

Errores garrafales en el uso del muestreo.

Apuntes de una experiencia

Luis Carlos Silva Ayçaguer

Instituto Superior de Ciencias Médicas de La Habana

Profesor Colaborador Asociado de la ENS

Mediante esta nota daré cuenta muy sintéticamente de una reciente experiencia desarrollada en la sede de la Organización Panamericana de la Salud (OPS) en Washington DC, que a mi juicio puede tener interés para los lectores y amigos del Boletín.

Empeñada en el desarrollo de sistemas de vigilancia de las enfermedades crónicas no transmisibles, la OPS indentificó la conveniencia de caracterizar el conocimiento disponible en materia de factores de riesgo de tales dolencias en América Latina y el Caribe. Si bien la descripción no constituye un procedimiento explicativo *per se*, es una forma de investigación biomédica no solo legítima sino ineludiblemente necesaria para pautar las tareas prácticas y las propias líneas futuras de la vigilancia puesto que las conclusiones derivadas de los estudios de prevalencia deberían, en buena lid, trascender lo

meramente cuantitativo y producir juicios útiles con vistas a las acciones correctoras por parte de los diversos actores de salud.

Durante los últimos años se han realizado y publicado cientos de estudios transversales donde se estiman prevalencias de tales factores de riesgo, de modo que en principio se trataría, básicamente, de hacer una mera compilación y resumen de datos procedentes de esos estudios. Sin embargo, aceptar acríticamente ciertos resultados por el solo hecho de que han sido publicados puede conducir a convicciones erróneas, a conclusiones falsas, y a decisiones equivocadas.

A pesar de tratarse de una tarea relativamente simple en lo conceptual, la variabilidad de recursos metodológicos empleados en los estudios abarca áreas tales como diseño muestral empleado, enfoques operacionales, criterios usados para el diag-

nóstico, rigor en la toma del dato primario o formas en que se hacen las estimaciones de las deficiencias que ellos pueden presentar. Tal espectro de posibilidades produce serias preocupaciones en cuanto al valor mismo de los datos que se informan y compromete de manera sustancial la posibilidad de sacar partido adecuado a las comparaciones que pudieran realizarse entre los distintos estudios.

Se impuso así la necesidad de valorar en qué medida la información disponible era suficientemente fidedigna (el dato primario reflejaba lo que realmente ocurría) y precisa (las medidas de resumen obtenidas se aproximaban adecuadamente a los parámetros poblacionales que se estudiaban) como para constituir un aporte informativo susceptible de ser tenido seriamente en cuenta a los efectos de resumir, cotejar, interpretar y

Errores garrafales en el uso del muestreo. Luis Carlos Silva Ayçaguer

valorar productivamente los trabajos realizados.

Ello condujo a su vez a la elaboración de un procedimiento que permitiera evaluar la calidad de la información contenida en informes de este tipo de investigación (generalmente concretados en artículos científicos publicados) y aportar de paso una metodología que contribuyera a estandarizar en el futuro algunos aspectos relacionados con estos esfuerzos, tanto en materia metodológica como en la presentación de informes.

UN INSTRUMENTO EVALUATIVO

En este contexto, empleando diferentes métodos de consenso, un grupo de expertos construyó un instrumento para aquilatar la calidad de los trabajos. Luego de múltiples versiones sucesivamente corregidas tras su aplicación a varias decenas de artículos y tras un proce-

so de validación de criterio (*criterion validity*) y de aspecto (*face validity*), se elaboró un cuestionario vertebrado en torno a 19 puntos concernientes a diversas áreas técnicas (objetivos declarados, diseño muestral, métodos de captación de información, procesamiento de la información, comunicación de resultados, etc.)

Los detalles de la metodología seguida para concretar esta tarea y del resultado conseguido pueden hallarse en el informe técnico correspondiente ¹, recientemente publicado. A los efectos de esta presentación me concentraré sólo en siete de los apartados, aquellos más directamente vinculados con el muestreo. A continuación se enumeran y comentan brevemente estas siete cuestiones:

1. ¿Se expone el diseño muestral aplicado?

No siempre se comprende que el presupuesto metodo-

lógico básico en este tipo de estudios es la calidad de la muestra. Los estudios de prevalencia se centran en la *estimación*, de modo que son teórica y prácticamente diferentes de aquellos cuyas preguntas exigen que se realicen *comparaciones* (con las que usualmente se procura responder a preguntas *explicativas*); y en un problema de estimación, la muestra es un protagonista del método, ya que los sesgos introducidos por los errores cometidos en la fase de selección ya no son susceptibles de enmienda en la etapa de análisis. Resulta medular que se hayan usado procedimientos adecuados de muestreo, los únicos capaces de legitimar el estudio de una población sobre la base de una parte que la represente.

