Medical statistics and epidemiology – the early history

Aina Schiotz

University of Bergen, Department of Global Public Health and Primary Care, Bergen, Norway
e-mail: aina.schiøtz@uib.no

In the 1850s and the 1860s, Eilert Sundt (1817-75), theologian and the pioneer of Norwegian social science, issued a number of papers about various conditions in Norway. Many of these papers, which focus on hygiene, mortality, nutrition, poverty, etc., have provided a broader understanding of public health in Norway and have been important reference works in Norwegian social science and historical research (1). Sundt was clearly inspired by European traditions, but he was also innovative on two different levels. First, he combined analyses of large amounts of quantitative numerical data with profound qualitative studies. He did so by referring to experiences from his travels throughout the country, from conversations he had with people, and from observations of the society and everyday life in different geographical areas and at different times. His intention was to find out what ‘was going on behind my numbers’ (1 (1975/1869, p. 155)). Second, he introduced concepts inspired by the natural sciences about laws and regularities in the study of society. According to Sundt, regularity was not caused by natural conditions acting equally on all individuals or by human destiny being somehow predetermined from birth. The individual was more likely to be influenced by the society’s common rules and opinions, by the culture, and it was this that caused certain regularities (2).

Norwegian statistics were subsequently influenced by these ways of knowing. Anders Nicolai Kiær (1838-1919), who was the first director of Statistics Norway (an independent entity since 1876), argued that statisticians had to have a thorough non-statistical knowledge of whatever they were going to describe and interpret. If they knew nothing about the conditions that lay behind the numbers, they ended up with tables that were of little worth. ‘For that which lies behind the numbers, the pulsating life and the driving forces, is as a rule unintelligible to anyone who does not know and understand the development at close hand so that they are able to explain it’, wrote Director Kiær in 1903 (2, p. 69). This is a type of insight that statistics and epidemiology in later years have been accused of having neglected.

**MEDICAL STATISTICS – THE EARLY PHASE**

Neither in Europe nor in Norway was statistics a new field in Sundt's day. It already had long traditions. The development of the field, which was closely linked to the origins of the early modern state in the 17th and 18th centuries, involved systematic gathering of demographic and economic data about the state of the realm. Numbers in statistics are an instrument with which to dissect and diagnose the society and to prescribe therapy, says the historian of ideas Karin Johannisson. The numbers were supposed to both explain and transform (3). Starting in the 17th century, a couple of Englishmen, the merchant John Graunt (1620-1674) and the economist, philosopher and professor of anatomy William Petty (1623-1687), were among the first who began to draw up reports about births and deaths, especially infant mortality and mortality from various diseases. One of their goals was to prevent disease if it was economically profitable for the state to do so; another was to estimate the worth of a human being according to her or his age (4,5). Still another work that is worth noting is the Italian physician Bernardino Ramazzini's (1633-1714) book on *De Morbis Artificum Diatriba* (*Diseases of Workers*), which was first published in 1700. Admittedly, this book does not deal with statistics, but it has nevertheless been a cornerstone of occupational medicine, an academic field with strong statistical and epidemiological traditions. Ramazzini considered a large number of occupational groups and examined the kinds of diseases to which they were exposed, including a number of ‘learned men's diseases’. Not surprisingly, he called attention to the unhealthy effect of sitting still too much. Beyond this, he depicted the workplaces, relevant health issues for workers and useful remedies, and he offered good advice (6).

What about Denmark-Norway and Sweden? Statistics in these countries have German roots. The actual concept of *statistics*, which replaced the concept of *political arithmetic*, was first introduced in 1749 by the German professor of philosophy and law, Gottfried Aachenwall (1719-72), and at the time it was generally understood to be political science. Its aim was to depict all political, social and economic conditions in the state (4,7). The aspect of utility played a key role in what was called the ‘mercantile’ government ideology, and in keeping with this the early German statisticians put considerable emphasis on how saving lives could contribute to population growth and a large, powerful state. On a world basis, the Swedes were the first country to institutionalise statistics. In 1749, *Tabellverket* (the Census in Sweden) was established, and here again the question of the state of health in the realm was introduced early as a basis for population growth (3,4). In Denmark-Norway, we got a similar institution when *Tabellkontoret* (a statistical...
office) was established in Rentekammeret (an organ in charge of the finances of the state) in Copenhagen in 1797. With this office, ‘the project of a technology of population’, as Michel Foucault called it, began to take shape. We got demographic estimates with calculations of age pyramids, life expectancy and mortality, estimates that were utilised to govern and mould the people and the society (8, p. 278).

