



Journal of Postgraduate Medicine

Official Publication of
The Staff Society of the Seth GS Medical College and KEM Hospital, Mumbai, India

October-December 2017 | Volume 63 | Issue 4

www.jpjgmonline.com

Statistical medicine: An emerging medical specialty

Indrayan A

Former Professor &
Head, Department of
Biostatistics and Medical
Informatics, University
College of Medical
Sciences, Delhi, India

Address for correspondence:
Dr. Indrayan A,
E-mail: a.indrayan@gmail.
com

Received : 05-04-2017
Review completed : 08-07-2017
Accepted : 12-08-2017

ABSTRACT

A large number of statistical tools are now used for medical decision in the core activities of diagnosis, treatment and prognosis. These tools provide undeniable help in improving medical outcomes. Prominent among them are uncertainty measurement by probability, medical indicators and indexes, reference ranges, and scoring systems. In addition are tools such as odds ratio, sensitivity, specificity and predictivities, area under the ROC curve, likelihood ratios, and cost-benefit analysis that are commonly applied in medical research but have implications for day-to-day clinical activities. These tools have so completely integrated into medical practice that statistical medicine by itself can stand alone as a medical specialty. Time has arrived to recognize statistical medicine as a medical specialty.

KEY WORDS: Clinical probabilities, medical uncertainties, reference ranges, scoring systems

Introduction

Statistical methods now widely pervade medical thought and practice on the strength of their ability to manage most empirical uncertainties^[1] and such management improves outcome in substantial cases. Yet, surprisingly, statistical medicine has not been proposed as a medical specialty, not even as an idea whose time has come. If academic medicine,^[2] which deals with medical education and not directly with health and disease in individuals or communities, can be accepted and flourish as a discipline, there is no reason that much more direct statistical medicine will not be able to make a respectable place for itself in the course of time, and establish itself as a self-sustained medical subject. Computer medicine arose in the nineties as a specialty although could not sustain the momentum due to unmanageable intricacies, but that is unlikely to happen with statistical medicine. This communication proposes and tries to justify that statistical medicine can be positioned as a specialty by itself that is capable of taking medical decisions in many cases primarily based on statistical arguments. Statistical medicine is not restricted to

the usual rigmarole of using group observations for decisions on individual patients in the hope that the future cases will follow nearly the same pattern but comprises a large number of statistical tools that can determine many medical decisions.

Realize at the outset that this article is not talking about medical statistics that is either understood as medical data as plural, nor about the intricacies of statistical methods applied to medicine as singular whereby inferences are drawn by confidence intervals and tests of significance. This article is also not discussing the details of various statistical methods such as ANOVA, quantitative regression, logistic regression, survival analysis, and meta-analysis that have found extremely useful applications in medical research. Many books and articles are available on these methods, and it would be a wastage of efforts to talk about such methods. This article is also not about statistical fallacies that so frequently occur in medical literature due to misuse and abuse of statistical methods as these also have been talked about at length by many workers and a book^[1] devotes full 32-page chapter on these fallacies. Instead, the focus of this article is the widespread use of statistical arguments that

Access this article online

Quick Response Code:



Website:

www.jpgmonline.com

DOI:

10.4103/jpgm.JPGM_189_17

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Indrayan A. Statistical medicine: An emerging medical specialty. J Postgrad Med 2017;63:252-6.

play a significant role in diagnosis and prognosis, in prevention and control of diseases, in promotion of health, and in medical research. This article proposes to establish that statistical medicine is not just semantically reverse of medical statistics but a leap forward in managing health and disease, both at individual level and at community level. The attempt here is to show that many medical decisions are best taken with the help of statistical tools, and statistical medicine can provide relief to the patients in particular and to the humanity in general. This aspect is seldom realized or appreciated.

