Concept and role of biostatistics as a basic component system of evidential medicine, history of development and value for practice of health protection.

Biostatistics (a portmanteau word made from biology and statistics; sometimes referred to as biometry or biometrics) is the application of statistics to a wide range of topics in biology. The science of biostatistics encompasses the design of biological experiments, especially in medicine and agriculture; the collection, summarization, and analysis of data from those experiments; and the interpretation of, and inference from, the results.
Biostatistics and the history of biological thought

Biostatistical reasoning and modeling were of critical importance to the foundation theories of modern biology. In the early 1900s, after the rediscovery of Mendel's work, the conceptual gaps in understanding between genetics and evolutionary Darwinism led to vigorous debate between biometricians such as Walter Weldon and Karl Pearson and Mendelians such as Charles Davenport, William Bateson and Wilhelm Johannsen. By the 1930s statisticians and models built on statistical reasoning had helped to resolve these differences and to produce the neo-Darwinian modern evolutionary synthesis.

The leading figures in the establishment of this synthesis all relied on statistics and developed its use in biology.

· Sir Ronald A. Fisher developed several basic statistical methods in support of his work "The Genetical Theory of Natural Selection"

· "Sewall G. Wright used statistics in the development of modern population genetics"

· J. B. S Haldane's book, "The Causes of Evolution", reestablished natural selection as the premier mechanism of evolution by explaining it in terms of the mathematical consequences of Mendelian genetics.

These individuals and the work of other biostatisticians, mathematical biologists, and statistically inclined geneticists helped bring together evolutionary biology and genetics into a consistent, coherent whole that could begin to be quantitatively modeled.

Despite the fundamental importance and frequent necessity of statistical reasoning, there may nonetheless have been a tendency among biologists to distrust or deprecate results which are not qualitatively apparent. One anecdote describes Thomas Hunt Morgan banning the Frieden calculator from his department at Caltech, saying "Well, I am like a guy who is prospecting for gold along the banks of the Sacramento River in 1849. With a little intelligence, I can reach down and pick up big nuggets of gold. And as long as I can do that, I'm not going to let any people in my department waste scarce resources in placer mining."[1] Educators are now adjusting their curricula to focus on more quantitative concepts and tools.[2]

Walter Frank Raphael Weldon FRS (15 March 1860, Highgate, London – 13 April 1906), Oxford, generally called Raphael Weldon, was an English evolutionary zoologist and biometrician.

Weldon was the second child of the journalist and industrial chemist, Walter Weldon (FRS 1882), and his wife Anne Cotton. Weldon père moved around the country so frequently that Raphael could not attend school until he was thirteen years old. Walter and Anne had three children; their first child was a girl, with Raphael born next followed by his younger brother Dante.

Raphael did receive some tutoring from a local clergyman before he was thirteen years old then, in 1873, he entered Mr Watson's boarding school at Caversham near Reading. After three years there, plus several months of private study, he entered University College London. Weldon spent the

academic year 1876/1877 at UCL, being taught by the zoologist E. Ray Lankester and the Danish mathematician Olaus Henrici. There he studied a wide range of subjects which he took in preparation for studying medicine. Henrici impressed Weldon more than any other lecturer; he later wrote that Henrici was the first naturally gifted teacher he had studied under.

Later in 1877 he transferred to King's College London and then to St John's College, Cambridge in 1878. There Weldon studied with the developmental morphologist Francis Balfour who influenced him greatly: Weldon gave up his plans for a career in medicine. In 1881 he gained a first-class honours degree in the Natural Science Tripos despite the loss of his brother Dante, who died suddenly. In the autumn he left for the Naples Zoological Station to begin the first of his studies on marine biological organisms.

Weldon married Florence Tebb, daughter of William Tebb of Rede Hall, Burstow in Surrey, on 13 March 1883. She played a large role in his scientific work, assisting him on many of his projects. He died in 1906 of acute pneumonia, and is buried at Holywell Church, Oxford.

Upon returning to Cambridge in 1882, he was appointed university lecturer in Invertebrate Morphology. Weldon's work was centred around the development of a fuller understanding of marine biological phenomena and selective death rates of these organisms.

After graduating he began research, going to Naples where he worked at the Zoological Station. He was appointed a demonstrator in zoology at Cambridge University in 1882, and became a Fellow of St John's College and a university lecturer in invertebrate morphology in 1884. His teaching was described in these glowing terms:-

"Seldom is it given to a man to teach as Weldon taught. He lectured almost as one inspired. His extreme earnestness was only equalled by his lucidity. He awoke enthusiasm even in the dullest, and had the divine gift of compelling interest."

