Do you see an elephant or just its trunk? The need of learning Modern Epidemiologic Methods: An introduction

G Babu

Citation

Abstract

INTRODUCTION

There is famous story wherein several blind men were asked what is the elephant is like. While one of them said it is like a pillar, another man said "snake" and each talked of it differently as they felt it. [1] In epidemiology, many of the hypotheses being evaluated in the interpretation of studies can be seen as auxiliary hypothesis in the sense that each blind person is feeling the elephant and describing individual experience. Particularly, each observation is independent of the presence, absence or direction of any causal connection between the study exposure and the disease. Much of the interpretation of epidemiological studies amounts to the testing of such auxiliary explanations for observed associations. [2]

Hence, it is important to understand that all epidemiological studies are only the testing parts of the observed association given that a whole set of factors (sociological, economic, environmental) are acting in the actual causal mechanism. [3] Similar to visualize the elephant by collective feedback, it is more important to understand concepts and principles that govern epidemiology to arrive at logical inferences. [4] This paper aims at introducing some of the modern methods in epidemiology with the objective of elaborating further in future papers.

BACKGROUND

The determinants of health and wellbeing in our environment are complex, interconnected and wide ranging, as the represented in wider determinants of health in Dahlgren and Whitehead's Social Model of Health. [5] Given such complex iterative process, the process of conceptualizing and arriving at logical conclusions in Epidemiology often involves intuition and prior information. However it is evident now, that subjective intuitions can be inaccurate in explaining uncertainties in predicting events. [6,7,8,9]

Hence, most of the inferences required for implementation of any new program or project in the field of public health cannot be provided by epidemiologists alone, but are dependant on further knowledge and explanations from different disciplines such as sociology, economics, environmental sciences, and historians to name a few. [3] This necessitates that we need to incorporate methodology in standard epidemiological training to understand not only concepts in public health but also training how to use such information to provide logical explanations. [3]

EPIDEMIOLOGIC THINKING- SCOPE

Epidemiology has evolved from just another scientific discipline to a professional practice area and as an information science. The following figure conceptualizes uses of Epidemiology and underlines the importance of modern concepts in decision-making and action in the discipline of public health.

Figure 1

Figure.1: Uses of Epidemiology in Public Health Practice and Research
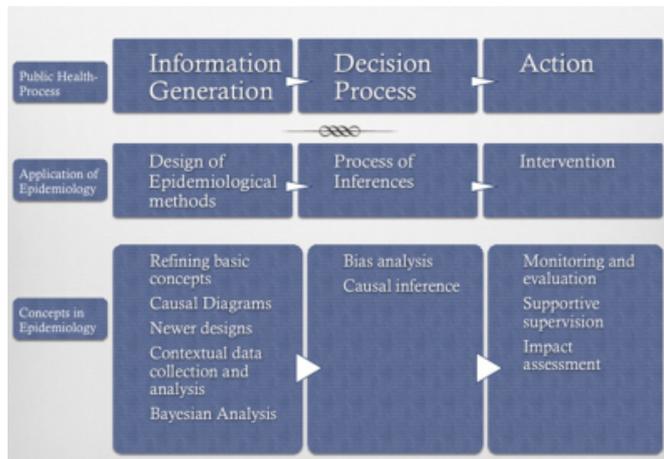
The current paper introduces concepts outlined under the broader concepts of design of epidemiological methods and process of inferences.

DESIGN OF EPIDEMIOLOGICAL METHODS

There are several tools and concepts developed in epidemiology over the past few decades. We are discussing only few of such concepts for the purposes of explaining how these concepts can contribute towards better information generation. It is important that some of basic concepts are understood properly ahead of understanding some of newer concepts in epidemiology.

REFINING BASIC CONCEPTS

The habit of terming any measure as rate and risk is widely in use even among the scientific discussion forums. A famous example is the term "Maternal Mortality Rate", which neither a rate (since there is no person time units in denominator) nor anything to do with measure of incidence. Measures of incidence are incidence time, incidence rate, and incidence proportion (or their ratios and differences) can be used only when one considers an deterministic event without recurrence in a closed population followed until everyone has experienced the event. In all other cases, interpretation of any incidence measures becomes difficult and requires extra assumptions. Authors, editors and reviewers of some scientific journals might have to consider whether naming any measure as incidence is justified under the above conditions.

The influence of temporal ambiguity, reverse causation, length bias and survivor bias will have to be given due attention in drawing any inferences from cross sectional studies.