Statistical medicine is not a virgin term. It has been loosely used in different contexts, primarily for medical data. According to Braud,^[3] Pierre Louis is remembered as an early ardent proponent of statistical medicine as a key element of his medical observations. This was cited to contrast statistical medicine with intuition, and the term seems to be used primarily for data based inferences. Cinteza and Jinga^[4] acknowledged in a recent article that today's medicine is statistical; although, they think that this would get obituary with personalized medicine. Their usage was to underscore that medical decisions are based on empiricism, but does not propose that statistical medicine can be a science by itself.

Earliest traceable reference to the term “statistical medicine” seems to be in 1822 in the First Report of the Royal Metropolitan Infirmary for Sick Children^[5] that has this title although it presented statistics of medical care. Cartwright^[6] in 1840 described that “statistical medicine furnishes the key which opens to public view in a manner the most convincing, simple and summary the actual results of the regular and empirical practice.” This use too apparently is for conventional medical data. Datta *et al.*^[7] recently emphasized statisticians' expertise in extracting information from data and converting it to medical knowledge and used the term statistical medicine for this process, and Shalabh^[8] used this term for statistical tests such as Chi-square. Mellman^[9] cited statistical fallacies to call statistically based medicine bogus and used the term statistical medicine for such anomalies. This article is not about such uses of the term but is about a subject as explained next.

Medical science is generally understood as a tool to prevent, diagnose, and manage a disease,^[10] where the disease is any aberration that restricts normal functioning, and management includes activities of treatment, disability limitation, and rehabilitation. At the individual level, medicine comprises efforts to put the physical and mental systems back on track when derailment occurs due to injury, pathogens, stresses, degenerations, etc., that our homeostatic system fails to manage. All these efforts produce varying results, and uncertainties remain a prominent constituent. These uncertainties arise from the fact that medicine for the most part evolves from the study of groups rather than isolated individuals, and knowledge gained from the past groups is applied to the future cases in the hope that they would behave in the same manner. Note that this argument is doubly probabilistic – first because group results are applied to individuals, and second because past experience is used on future cases. However, this is just part of the story: the other

part is the increasing use of statistical tools such as scoring systems that by themselves can drive medical decisions.

The foregoing discussion may have given an indication of how statistical methods form the basis of some aspects of medicine because of omnipresent uncertainties. A universally accepted measure of uncertainties is the probability, which is the sheet anchor of the science of statistics. Management of empirical uncertainties in clinical activities and in medical research requires a science that understands randomness and uncertainties – which statistics is known for.

With this background, statistical medicine can be defined as that part of medical science that uses statistical tools and methods to take decisions regarding health and disease in individuals and communities. Thus, it comprises those statistical tools and methods that help to improve the medical outcomes. This is different from clinical epidemiology that uses epidemiological principles for clinical decisions. Clinical epidemiology does require some statistical methods but is not a direct application of the statistical tools to medical care as proposed in this article. The specifics and examples of statistical medicine are as follows, which show how it is practiced.

Statistical Medicine in Clinics

A large number of clinical issues can be cited that are aptly handled with statistical tools. These can be divided into four broad groups – reference values, medical indicators, scoring systems, and probabilities. All these help in diagnosis, treatment, and prognosis – some even define the health conditions and determine the consequences.

Reference ranges

Reference ranges for various medical parameters are used every day in clinics for determining whether a new patient has values well within the normal range, at borderline, or outside permissible limits for healthy persons, and how much – thus whether the patient needs immediate intervention and how much, or can wait and watch.