2. ¿Fue la muestra probabilística?

Un paso de crucial relevancia en el proceso histórico de consolidación de la teoría del muestreo se produjo

¹ Silva LC, Orduñez P, Rodríguez P, Robles S (2001) *A tool for assessing the usefulness of prevalence studies done for surveillance purposes: the example of hypertension*. Pan American Journal of Public Health; 10 (3): 152-160.

Errores garrafales en el uso del muestreo. Luis Carlos Silva Ayçaguer

con la introducción del azar en el acto de selección muestral. Ésta es una demanda de máxima importancia, pues los procedimientos probabilísticos —que necesariamente involucran al azar— satisfacen la exigencia intuitiva de eliminar, o por lo menos mitigar, la carga subjetiva que podría influir en la elección de los elementos que se van a examinar (y, por lo tanto, en las conclusiones a que ellos dan lugar). Suele no comprenderse que *sólo cuando el método es probabilístico se podrá solucionar la tarea de medir el grado de precisión con que se realizan las estimaciones.*

3. ¿Se plantea que se quiere una caracterización tanto cuantitativa como cualitativa?

Un verdadero estudio descriptivo debe trascender la simple cuantificación neutra de la realidad. Los «trabajos contemplativos», muchas veces indistinguibles de un informe administrativo que no va más allá que lo que una computadora puede producir y que se detiene

justamente allí donde nuestra cultura y capacidad de interpretación resultan imprescindibles, no constituyen verdadera investigación científica. Ésta solo tiene lugar cuando nace de una vocación crítica, capaz de transformar los resultados numéricos en juicios de valor que desbrocen el camino hacia la toma de decisiones. Es imprescindible que los resultados sean extrapolables y precisos, pero su traducción al plano conceptual, su conversión en valoraciones, exige el concurso de la reflexión científica de los investigadores, que trasciende el muestreo u otros instrumentos estadísticos. Cabe exigir por tanto que el investigador coloque su propósito en un marco conceptual y exponga unos objetivos orientados a la caracterización crítica de la realidad y no al mero conocimiento de unos cuantos números.

4. ¿Se aplicó control de la calidad de los datos primarios?

Se trata de una condición medular para garantizar la

calidad del dato primario y para disminuir sesgos intra e interobservadores. El control de la calidad es un ingrediente metodológico fundamental. Un aspecto clave, por poner un ejemplo, que debe integrar dicho control es la *reencuesta* (idealmente realizada por otro observador usando una submuestra aleatoria).

5. ¿Se hicieron las estimaciones con acuerdo al diseño muestral?

Con mucha frecuencia los diseños no son equiprobabilísticos; o sea, no otorgan igual probabilidad de integrar la muestra global a todos los sujetos de la población. Esto es perfectamente legítimo (y ocasionalmente conveniente), pero en tal caso es imprescindible corregir las estimaciones; es decir, hay que ajustar las estimaciones de las prevalencias mediante *ponderaciones* que corrijan los efectos de haber abandonado la equiprobabilidad. Atribuir a la población total la prevalencia muestral puede constituir en tales casos un costoso error.

Errores garrafales en el uso del muestreo. Luis Carlos Silva Ayçaguer

6. ¿Se comunican los errores de las estimaciones con acuerdo al diseño muestral?

La precisión puede darse indistintamente a través de intervalos de confianza o de modo separado, consignando el error máximo que puede estar afectando a las estimaciones, pero es inaceptable que se prescinda del único recurso que permite enjuiciar el grado en que realmente conocemos el parámetro que quiere estimarse. El Comité Internacional de Editores de Revistas Médicas, en sus universalmente aceptados *Requisitos Uniformes para los Manuscritos Enviados a Revistas Biomédicas* (sección destinada a normar cómo debe manejarse la información estadística) es categórico en este sentido y consigna textualmente: «cuantifique los hallazgos y preséntelos con los indicadores adecuados de medición de error o inseguridad (tal como los intervalos de confianza). Evite confiar exclusivamente en las pruebas de hipótesis estadísticas, tales como el empleo de valores de

p , que no ofrecen información cuantitativa de importancia». Debe advertirse que, con frecuencia se realiza un diseño probabilístico (a veces complejo), pero luego no se calculan errores de muestreo, o se hacen vagas declaraciones en torno a ellos. Es decir, se trabaja con rigor en la fase de diseño, se acude incluso a un especialista con el fin de conseguir que la muestra sea probabilística, pero en la fase de análisis se prescinde del cálculo de errores. Se reducen así los esfuerzos iniciales a un formalismo.