After he was married to Florence, Weldon took all his holidays with his wife in places where they could study marine biology. In particular they visited the Bahamas in 1886, which was scientifically very profitable. The Marine Biological Association set up a laboratory in Plymouth, and Weldon and his wife began spending all their vacations there undertaking research. By 1888 they were spending as much time there as his duties at Cambridge would allow, and he only went to the university to give his lectures. He undertook research June to January, teaching at Cambridge for two terms each year.

In 1889 Weldon succeeded Lankester in the Jodrell Chair of Zoology at University College London, and was elected to the Royal Society in 1890. Royal Society records show his election supporters included the great zoologists of the day: Huxley, Lankester, Poulton, Newton, Flower, Romanes and others.

His interests were changing from morphology to problems in variation and organic correlation. He began using the statistical techniques that Francis Galton had developed for he had come to the view that "the problem of animal evolution is essentially a statistical problem." Weldon began working with his University College colleague, the mathematician Karl Pearson. Their partnership was very important to both men and survived Weldon's move to the Linacre Chair of Zoology at Oxford University in 1899. In the years of their collaboration Pearson laid the foundations of modern statistics. Magnello emphasises this side of Weldon's career. In 1900 he took the DSc degree and as Linacre Professor he also held a Fellowship at Merton College, Oxford.

By 1893 a Royal Society Committee included Weldon, Galton and Karl Pearson 'For the Purpose of conducting Statistical Enquiry into the Variability of Organisms'. In an 1894 paper Some remarks

on variation in plants and animals arising from the work of the Royal Society Committee, Weldon wrote:-

"... the questions raised by the Darwinian hypothesis are purely statistical, and the statistical method is the only one at present obvious by which that hypothesis can be experimentally checked."

In 1900 the work of Gregor Mendel was rediscovered and this precipitated a conflict between Weldon and Pearson on the one side and William Bateson on the other. Bateson, who had been taught by Weldon, took a very strong line against the biometricians. This bitter dispute ranged across substantive issues of the nature of evolution and methodological issues such as the value of the statistical method. Will Provine gives a detailed account of the controversy. The debate lost much of its intensity with the death of Weldon in 1906, though the general debate between the biometricians and the Mendelians continued until the creation of the modern evolutionary synthesis in the 1930s.

Karl Pearson FRS (March 27, 1857 – April 27, 1936[1]) established the discipline of mathematical statistics.[2]

In 1911 he founded the world's first university statistics department at University College London. He was a controversial proponent of eugenics, and a protégé and biographer of Sir Francis Galton.

A sesquicentenary conference was held in London on 23 March 2007, to celebrate the 150th anniversary of his birth.[2]

Carl Pearson, later known as Karl Pearson (1857-1936) was born to William Pearson and Fanny Smith, who had three children, Arthur, Carl and Amy. William Pearson also sired an illegitimate son, Frederick Mockett.

Pearson's mother, née Fanny Smith, came from a family of master mariners who sailed their own ships from Hull; his father read law at Edinburgh and was a successful barrister and Queen's Counsel (QC). William Pearson's father's family came from the North Riding of Yorkshire. The family grave is at Crambe, near York. Its motto, "ERIMUS" means "We shall be", and is also the motto of the Middlesbrough coat-of-arms.

"Carl Pearson" inadvertently became "Karl Pearson" when he enrolled at the University of Heidelberg in 1879, which changed the spelling. He used both variants of his name until 1884 when he finally adopted Karl - supposedly also after Karl Marx, though some argue otherwise.[3] Eventually he became universally known as "KP".

Karl Pearson had two daughters, Sigrid Loetitia Pearson and Helga Sharpe Pearson, and one son, Egon Sharpe Pearson. Egon Pearson became an eminent statistician himself, establishing the Neyman-Pearson lemma. He succeeded his father as head of the Applied Statistics Department at University College.

Karl Pearson was educated privately at University College School, after which he went to King's College, Cambridge in 1876 to study mathematics. He then spent part of 1879 and 1880 studying medieval and 16th century German literature at the universities of Berlin and Heidelberg – in fact, he became sufficiently knowledgeable in this field that he was offered a Germanics post at Kings College, Cambridge.