How do we know that the normal range of, say, total bilirubin levels is 0.20–1.10 mg/dL? Barring few exceptions such as blood pressure (BP) levels and blood sugar levels that have clinical thresholds based on prognostic implications, normal range for quantitative medical measurements is generally established as 2.5th and 97.5th percentiles (mean \pm 2 standard deviation in the case of Gaussian distribution) of the values seen in the healthy segment of the target population.^[11] For example, Prsa *et al.*^[12] recently used this method to obtain reference ranges of parameters that measure blood flow in the major vessels of the normal human fetal circulation at term. Since these limits promise to include 95% values in healthy subjects, anybody crossing these limits is suspected to be not normal and is a candidate to start therapy. Thus, the clinical decision is almost exclusively based on statistical consideration in these cases: complaints, if any, provide supplementary information. This process seems to be doing well as there is hardly any complaint regarding the validity of these limits despite the limitation that

they exclude 5% healthy values equally divided at both the extremes and despite that some non-healthy people may also have values within these limits. Such a possibility of missed diagnosis and misdiagnosis exists with the clinical thresholds as well, since some people with, say, BP >140/90 mmHg may be absolutely healthy and some requiring intervention despite BP ≤140/90 mmHg.

When the ranges of levels in healthy and disease people are available, they will most likely overlap, and the best cutoff can be obtained at the intersection of the two distributions [Figure 1] to minimize the misclassifications. This too is a statistical exercise and does not rule out misdiagnosis and missed diagnosis. The method of receiver operating characteristics (ROC) curves that locates the threshold of a quantitative tests where the sum of sensitivity and specificity is the highest (Youden index) is also statistical, and determines values to be directly used for medical decisions when the area under the ROC curve is high and the sensitivity–specificity at the threshold is satisfactory. Karasahin *et al.*^[13] used Youden index to find a cutoff of 78.31 mg/L of C-reactive protein beyond which 30-day mortality in patients undergoing percutaneous endoscopic gastrostomy increased by about 9 times. Similar is the usage of Z-score and T-score for, say growth assessment of children and for bone mineral density. They assess measurement of a patient in relation to healthy subjects in the population and allow us to take a graded medical decision regarding the kind of intervention required for treatment, including none at all.

Medical indicators and indexes

Although practical usages of the terms indicator and index substantially overlap, an indicator is a univariate quantitative measure of a specific aspect of health, and an index is a meaningful combination of two or more indicators for enhanced context. In this sense, all directly obtained measurements are indicators and those calculated are indexes. Thus, BP level is an indicator and shock index is an index, birth weight is an indicator and birth weight ratio (ratio actual weight to expected weight for gestational age) is an index. Besides an obvious body mass index, measures such as waist-hip ratio and waist-height ratio are indexes for obesity. My review of the medical literature suggests that there might be more than 1000 kind of indexes in use. In many situations, these indexes can be used directly for medical decisions regarding starting or stopping treatment, to discharge or not from the hospital, to warn about grave prognosis, etc. For example, the bispectral index can be used for distinguishing levels of consciousness in severely damaged brain patients^[14] and ankle-brachial index for the detection of

peripheral arterial disease of lower extremities.^[15] Besides their statistical content, all these indexes need to be checked for their reliability and validity for providing usable results – both of which are statistical measures.

Scoring systems for diagnosis and for assessing severity

There is an increasing tendency around the world to depend more on quantitative measurements than on qualitative assessment since this minimizes subjective element in the sense that a score of, say, 7.2 is more than 7.1; although, both may look qualitatively same. Scoring in health and disease requires that qualities are somehow turned into quantities, and unmeasurable continua such as severity of disease are measured on some objective basis. When properly validated, such scores help in reducing some of the epistemic uncertainties^[1] that can arise from the inadequate realization of how much weight is to be given to various pieces of information for correct medical decisions.

Among simplest medical scores is Apgar that assigns quantities to the presence of specific signs, but the most popular possibly is the APACHE score. Glasgow coma scale, Yale observation scale, and peritonitis severity score are the other examples. These scores measure the severity of the condition and are widely accepted guide for appropriate clinical action.

There are a large number of scores for diagnosis also. For example, these have been developed for the diagnosis of benign paroxysmal positional vertigo,^[16] for necrotizing soft-tissue infections,^[17] for chronic lymphocytic leukemia,^[18] and for acute appendicitis.^[19] All such scoring systems are statistical tools, and now increasingly used for diagnosis of diseases and for prognosis assessment, lending credence to our plea that statistical medicine is an appropriate candidate to be considered a medical specialty.