7. ¿Se comunica el marco de extrapolación o se reflexiona en torno a él?

La legitimidad de las extrapolaciones es uno de los temas más espinosos en la práctica. Entre los estudiantes de muestreo es bien conocida la regla que establece que *las inferencias que se produzcan a partir de una muestra han de circunscribirse a la población que fue objeto del acto de selección*. Desde una óptica rigurosa, esto es impecablemente cierto. Sin embargo, un apego

estricto a dicha regla tendría efectos tan paralizantes que en la práctica ella suele pasarse por alto. Hay que reconocer que con frecuencia el método de selección es tal que, por una razón u otra, no todos los individuos de la población objeto de análisis tienen oportunidad de integrar la muestra y, sin embargo, la inferencia objetivamente realizada abarca a toda la población y no sólo a la porción de la que procede la muestra. El grado en que una transgresión como ésa resulte «perdonable» no es por lo general un asunto de naturaleza estadística, sino algo inherente al problema que se aborda: depende de la valoración que hagan los investigadores, basada en su sentido lógico y en su «cultura» sobre el problema, elementos a partir de los que se dirá la última palabra. El nivel de extrapolación adecuado suele ser un problema en que participan por igual las técnicas formales de muestreo y los recursos centrados en el sentido común. Pero lo que no debe aceptarse es que no se realice una reflexión explícita

Errores garrafales en el uso del muestreo. Luis Carlos Silva Ayçaguer

sobre el marco poblacional concreto al cual pueden extenderse los resultados.

UNA VALORACIÓN CRÍTICA

El enfoque arriba expuesto se aplicó para caracterizar a grandes rasgos la realidad informativa vigente a partir de lo que se ha publicado sobre datos de América Latina y el Caribe en los últimos 20 años en materia de un factor de riesgo de alta

relevancia: la hipertensión arterial. Las preguntas se respondieron a partir de lo que fuera comunicado explícitamente en cada trabajo (recuérdese que lo que no se informa es como si no se hubiera hecho).

Una búsqueda bibliográfica en tres bases de datos (PubMed, Medline, Bireme) de artículos sobre prevalencia de hipertensión en países latinoamericanos y caribeños publicados en las últimas dos décadas arrojó la existencia de 58 trabajos, 48 de

los cuales (83%) vieron la luz en fecha posterior a 1990. Los artículos proceden de revistas que se ubican en un amplio recorrido, desde publicaciones de gran prestigio y alto factor de impacto como *Circulation* hasta algunas menos conocidas como la boliviana *Revista Médica de La Paz* o *Acta Andina de Chile*. Hecha la valoración detallada de todos los trabajos, se obtuvieron los siguientes porcentajes de respuestas positivas a las 7 preguntas arriba enumeradas:

Porcentaje de estudios sobre prevalencia de hipertensión en América Latina y el Caribe para siete demandas metodológicas relacionadas con el muestreo.

Artículos publicados en 1980-2000

REQUISITO METODOLÓGICO	PORCENTAJE
Se expone a grandes rasgos el diseño muestral que se empleó	74.1
Fue la muestra probabilística	69.0
Se plantea que se quiere una caracterización tanto cuantitativa como cualitativa	63.8
Se aplicó control de la calidad de los datos primarios	41.4
Se hicieron las estimaciones con acuerdo al diseño muestral	25.9
Se computan los errores de las estimaciones con acuerdo al diseño muestral	10.3
Se comunica el marco de extrapolación o se reflexiona en torno a él	53.4

Errores garrafales en el uso del muestreo. Luis Carlos Silva Ayçaguer

La tabla «habla» por sí misma. A diferencia de lo que ocurre, digamos, en el campo de los ensayos clínicos (donde no es fácil hallar tamaña frecuencia de estudios que, por ejemplo, no operen con un grupo de control), en materia de muestreo se cometen generalizadamente pifias de calibre similar, como pone de manifiesto el hecho de que sólo uno de cada 10 trabajos acompañe las estimaciones de sus errores muestrales (máxime

cuando casi el 70% se basaron en muestras probabilísticas). Una par de pinceladas más: sólo uno de los 58 trabajos cumple todos los requisitos y hay 10 trabajos que no cumplen ninguna de las condiciones.

La naturaleza de las deficiencias encontradas en el orden metodológico y en la forma de articular por escrito la información concerniente a la investigación descriptiva, conduce a subrayar la necesidad de acrisolar el

rigor en relación con aspectos que son en principio fácilmente solubles pero cuya desatención puede, simplemente, invalidar esfuerzos humana y materialmente costosos.