

EPIDEMIOLOGY OR STATISTICS?

Dario Consonni¹

Epidemiology is not just having some data (even a large amount) and doing statistical analysis (even sophisticated one). Unfortunately, still today, we see health professionals carrying the data to the statistician and asking: “here is the data; do the analysis and tell me if the p-value is significant”. This is bad practice. In other situations, health professionals think that epidemiologists and statisticians are superfluous: after learning some statistics in basic textbook or in the web, they pick-up statistical software and do the analysis by themselves. Although this practice may occasionally be successful, serious errors or low quality statistical analysis are the rule. The best way to carry out a scientific study is when health professionals, epidemiologists, and statisticians work together since study conception.

“Epidemiology” and “statistics” are often used interchangeably and the different roles of epidemiologists and statisticians are not always appreciated. An epidemiologist must have a sound knowledge of study design and some notions of data analysis, but not necessarily of advanced statistical techniques. Conversely, a statistician must have a sound knowledge of theory and practice of data analysis, even though s/he might not master very well study design issues. Of course, there are people who know well both epidemiology and statistics. An epidemiological study involves a series of activities carried out to answer a scientific or public health question. We may envision five steps.

Step 1. *Study design* involves a protocol which states *why* the study is done, *who* will be enrolled, *what* information will be collected, *where*, and *when*. This step is crucial, and it's here that epidemiologists can give their valuable contribution. If the study aim is assessment of a *cause-effect* relationship (e.g., to establish if a chemical is carcinogenic), the key issue is *validity* (i.e., the study should be relatively free from have little selection, information, and confounding *biases*), regardless of representativeness (e.g., many carcinogens were discovered in males, but there is no reason to believe that they are not carcinogenic in females). Only when if the study aim is *description* (e.g., the rate of work injuries in a country), *representativeness* becomes the key issue. When possible, the study should also have a fair sample size to be sufficiently *precise* (i.e., with little random error, measured by the width of the confidence interval).

Step 2. In *study conduction* subjects are effectively enrolled and information is collected. Here the objective is to reach a high subjects' *participation* (to avoid selection bias), and measures must be taken to strictly follow the procedures of data collection defined in the protocol (to avoid information bias).

Step 3. *Statistical analysis* is meaningful only if previous steps were successful: sophisticated analyses cannot be a remedy of bad study planning (step 1) and/or conduction (step 2). The analysis should be performed by a statistician, but planning of analyses should be done by the statistician *together* with epidemiologists and health professionals who are expert of the study matter.

Step 4. *Presentation of results*. Too often a study reports only p-values (which focus on “statistical significance”) rather than presenting *measures of disease occurrence* (which duly focus on clinical, biological, or public health significance). These are incidence rates, risks, and odds, along with their *ratios* or *differences* and *confidence intervals*.

Step 5. *Study interpretation* involves discussion of study validity, precision, and results in light of current knowledge. One should strive to resist to the sirens of “statistical significance” (the $p < 0.05$ myth). A statistical test/analysis is like a diagnostic test and the study is like the patient. The health professionals apply many different diagnostic tests, then comes their interpretation in the light of patient's history and other evidence (epicrisis). The same should be when conducting a study: the many tests/analyses done should be interpreted in the light of previous evidence to formulate a final evaluation.

In summary, epidemiology and statistics are two different disciplines, each one with their precise role. To do studies of good quality a strict collaboration between epidemiologists, statisticians, and health professionals expert of study matter is required

¹Medical Epidemiologist in the Epidemiology Unit, Department of Preventive Medicine, Fondazione IRCCS Ca' Granda - Ospedale Maggiore Policlinico, Milan – Italy. Doctor in Medicine.

EPIDEMIOLOGIA OU ESTATÍSTICA?

Dario Consonni¹

Epidemiologia não é apenas, obter dados (mesmo que seja um número considerável) e fazer análise estatística (mesmo que seja sofisticada). Infelizmente, ainda hoje, vemos profissionais da área de saúde levando os dados para um estatístico e dizer: “aqui estão os dados; faça a análise e diga-me se o valor-p é significativo”. Esta é uma prática ruim. Em outras situações, profissionais de saúde pensam que epidemiologistas e estatísticos são supérfluos: após aprender um pouco de estatística num livro básico ou na ‘internet’, eles pegam um software e fazem a análise por si próprios. Embora esta prática possa ter sucesso ocasionalmente, erros consideráveis ou baixa qualidade de análise estatística são a regra. A melhor maneira de realizar um estudo científico é quando profissionais de saúde, epidemiologistas e estatísticos trabalham juntos desde a concepção do estudo.