CAUSAL DIAGRAMS

Directed acyclic graphs (DAG's) have been widely used in epidemiological research for several purposes. [10] Most importantly, they can aid in designing study, ordering temporality, to depict basic biology with diligent attention to epidemiologic principles. [11] [12] DAG's are relatively simple to construct with the use of any line or arrow connecting several variables. There are several rules and assumptions governing the use of DAG as outlined in seminal papers. [10] DAG's help in identification and control of confounding, selection bias and ordering of temporal relations and hence are useful for all study designs. Modern statistical methods such as G-estimation and Propensity score find more utility with relevant DAG's. [13]

NEWER STUDY DESIGNS

Case Control studies have evolved over a period of time and can be even efficient than cohort studies when one adopts modern sampling strategies such as selection of controls as in case cohort study and/or density sampling. The advantage of adopting these sampling techniques is that one can estimate incidence measures (risk ratio and incidence rate ratio) without the need of rare disease assumption. [14] In traditional cumulative case control studies, the study may address a risk factor that ends before the subject selection begins and selection of controls will be from the portion of population that remains after eliminating the accumulated cases.

In case cohort studies, cases are all incident cases in a given risk period while controls are a random sample from the population at risk at the start of the risk period. In density or risk-set case-control study design, each person in the source population has a probability of being selected proportional to his or her own person-time contribution to the denominators of the incidence rates. [14]

Further, case-crossover studies are special type of matched case control studies where we can estimate the effect of a time-varying, short-acting exposure. [14] There are also case only and case-time control study designs, which are very useful in studying interactions in genetic epidemiology.

CONTEXTUAL DATA COLLECTION AND ANALYSIS

Use of prior information and DAG's can help in delineating

several variables while designing the study. Applying the modern principles of data-analysis, attention has to be paid to avoid several misinterpretations. In particular, there are several misinterpretations of P-value for testing the null hypothesis. [15] One of the common misinterpretations of a two-sided P-value is that it represents the probability that the data would show as strong an association as observed or stronger, if the null hypothesis were correct. [15] It is to be noted that the size of a P-value depends on the size of the test statistic, which in turn depends on both the size of the association estimate and standard error of the estimate. [15] Other misinterpretations include that P-value represents the probability of the test hypothesis, the probability of the observed data under the test hypothesis and that the P-value represents the alpha level of a statistical hypothesis test. [15] For a detailed description of correct interpretation of P values, readers should refer to modern methods in epidemiology. [15] There are misinterpretations involving confidence intervals too. [15] It is important to remember that confidence intervals only depict partial uncertainty under the assumption that statistical model involved in the analysis is correct. [16]

BAYESIAN ANALYSIS

Hume and Popper successfully demonstrated that we cannot deductively 'prove' hypothesis while others have argued that the deduction has limited scientific utility because we cannot ensure the truth of all the premises, even if logical argument is valid. Hence theory formation and enumerative induction remain an essential part of scientific explanation. [17] As aids to these processes, modern epidemiology offers deductive methodology of Bayesian probability logic. This methodology translates personal probabilities of the premises of valid arguments into personal probabilities about deductive conclusions [18],[19] and bias analysis, to combine Bayesian and sensitivity-analysis concepts to evaluate the plausibility of alternative explanations.

PROCESS OF INFERENCES

One of the most complicated steps involved in epidemiology is to draw inferences from studies. This is complex iterative process that seeks public health professionals to conceptualize a causal mechanism, given that epidemiological observations can provide crucial tests of competing explanations. We have outlined above that causal diagrams [20] can be used to depict how hypothesized causal networks translate into testable associations. However, understanding bias analysis and causal inference is required

to draw unbiased, justifiable and true inferences.

BIAS ANALYSIS

There is widespread misconception that estimates obtained from large studies are trustworthy. However, large studies only offer protection from random error and hence analyses of systematic errors are required in all studies (whether large or small). [21] Bias analysis involves analysis of unmeasured confounders, misclassification and selection bias. There are several techniques including probabilistic analysis (eg.,Monte-Carlo Sensitivity analysis), Bayesian and semi-Bayesian analysis to perform bias analysis. [22]

CAUSAL INFERENCE

The task of epidemiologists is to test hypotheses, most of which might be difficult to test and hence historically this task has been done based stating and rejecting a null hypotheses. [23] The so-called Hill's causal criteria have also been used to either prove or negate causation. However, as Bradford Hill himself has suggested, these factors were not to be used as prescriptions for either proving or rejecting causal criteria. [24] On the contrary, we often find many publications, which even now use these factors for the same purpose. Among the eight criteria used by Bradford Hill, except temporality none of other can be either sufficient or necessary criterion for determining whether an observed association is causal. [24] The researchers will have to be cautious in drawing causal inferences and this should be based on modern epidemiological principles.