He graduated from Cambridge University in 1879 as Third Wrangler in the Mathematical Tripos. He then travelled to Germany to study physics at the University of Heidelberg under G H Quincke and metaphysics under Kuno Fischer. He next visited the University of Berlin, where he attended the lectures of the famous physiologist Emil du Bois-Reymond on Darwinism (Emil was a brother

of Paul du Bois-Reymond, the mathematician). Other subjects which he studied in Berlin included Roman Law, taught by Bruns and Mommsen, medieval and 16th century German Literature, and Socialism. He was strongly influenced by the courses he attended at this time and he became sufficiently expert on German literature that he was offered a post in the German Department of Cambridge University. On returning to England in 1880, Pearson first went to Cambridge:- Back in Cambridge, I worked in the engineering shops, but drew up the schedule in Mittel- and Althochdeutsch for the Medieval Languages Tripos.

In his first book, The New Werther, Pearson gives a clear indication of why he studied so many diverse subjects:- I rush from science to philosophy, and from philosophy to our old friends the poets; and then, over-wearied by too much idealism, I fancy I become practical in returning to science. Have you ever attempted to conceive all there is in the world worth knowing - that not one subject in the universe is unworthy of study? The giants of literature, the mysteries of many-dimensional space, the attempts of Boltzmann and Crookes to penetrate Nature's very laboratory, the Kantian theory of the universe, and the latest discoveries in embryology, with their wonderful tales of the development of life - what an immensity beyond our grasp! ... Mankind seems on the verge of a new and glorious discovery. What Newton did to simplify the planetary motions must now be done to unite in one whole the various isolated theories of mathematical physics.

Pearson then returned to London to study law so that he might, like his father, be called to the Bar. Quoting Pearson's own account: Coming to London, I read in chambers in Lincoln's Inn, drew up bills of sale, and was called to the Bar, but varied legal studies by lecturing on heat at Barnes, on Martin Luther at Hampstead, and on Lasalle and Marx on Sundays at revolutionary clubs around Soho.

His next career move was to Inner Temple, where he read law until 1881 (although he never practised). After this, he returned to mathematics, deputizing for the mathematics professor at King's College London in 1881 and for the professor at University College London in 1883. In 1884, he was appointed to the Goldsmid Chair of Applied Mathematics and Mechanics at University College London. 1891 saw him also appointed to the professorship of Geometry at Gresham College; here he met Walter Frank Raphael Weldon, a zoologist who had some interesting problems requiring quantitative solutions. The collaboration, in biometry and evolutionary theory, was a fruitful one and lasted until Weldon died in 1906. Weldon introduced Pearson to Charles Darwin's cousin Francis Galton, who was interested in aspects of evolution such as heredity and eugenics. Pearson became Galton's protégé — his "statistical heir" as some have put it — at times to the verge of hero worship.

After Galton's death in 1911, Pearson embarked on producing his definitive biography—a three-volume tome of narrative, letters, genealogies, commentaries, and photographs—published in 1914, 1924, and 1930, with much of Pearson's own financing paying for their print runs. The biography, done "to satisfy myself and without regard to traditional standards, to the needs of publishers or to the tastes of the reading public", triumphed Galton's life, work, and personal heredity. He predicted that Galton, rather than Charles Darwin, would be remembered as the most prodigious grandson of Erasmus Darwin.

Pearson's thinking underpins many of the 'classical' statistical methods which are in common use today. Some of his main contributions are:

1. Linear regression and correlation - Pearson was instrumental in the development of this theory. One of his classic data sets (originally collected by Galton) involves the regression of sons' height upon that of their fathers'. Pearson built a 3-dimensional model of this data set (which remains in the care of the Statistical Science Department) to illustrate the ideas. The Pearson product-moment correlation coefficient is named after him, and it was the first important effect size measure to be introduced into statistics.

2. Classification of distributions - Pearson's work on classifying probability distributions forms the basis for a lot of modern statistical theory; in particular, the exponential family of distributions underlies the theory of generalized linear models.

3. Pearson's chi-square test - A particular kind of chi-square test, a statistical test of significance.

4. Coefficient of correlation and two coefficients of skewness

---

The basic concepts of medical statistics are:

A statistical aggregate is the common number of units of supervision, taken in the set borders of space and time.

A general statistical aggregate is an aggregate, which includes all units of supervision. For example, all morbidity on the earth.

A selective statistical aggregate is an aggregate, which includes the certain part of units of supervision, but this part is able to represent all general aggregate.

Unit of supervision is every special case of the phenomenon, that is studied, that it is every primary element of aggregate, which belongs to the account (for example, every case of disease, birth, deaths, hospitalizations and others like that). Such registration elements of aggregate divide into attributive (expressed verbally) and quantitative (expressed by a number).