“Epidemiologia” e “estatística” são frequentemente usados permutavelmente e os papéis dos epidemiologistas e estatísticos não são sempre apreciados. Um epidemiologista deve ter um conhecimento profundo do projeto do estudo e algumas noções de análise de dados, mas não necessariamente conhecimento de técnicas avançadas de estatística. Por outro lado, um estatístico deve ter um conhecimento profundo da teoria e prática da análise de dados, mesmo que ele ^(a) não domine muito bem desenho de estudo. Claro que existem pessoas que conhecem bem ambos: epidemiologia e estatística. Um estudo epidemiológico envolve uma série de atividades realizadas para responder um estudo científico ou questão de saúde pública. Nós podemos visualizar cinco passos.

Etapa 1. Delineamento do estudo envolve um protocolo que estabelece o por quê o estudo é feito, quem estará registrado, quais informações serão coletadas, onde e quando. Esta etapa é crucial e é aqui que os epidemiologistas podem dar a sua valiosa contribuição. Se o objetivo do estudo é uma avaliação da relação de causa-efeito (ex., estabelecer se um produto químico é cancerígeno), o elemento chave é validade (i.e., o estudo deveria ser relativamente livre desta pequena seleção, informação e tendência confusa), independentemente de representatividade (ex., muitas substâncias cancerígenas foram descobertas em homens, mas não há motivos para acreditar que eles não são cancerígenos em mulheres). Somente quando o objetivo do estudo é descrição (ex., o índice de acidentes de trabalho em um país), representatividade torna-se um elemento chave. Quando possível, o estudo deveria ter um tamanho da amostra razoável e ser suficientemente precisa (i.e., com poucos erros mensurados pela largura do intervalo de confiança).

Etapa 2. Em estudos de condução sujeita são efetivamente registrados, e informações coletadas. Aqui o objetivo de alcançar uma alta participação dos sujeitos (para evitar seleção tendenciosa), e providências devem ser tomadas para seguir estritamente os procedimentos de coleta de dados definida no protocolo (para evitar informação tendenciosa).

Etapa 3. Análise Estatística só tem significado se as etapas prévias tiverem sucesso: análises sofisticadas não podem ‘ser’ o remédio para um planejamento de estudos inadequado (etapa 1) e/ou condução (etapa 2). A análise deveria ser executada por um estatístico, mas o planejamento da análise deveria ser feito por um estatístico juntamente com epidemiologistas e profissionais de saúde especialistas na área de estudo.

Etapa 4. Apresentação de resultados. Estudos registram frequentemente demais somente o valor-p (que foca em “significado estatístico”) preferivelmente deveriam apresentar medidas de ocorrência de doença (que foca devidamente em significados clínicos, biológicos, ou de saúde pública). Estes são índices de incidência, riscos e probabilidades, junto com a proporção ou diferenças e intervalos de confiança.

Etapa 5. Interpretação de estudo envolve discussão da validade, precisão e resultados perante o conhecimento atual. Dever-se-ia resistir a tentação de “significados estatísticos” (o mito de $p < 0,05$). Um teste/análise estatística é como um teste diagnóstico e o estudo de como o paciente está. Os profissionais de saúde aplicam muitos testes diagnósticos, então obtêm a interpretação deles perante o histórico do paciente e outras evidências (epicrisis). O mesmo deveria ser quando conduzindo um estudo: os muitos testes/análises feitas deveriam ser interpretadas perante as evidências prévias para formular a avaliação final.

Em suma, epidemiologia e estatística são duas disciplinas diferentes, cada uma com o seu papel preciso. Para elaborar estudos de boa qualidade, é necessário uma colaboração precisa entre epidemiologista, estatísticos e profissionais de saúde especializados no assunto.

¹Médico Epidemiologista na Unidade de Epidemiologia, Departamento de Medicina Preventiva, Fondazione IRCCS Ca' Granda - Ospedale Maggiore Policlinico, Milan – Italy. Doutor em Medicina.

EPIDEMIOLOGÍA O ESTADÍSTICA?

Dario Consonni¹

El campo de la Epidemiología no se traduce solo en obtener datos (mismo que sea un número considerable) y hacer análisis estadístico (mismo que sofisticado). Infelizmente, todavía hoy, es posible mirar profesionales del área de salud llevando datos a un estadístico y diciendo: “aquí están los datos; haz el análisis y dime si el valor-p es significativo”. Este es un ejemplo de una mala práctica.

