de varios estudios de pocos casos no es equivalente a un gran estudio multicéntrico, pero indudablemente es un instrumento útil. El problema aquí no es la combinación de resultados, sino de artículos, que pueden ser muy dispares en su metodología.
- **Valor de  $r$**  (coeficiente de correlación). Igual que en el caso del valor de  $p$ , el hecho que  $r$  alcance significación estadística, o sea que sea diferente de cero, depende fundamentalmente del tamaño de la muestra. Muestras pequeñas con  $r$  grandes pueden no alcanzar significación estadística y, por el contrario, muestras grandes con  $r$  pequeños pueden alcanzar significación estadística aunque no la tengan clínicamente. Más allá que un simple número y de la  $p$  asociada, el  $r$  exige mayor información. Esto es explicado detalladamente en el artículo "Correlación".
- **Conclusiones.** La inferencia estadística generalmente es limitada y el autor tiende a sacar conclusiones que no corresponden a los resultados de su estudio. Muchas veces es sorprendente, especialmente al escuchar los temas libres en un congreso, la cantidad de conclusiones que se sacan y que no concuerdan o definitivamente no tienen nada que ver con los resultados presentados. Habitualmente, mientras más selectiva sea la conclusión de un estudio, mejor refleja el real aporte de una investigación: sólo debe responderse la pregunta planteada en la hipótesis.

Así como estos, se detectan muchos otros errores en la literatura anestesiológica. Además de los mencionados: definiciones operacionales deficientes, mal manejo del sesgo y variables de confusión, resultados vagos y descontrol con los números, hipótesis errantes y conclusiones erráticas. Saber encontrarlos es tarea de un lector crítico. Esperamos que esta edición contribuya a este propósito.

El esfuerzo de esta edición fue tan grande y el material tan valioso, que sugeriría al Directorio de nuestra Sociedad editarlo como un libro en un solo volumen que pudiera ser adquirido en forma particular, debido al escaso tiraje de nuestra Revista (400 ejemplares) y a que su contenido puede colaborar en la formación de muchos futuros especialistas. Esto podría hacerse con muchos otros temas anestesiológicos para colaborar en la formación continua de los anestesiólogos y a la vez uniformar en parte la tan dispersa formación académica que actualmente tienen nuestros alumnos de postgrado.

**Dr. Ricardo Bustamante Bozzo**  
Editor

**REFERENCIAS**

1. Dagnino J. Bioestadística. Revista Chilena de Anestesia 1997; 1: 5-58.
2. Snedecor GW, Cochran WG. Statistical Methods. Octava Edición. Iowa State University Press, Ames, Iowa 1989.
3. Katz M. Study design and statistical analysis. A practical guide for clinicians. Tercera Edición, Cambridge: CUP, 2010.
4. Norman GR, Streiner DL. Biostatistics. The bare essentials. Tercera Edición, Hamilton: BC Decker, 2008.
5. Anderson B. Methodological Errors in Medical Research. Oxford: Blackwell Scientific Publications. 1990.
6. Armitage P, Berry G: Statistical Methods in Medical Research, Tercera Edición. Oxford:Blackwell Scientific, 1990.
7. Fletcher RH, Fletcher SW. Clinical epidemiology. The essentials. 4th ed. Baltimore: Wolters, Kluwer, Lippincott, Williams & Wilkins; 2005.
8. Rothman KJ. Epidemiology. An Introduction. Oxford OUP, 2002.
9. Ferreiro O, Fernández P. La estadística, una ciencia en la controversia. Revista Universitaria 1988; 25.
10. Beach ML. Statistics for the Clinician. ASA Refresher Courses 2001; 29: 21-29.
11. Avram MJ, Shanks CA, Dykes MH, et al. Statistical methods in anesthesia articles: an evaluation of two American Journals during two six-month periods. Anesth Analg 1985; 64: 607-611.
12. Gore SM, Jones IG, Rytter EC. Misuses of Statistical statistical methods. Critical assessment of articles in BMJ from January to March 1976, BMJ 1977; 1: 85-87.
13. Altman D. The scandal of poor medical research. BMJ 1994; 308: 283.
14. Smith R. Medical Research-still a scandal, en BMJ blog: (Accesado el 20 de octubre de 2014 en: <http://blogs.bmj.com/bmj/2014/01/31/richard-smith-medical-research-still-a-scandal/>).