Group properties of statistical totality:

1. Distribution of characteristic (criterion – relative sizes);

2. Average level of index (criterions – Mo-mean, Me-median, arithmetical mean);

3. Variety of characteristic (criterions – lim- limit, am – amplitude, σ – average deviation);

4. Representation (criterions – mM – mistake of average sizes, m% - mistake of relative sizes);

5. Mutual connection between characteristics (criterion – rxy - coefficient of connection.

The important interest of medical statistics is quantitative and qualitative analyses of activity of a treatment-and-prophylactic network, an estimation of this activity through the mechanism of influence on a state of health with the obligatory account of complex influence of different factors.

The stages of statistic investigation.

1st stage – composition of the program and plan of investigation

2nd stage – collection of material

3ed stage – working up of material

4th stage – analysis of material, conclusions, proposals

5th stage – putting into practice

The program of statistical research shows the basic directions of research and information, which it is necessary to collect. The programs are official (medical statistical forms of document) and special (which are folded by a researcher).

The program of statistic investigation consists of the program of material collection, the elaboration program and the program of analysis.

The collection program is the program of statistic observation, the form with the list of signs, that have to be registered (registration' signs) with an indication on whom it will be filled in, that is with a determination of the unit of observation. There are the official programs of material collection to study the health and activity of medical establishments, and special ones, composed by the investigater.

The material elaboration program is the composition of models of tables. There are three types of tables: simple (that give the material bringing together only by one sign); group (that give the material bringing together only by two sign); combinational (connection of three or more signs).

The plan of statistic investigation foresees the organizational elements of work. It consists of: 1) the determination of the object of investigation; 2) the place of investigation; 3) the time of investigation; 4) the volume of investigation; 5) the method of material elaboration (manual, with the help of ECM); 6) the terms of the carrying out of the work; 7) executors; 8) composition of the instructions of the methods of work; 9) carrying out of the seminar for the executors.

The ways of formation of the object of investigation: by the scope of observation are continuous or selective; by the time of observation – flowing or one-moment; by the way of obtaining the information – direct observation, copying, filling in a form.

Registration and accounting medical documents can serve as programs of medic-statistical research.

Medic-statistical research can be complete or selective.

Complete or continuous research covers all observation units.

Selective research covers a representative part of the supervision units, which enables to evaluate phenomenon in whole.

Research is of great importance. The territory strongly influences the results of research.

The next question is the time and the term of the research. Research can last constantly, that is to be current, to be carried out periodically, during certain time or to be one-stage.

Constant researches are: studying of natural movement of the population, periodic — studying of prevalence of chronic diseases, one-stage — population census, fixing of a condition of medical service.

So, after collection of the statistical data their working up is being processed. This process includes the quantitative and qualitative check, coding and grouping of these data. The quantitative check means check of correctness of statistical record of documents, the qualitative is the logic comparison of the data, for example, age and diagnosis, age and employment, growth and weight of a body, etc. Later there is coding. To each quantitative or qualitative characteristic of the phenomenon certain code is given.

Grouping may be a distribution of the data according to the quantitative or qualitative characteristics with the purpose of their analysis. The quantitative characteristics are: age, growth, weight, etc. The qualitative characteristics are background, social status, occupation, disease, etc.

Grouping is simple (according to one characteristic), complex (according to many characteristics, which are combined among themselves) and repeated (grouping before the divided earlier groups with the purpose of deeper studying the phenomenon).
The ways of formation of the statistic integrity you see on the slide.

The stages of development of statistical material are following:

- control /logical and technical/;
- enciphering /code/ of registered signs by numbers, letters of alphabet;
- lay-out of cards on groups for the subaccount or groupment;
- report of material;
- deduction of statistical criteria /indexes/, their graphic image.

RANDOMIZED CONTROLLED TRIALS − AS Element of evidential medicine

Animal experiments

Throughout history animals have played an important role in men's quest for knowledge about himself and his environment. Animal studies have contributed to our knowledge of anatomy, physiology, pathology, microbiology, immunology, genetics, chemotherapy and so many others. At the beginning of this century, Webster in United States and Topley, Wilson and Greenwood in England had carried out classical animal experiments. Their studies centred round inducing epidemics in animals and in studies of herd immunity under laboratory conditions.

More important application of animal experiments have been in (a) experimental reproduction of human disease in animals to confirm aetiological hypotheses and to study the pathogenetic phenomena or mechanisms (b) testing the efficacy of preventive and therapeutic measures such as vaccines and drugs, and (c) completing the natural history of disease. For example, naturally occurring leprosy has been found in armadillos. Data obtained from studying these animals indicate that lepra bacilli might exist outside of humans.