En otras situaciones, profesionales de salud creen que epidemiólogos y estadísticos son desnecesarios; así, después de investigar un poco de estadística en un libro básico o en internet, empleando un software, pasan ellos mismos a hacer solos el análisis. A pesar de ser posible esta práctica tener suceso ocasionalmente, errores considerables o baja cualidad en el análisis estadístico son la regla. La mejor forma de realizar un estudio científico es cuando profesionales de salud, epidemiólogos y estadísticos trabajan juntos desde la concepción del estudio.

Los términos “epidemiología” y “estadística” son de modo frecuente usados de manera equivocada y los distintos papeles de los epidemiólogos y estadísticos ni siempre son apreciados. Un epidemiólogo hay que tener gran conocimiento del proyecto del estudio, y algunas nociones de análisis de datos, pero no, necesariamente, conocimiento de técnicas avanzadas de estadística. Por otro lado, un estadístico debe tener gran conocimiento de la teoría y práctica del análisis de datos, mismo que no domine completamente el diseño de estudio. Hay personas que conocen bien los dos áreas: epidemiología y estadística. Un estudio epidemiológico abarca una serie de actividades realizadas para contestar a una pregunta científica, o de salud pública. Es posible, entonces, observar cinco etapas.

Etapa 1. Delineación del estudio – involucra un protocolo que establece el porqué del estudio ser hecho, quien estará registrado, cuales informaciones serán recogidas, donde y cuando. Esta etapa es fundamental y en ella epidemiólogos pueden dar valiosa contribución. Si el objetivo del estudio es una evaluación de la relación de causa-efecto (ej., establecer se un producto químico es cancerígeno), el elemento clave es validez (i.e., el estudio debería ser relativamente libre de sesgo de selección, de información y de pérdidas), independientemente de representatividad (ej., muchas sustancias cancerígenas fueron descubiertas en hombres, pero no hay motivos para creer que ellos no son cancerígenos en mujeres). Solo cuando el objetivo del estudio es descripción (ej., el índice de accidentes de trabajo en un país), representatividad se vuelve un elemento clave. Cuando posible, el estudio debe presentar un tamaño de muestra razonable y ser suficientemente precisa (i.e., con pocos errores aleatorios mensurados por el intervalo de confianza).

Etapa 2. En estudios de conducción, los participantes son efectivamente registrados, y las informaciones obtenidas. Aquí el objetivo es alcanzar alta participación de los sujetos (para evitar sesgo de selección), y providencias deben ser tomadas para seguir estrictamente los procedimientos para obtener datos establecidos en el protocolo (para evitar sesgo de información).

Etapa 3. El análisis estadístico solo tiene significado si las etapas anteriores tuvieron suceso; análisis sofisticados no pueden ‘ser’ el remedio para un planeamiento de estudio inadecuado (etapa 1) y/o su conducción (etapa 2). El análisis debería ser ejecutada por un estadístico, pero el planeamiento del análisis debería ser hecho por un estadístico juntamente con epidemiólogos y profesionales de salud especialistas en el área de estudio.

Etapa 4. Presentación de resultados – Estudios registran frecuentemente solo el valor-p (que foca en “significado estadístico”), preferiblemente deberían presentar medidas de ocurrencia de enfermedad (que foca debidamente en significados clínicos, biológicos o de salud pública). Estos son índices de incidencia, riesgos y probabilidades, junto con sus proporciones o diferencias e intervalos de confianza.

Etapa 5. La Interpretación de estudio abarca discusión de la validez, precisión y resultados delante del conocimiento actual. Se debería resistir a la tentación de “significados estadísticos” (el mito de $p < 0,05$). Un test/análisis estadístico es como un test diagnóstico y el estudio de cómo el paciente está. Los profesionales de salud aplican muchos testes diagnósticos diferentes, así obtienen la interpretación de ellos delante del histórico del paciente y otras evidencias (epicrisis). Lo mismo debería ser cuando conduciendo un estudio, los muchos testes/análisis hechos deberían ser interpretados considerando las evidencias para formular la evaluación final.

Así, epidemiología y estadística son dos materias distintas, y cada una tiene su papel preciso. Para elaborar estudios de buena cualidad, es necesaria la colaboración precisa entre epidemiólogo, estadísticos y profesionales de salud especialistas en el asunto.

¹Médico Epidemiólogo en la Unidad de Epidemiología, Departamento de Medicina Preventiva, Fondazione IRCCS Ca' Granda – Ospedale Maggiore Policlinico, Milán – Italia. Doctor en Medicina.