Probabilities in diagnosis, treatment, and prognosis

Be it univariate disease such as anemia, hypertension and diabetes, or multifactorial conditions such as cancer and coronary artery disease, the diagnosis is always a statistical entity, as this is a name to given to statistically more extreme values in one case, and to a cluster of signs-symptoms-measurements that occur more frequently together and follow the same course, in the other case. Variations occur, and the chance remains an integral part in either situation. When a diagnosis is reached on the basis of complaints and physical examination, this is generally only the most likely diagnosis. As the investigation reports become available or the response to the therapy is known, the probability changes under Bayes rule – sometimes even the most likely diagnosis also changes. Bayes rule is indispensable in sensitizing clinicians that probability of complaints in disease, $P(C|D)$, can be very different from the clinically suitable probability of disease in a case with given complaints, $P(D|C)$, depending on the prevalence of the disease. In any case, probability serves as a crucial tool in converting unpredictable uncertainties to predictable uncertainties. We cannot predict individual toss of a coin but can predict that out of 1000 tosses, nearly half will be the head.

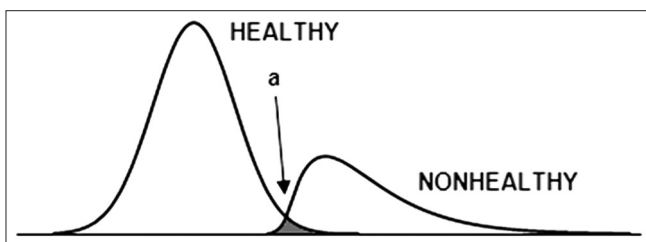


Figure 1: The pattern and overlap of measurement in healthy and nonhealthy subjects (the best cutoff is at the point a)

Sensitivity–specificity, and predictivities for local adoption are entirely statistical considerations that independently validate the medical tests – not only just laboratory and radiological investigations but also signs–symptoms syndromes that form the backbone of diagnosis. However, a clinician has to realize that they can be misleading too. Indrayan^[20] has cited a telling example where the sensitivity and specificity of pap smear is nearly 95% each, but the positive predictivity is only 48% because of the low prevalence of cervical cancer even among those who are screened. This is an extremely useful information for a clinician for application to individual cases since it tells that positive pap smear does not tell so much about the presence of disease in a case as is generally believed, and further investigations are required to confirm or exclude the disease. Sometimes, likelihood ratios are calculated for positive and negative results that measure the utility of a diagnostic test in increasing or decreasing our confidence one way or the other in a suspected case.

When several possible modalities are available for treating a disease, the choice mostly is based on probabilities, since the one that is most likely to provide the best relief to the patient is chosen. This is more so when the treatment is started on the basis of signs–symptoms in situations where the time-elapsing in waiting for confirmation of diagnosis can be hazardous. These uncertainties are not only just due to probability attached to the diagnosis but also due to individual's uncertain response to therapy. All this affects the efficacy of treatment for which we need a measure that can guide. Relative risk and odds ratio have become invaluable tools to measure efficacy of various treatment modalities on one hand and to assess the relative importance of the risk factors on the other. For example, Oresanya *et al.*^[21] used odds ratios and relative risks to find which geriatric pre-operative conditions are more intimately associated with adverse surgical outcomes – a useful result for application in surgeries of older patients.

Being an exercise in predicting the future, the prognosis can never be free of uncertainties. While correlating the spectrum of possible outcomes with the existing state, statistical chances unwittingly play a significant role. Conventional scoring systems such as APACHE ignores process variables such as correct diagnosis, promptness of treatment, type of patient, the attention of medical personnel, and their competency, yet seem to predict prognosis very well. Next generation statistical medicine may incorporate all these and come up with a comprehensive prognostic score for various health conditions.