Animal experiments have their own advantages and limitations. The advantages are that the experimental animals can be bred in laboratories and manipulated easily according to the wishes of the investigator. A more important point is that they multiply rapidly and enable the investigators to carry out certain experiments (e.g., genetic experiments) which in human population would take several years and involve many generations. The limitations of animal experiments are that not all human diseases can be reproduced in animals. Secondly, all the conclusions derived from animal experiments may not be strictly applicable to human beings. An excellent example to illustrate this point is the WHO trial of typhoid vaccine in Yugoslavia in the mid-1950s. Laboratory tests in animals showed the alcohol-killed and preserved vaccine to be more effective than the traditional heat-killed phenol-preserved vaccine. But randomized controlled trials in human beings demonstrated that, contrary to laboratory evidence, the alcohol-preserved vaccine was found to be less than half as effective in preventing typhoid fever as the traditional phenol-preserved vaccine introduced by Almorth Wright. This highlights the difficulties encountered in extrapolating findings from animal experiments in man.

Human experiments

Human experiments will always be needed to investigate disease aetiology and to evaluate the preventive and therapeutic measures. These studies are even more essential in the investigation of diseases that cannot be reproduced in animals.

Historically, in 1747, James Lind performed a human experiment (clinical trial) in which he added different substances to diet of 12 soldiers who were suffering from scurvy. He divided his patients into 6 pairs and supplemented the diets of each pair with cider, elixir vitriol, vinegar, sea water; a mixture of nutmeg, garlic, mustard and tamarind in barley water; and two oranges and one lemon daily. All the subjects were studied for 6 days. At the end of 6 days the LIMEYS recovered from scurvy and were found fit for duty. Then came Edward Jenner's experiment with cowpox in 1796. Other classical experiments are Finlay and Reed's experiments (1881-1900) to elucidate the mosquito-borne nature of yellow fever and Goldberger's classical experiments in 1915 inducing pellagra by diets deficient in nicotinic acid, thereby proving pellagra to be a nutritional deficiency disease, not an infectious disease as was then supposed. Since then, human beings have participated in studies of malaria, syphilis, hepatitis, measles, polio and others. These experiments have played decisive roles in investigating disease aetiology and in testing preventive and therapeutic measures.

Although the experimental method is unquestionably the most incisive approach to scientific problem, ethical and logistic considerations often prevent its application to the study of disease in humans. Therefore, before launching human experiments, the benefits of the experiment have to be weighed against risks involved. The volunteers should be made fully aware of all possible consequences of the experiment. Thus when an illness is fatal (e.g., excessive haemorrhage) and the benefit of treatment (e.g., blood transfusion) is self-evident, it would be ethically unacceptable to prove or disprove the therapeutic value of blood transfusion. However, such instances represent only a small part of the total research effort. On the other hand, in the present era of scientific medicine, many unscientific or scientifically unsound procedures are still being carried out. For instance, in the study of prescription drugs, a panel of experts in USA found that only 23 per cent of some 16,000 drugs could be classified unequivocally as "effective" (36). It is now conceded that it is equally unethical if a drug or procedure is brought into general use without establishing its effectiveness by controlled trials. The thalidomide disaster and the occurrence of carcinoma of the vagina in the offspring of pregnant women treated with diethylstilbestrol highlight the unfortunate consequence of therapy on the basis of uncontrolled observations. The WHO in 1980 has laid down a strict code of practice in connection with human trials.

Experimental studies are of two types:

a. Randomized controlled trials (i.e., those involving a process of random allocation); and

b. Non-randomized or "non-experimental" trials (i.e., those departing from strict randomization for practical purposes, but in such a manner that non-randomization does not seriously affect the theoretical basis of conclusions).

RANDOMIZED CONTROLLED TRIALS

Too often physicians are guided in their daily work by clinical impressions of their own or their teachers. These impressions, particularly, when they are incorporated in textbooks and repeatedly quoted by reputed teachers and their students acquire authority, just as if they were proved facts. Similarly many public health measures are introduced on the basis of assumed benefits without subjecting them to rigorous testing. The history of medicine amply illustrates this. For instance, it took centuries before therapeutic blood letting and drastic purging were abandoned by the medical profession.

The design of a randomized controlled trial is given in Figure 1. For new programmes or new therapies, the RCT is the No.l method of evaluation. The basic steps in conducting a RCT include the following:

1. Drawing up a protocol

2. Selecting reference and experimental populations

3. Randomization

4. Manipulation or intervention

5. Follow-up

6. Assessment of outcome