SUMMARY

In this paper, I have introduced some of important and modern concepts of epidemiology, which are useful in designing studies and drawing inferences. In this pursuit, I have explained that these concepts are useful in information generation and decision process involved in public health. This paper does not discuss application of epidemiological principles for measuring intervention for the purposes of brevity and defined scope. It is important that the modern methods of epidemiology are used to practice evidence based public health planning and hence there is an urgent need to inculcate these in teaching public health practitioners. [24]

References

1. Gupta, Mahendranath Chapter V – Vaishnavism and sectarianism – harmony of religions". Kathamrita. Vol. II. Downloaded from http://www.kathamrita.org/kathamrita2/k2sec02.htm
2. Rothman KJ, Greenland S, Poole C, Lash TL. Causation

and Causal Inference. Chapter 2. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 5 – 31.

3. Babu GR, Comment on 'From risk factors to explanation in public health', J Public Health 2008 30: 515-516.

4. McDowell I. From risk factors to explanation in public health. J Public Health 2008;30(3):219-23.

5. Dahlgren, Göran & Margaret Whitehead, Policies and strategies to promote social equity in health, Background document to WHO – Strategy paper for Europe, Institute for Future studies. Copenhagen; OMS, 1992, 62 p. Ilus, Tab

6. Kahneman D, Slovic P, Tversky A. Judgment Under Uncertainty: Heuristics and Biases. New York: Cambridge University Press, 1983.

7. Gilovich T, How We Know What Isn't is So. New York: Free Press, 1993.

8. Piattelli-Palmarini M. Inevitable Illusions. New York: Wiley, 1994.

9. Nakkeeran N. Knowledge, truth, and social reality: An introductory note on qualitative research. Indian J Community Med 2010;35:379-81

10. Greenland S, Pearl J, Robins JR. Causal Diagrams for Epidemiologic Research Epidemiology, Vol. 10, No. 1 (Jan., 1999), pp. 37-48

11. Robins JR. Data, Design, and Background Knowledge in Etiologic Inference. Epidemiology 2001;11:313–320

12. Greenland S, Brumback B. An overview of relations among causal modelling methods. International Journal of Epidemiology 2002;31:1030–1037.

13. Hernan MA, Robins JM. Instruments for Causal Inference An Epidemiologist's Dream?. Epidemiology 2006;17: 360 –372.

14. Rothman KJ, Greenland S, Lash TL. Case-Control Studies. Chapter 8. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 122 – 126.

15. Rothman KJ, Greenland S, Lash TL. Precision and Statistics in Epidemiologic Studies. Chapter 10. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 152 –153.

16. Rothman KJ, Greenland S, Lash TL. Precision and Statistics in Epidemiologic Studies. Chapter 10. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 156–157.

17. Greenland S. Induction versus Popper: Substance versus semantics. Int J Epidemiol 1998;27:543 – 8.

18. Greenland S. Probability logic and probabilistic induction. Epidemiology 1998;9:322 – 32.

19. Greenland S. Bayesian perspectives for epidemiologic research. I. Foundations and basic methods (with comment and reply). Int J Epidemiol 2006;35:765 – 78.

20. Glymour Maria M, Greenland S. Causal diagrams. Chapter 12. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 183 – 209.

21. Greenland S, Lash TL. Bias Analysis. Chapter 19. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 345–346.

22. Greenland S, Lash TL. Bias Analysis. Chapter 19. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 363–380.

23. Rothman KJ, Greenland S, Poole C, Lash TL. Causation and Causal Inference. Chapter 2. In: Rothman KJ, Greenland S, Lash TL (eds). Modern Epidemiology, 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008, 25–31.

24. Babu GR. India's tryst with creation of public health cadre. Ann Trop Med Public Health [serial online] 2011 [cited 2011 Oct 13];4:143-4. Available from: http://www.atmph.org/text.asp?2011/4/2/143/85774

Author Information

Giridhara R Babu

Department of Epidemiology, University of California

Do you see an elephant or just its trunk? The need of learning Modern Epidemiologic Methods: An introduction

5 of 5