Most Medical Research Results are Statistical

The objective of the medical research is to devise new medical methods that can be used on individual patients and communities for improved outcomes. Most such efforts require empirical investigations where data do the talking. Overriding role of statistical considerations in research also stems from uncertainties that are an integral part of such medical research framework. This, coupled with the limitation of our knowledge about biological processes, throws indomitable challenge in

reaching to a definitive conclusion. Luckily, statistical methods help discern signals from noise, waves from turbulence, and trends from chaos, despite limitations. Thus, results are obtained that can be confidently used in clinics and communities when the study is carried out with accepted scientific principles.

An essential ingredient in almost all primary medical research is an observation of what goes on naturally, or after a deliberate intervention, but such observations seldom provide infallible evidence. Laboratory experiments outscore over observational studies and clinical trials in providing more valid evidence of cause-effect relationship because of controlled conditions; although, the result is never 100% perfect in this setup also. A clinical trial in any case has a large number of interfering factors that can hardly be taken care of together. Epistemic bottlenecks confound the problem further – thus the results have to be necessarily presented in terms of probability. These factors are even more prominent in epidemiological research in communities and clinics.

Proper study designs, including for clinical and prophylactic trials, improved medical tools, adequate statistical analysis and correct interpretation of results are advised^[1] to control these uncertainties and to come up with a reliable and valid conclusion. All these steps belong to the domain of statistics, and make empirical medical research results primarily statistical in content.

Whereas counterfactuals can be used to disprove a hypothesis, data-based medical research has to pass through the rigors of statistical confidence intervals or tests of hypotheses to be confident that the results are not due to sampling fluctuations, and that they are *prima facie* repeatable. In addition to the Type-I and Type-II errors, fallacies do occur as in any setup, but that does not deter us from moving forward. Furthermore, variation in results propel research synthesis through meta-analysis and systematic reviews, but conclusions here too remain probabilistic than definitive in most cases.

An important area of statistical applications to medical care now emerging is the cost-benefit and cost-effectiveness analysis of various interventions. These analyses help clinicians to determine which intervention could be more acceptable for a given patient. For example, in a small study, Henson *et al.*^[22] reported a potential savings of \$142,822 per month to a hospital in the U.S. by a reduction of 2.9 meticillin-resistant *Staphylococcus aureus* hospital-acquired infection/month associated with polymerase chain reaction screening.

Conclusion

Positive developments are occurring with personalized medicine that could kill statistical medicine at the individual level^[4] but that may take decades to get a firm foothold. Till then, statistical medicine will continue to have a say and deserves to be recognized as a medical specialty. We still have to see whether tools such as scores would find any place in personalized medicine. If yes, statistical medicine would continue to have a definite role in foreseeable future.

Financial support and sponsorship
Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Indrayan A. Medical Biostatistics. 3rd ed. New York: CRC Press; 2012.
2. Association of American Medical Colleges. Planning in Academic Medicine. Available from: <https://www.aamc.org/members/gip/strategicplanning/353164/planninginacademicmedicine.html>. [Last accessed on 2016 Nov 17].
3. Braud HD. Intuition in Medicine: A Philosophical Defense of Clinical Reasoning. Chicago: University of Chicago Press; 2012. p. 107.
4. Cinteza M, Jinga DC. The tomorrow's personalized medicine - The killer of today's statistical medicine. *Maedica (Buchar)* 2014;9:119-20.
5. Anonymous. Statistical Medicine: First report of the royal metropolitan infirmary for sick children. *Lond Med Phys J* 1822;48:120-4. Available from: https://books.google.co.in/books?id=2tiXgqil4QC&dq=Statistical+Medicine:+First+report+of+the+Royal+Metropolitan+Infirmary+for+Sick+Children&source=gbs_navlinks_s. [Last accessed on 2016 Nov 18].
6. Cartwright SA. Remarks on statistical medicine, contrasting the result of the empirical with the regular practice of Physic, in Natchez. *West J Med Surg* 1840;2:1-21. Available from: https://www.books.google.co.in/books?id=2wQDAAAAYAAJ&pg=PA1&dq=Statistical+Medicine++Drake+Yandell&source=gbs_toc_r&cad=3#v=onepage&q=Statistical%20Medicine%20%20Drake%20Yandell&f=false. [Last accessed on 2016 Nov 18].
7. Datta S, Xia XQ, Bhattacharjee S, Jia Z. Advances in statistical medicine. *Comput Math Methods Med* 2014;2014:316153.
8. Shalabh HT. Statistical Analysis of Designed Experiments. Springer; 2006. p. 7.
9. Mellman M. Bogus Statistical Medicine. Available from: <http://www.thehill.com/opinion/mark-mellman/237558-mark-mellman-bogus-statistical-medicine>. [Last accessed on 2016 Nov 17].
10. English Oxford Living Dictionaries. Available from: <https://www.en.oxforddictionaries.com/definition/medicine>. [Last accessed on 2016 Nov 16].
11. Tembe N, Joaquim O, Alfai E, Siteo N, Viegas E, Macovela E, *et al*. Reference values for clinical laboratory parameters in young adults in Maputo, Mozambique. *PLoS One* 2014;9:e97391.
12. Prsa M, Sun L, vanAmerom J, Yoo SJ, Grosse-Wortmann L, Jaeggi E, *et al*. Reference ranges of blood flow in the major vessels of the normal human fetal circulation at term by phase-contrast magnetic resonance imaging. *Circ Cardiovasc Imaging* 2014;7:663-70.
13. Karasahin O, Tasar PT, Timur O, Binici DN, Yilmaz TK, Aslan A, *et al*. High C-reactive protein and low albumin levels predict high 30-day mortality in patients undergoing percutaneous endoscopic gastrostomy. *Gastroenterology Res* 2017;10:172-6.
14. Schnakers C, Ledoux D, Majerus S, Damas P, Damas F, Lambermont B, *et al*. Diagnostic and prognostic use of bispectral index in coma, vegetative state and related disorders. *Brain Inj* 2008;22:926-31.
15. Niazi K, Khan TH, Easley KA. Diagnostic utility of the two methods of ankle brachial index in the detection of peripheral arterial disease of lower extremities. *Catheter Cardiovasc Interv* 2006;68:788-92.
16. Imai T, Higashi-Shingai K, Takimoto Y, Masumura C, Hattori K, Inohara H, *et al*. New scoring system of an interview for the diagnosis of benign paroxysmal positional vertigo. *Acta Otolaryngol* 2016;136:283-8.
17. McGillicuddy EA, Lischuk AW, Schuster KM, Kaplan LJ, Maung A, Lui FY, *et al*. Development of a computed tomography-based scoring system for necrotizing soft-tissue infections. *J Trauma* 2011;70:894-9.
18. Promsuwicha O, Sontgmuang W, Auewarakul CU. Utilization of a scoring system for diagnosis of chronic lymphocytic leukemia in Thai patients. *J Med Assoc Thai* 2011;94 Suppl 1:5232-8.
19. Di Saverio S, Birindelli A, Kelly MD, Catena F, Weber DG, Sartelli M, *et al*. WSES jerusalem guidelines for diagnosis and treatment of acute appendicitis. *World J Emerg Surg* 2016;11:34.
20. Indrayan A. Basic Methods of Medical Research. 3rd ed. Delhi: AITBS Publishers; 2013.
21. Oresanya LB, Lyons WL, Finlayson E. Preoperative assessment of the older patient: A narrative review. *JAMA* 2014;311:2110-20.
22. Henson G, Ghonim E, Swiatlo A, King S, Moore KS, King ST, *et al*. Cost-benefit and effectiveness analysis of rapid testing for MRSA carriage in a hospital setting. *Clin Lab Sci* 2014;27:13-20.