Norwegian medical statistics took shape early in the 19th century. In 1804 all doctors in Norway were ordered to submit annual reports on the state of health in their district. Among other things, these reports, which were initially supposed to be sent to the Health Board (Sundheds-collegiet) in Copenhagen and later to the Norwegian central government administration, were supposed to provide an overview of ‘the diseases that have been most prevalent, especially small pox, scabies, venereal diseases and in Norway the ‘råde’ disease (radesyke – probably a sexually transmitted skin disease, but it might include several diagnoses such as leprosy, scabies etc.)’ and likewise an overview of mortality ‘especially of infants and post-natal mothers’ (9, p. 36, 10, for more information about radesyke see 11). These reports were called ‘Medical reports’ (Medicinalberetninger), and starting in 1853, when overviews for each county were compiled, they were called ‘Reports on the State of Health and Medical Conditions in Norway (Beretninger om Sundhedsstillstanden og Medicinalforholdene i Norge). These reports have formed the basis for the health policy efforts in Norway all the way up to our own day. As historical sources they are invaluable, not least in the research on public health and preventative issues (cf. 9,12,13). Norwegian doctors gathered and analysed data material in another field as well. These studies, which were initiated in the 1820s, dealt with the mentally ill and the conditions under which they lived. In order to understand mental illness and to be able to make arrangements for adequate therapy, there was a need to approach the question from a scientific point of view, argued the physician, Frederik Holst (1791-1871). In this case, three principles applied: 1. a census must be taken of those who were ill, 2. a classification and diagnostic system had to be drawn up, and 3. those who were ill had to be observed and checked over a period of time. The objective was to ‘study the nature of these evils, to determine the ways to cure them and to develop purer and clearer concepts with regard to that,’ as Holst stated in 1828 (quoted from 14, p. 283, with regard to psychiatry, see 2).

ANALYTICAL STATISTICS

What we can call analytical statistics in medicine was first developed by the French doctor Pierre Charles Alexandre Louis (1787-1872) in the 1820s and 1830s. Louis’s contribution to the numerical or mathematical method was based on the clinic and the treatment of patients and not on large demographic groups as in epidemiology, introduced a few years later. Louis is best known for his evaluation of the use of phlebotomy/bloodletting in the treatment of pneumonia, where he found that the method was unfit. Here as in other studies, he observed and noted clinical facts from a large body of patient data and from experiences with autopsies and mortality rates among patients. Based on an analysis of the gathered data, he made cautious generalisations (15). In so doing, he argued that it was not possible to utilise a therapeutic method without risk unless it could be demonstrated that it had a general effect in similar cases, and ‘… therefore I conceive that without the aid of statistics nothing like real medical science is possible’ (16, p. 489). Louis would have a significant impact on the development of the mathematical method in clinical medicine, but also on the coming field of epidemiology (cf. later).

Another well-known personage who has made a name for herself in the field of medical statistics is Florence Nightingale (1820-1910), not because she developed statistical methods as an adviser in clinical research, but because of her use of statistics as a political tool. Her efforts in the establishment of modern nursing are well-known, but she had many other interests. She made a pioneering effort in the development of hospital architecture and not least in medical statistics. ‘Where other young ladies might find pleasure in a novel, she found greater enjoyment in studying a book of statistical tables’, says her biographer Mark Bostridge (17, p. 171). Nightingale lived in an age of strong Christian belief, and she herself put statistics in a religious context. ‘We learn the purpose of God by studying statistics’, she argued (18, p. 20; cf. 17 as well). Thus, statistical studies were both a religious duty and a moral imperative. In her opinion, statistics were ‘the most important science in the world’, for it was the only one that could give the true picture of human experience (18, p. 17; 19). Thus, she understood quite early how to use statistics as evidence in the efforts to get politicians and other people in power to give priority to hospital reforms and public-health measures. In particular, she developed statistical surveys that showed how large a percentage of the mortality in hospitals was due to poor hygienic conditions and a lack of administration and leadership. She was also among the first to develop visual aids, such as diagrams and maps, in order to be able to present the data clearly and distinctly. From the 1870s Nightingale and her close friend and Oxford-professor Benjamin Jowett (1817-1893), discussed the possible foundation at Oxford University of a Chair in Statistics, and were even willing to endow £4000 for the purpose. This, according to Nightingale, would educate future politicians and civil servants in the use of statistics. In the end Oxford was reluctant to establish such a Chair (17, pp. 510-11).

A third area within medical statistics that was developed in the 1840s and 1850s deals with medical meteorology. Interest in the relationship between climate and the outbreak of disease, especially epidemics, has a long pre-history. We find examples of this dating all the way back to the Hippocratic texts. By the early 19th
century, interest in this matter was increasing, and many people were convinced that the weather or a special atmospheric condition were the most important cause of the epidemic outbreak of disease. The Danish medical historian, Morten Skydsgaard, claims that studies of this relationship were usually based at the outset on the doctors' qualitative assessments of the local climate (20).

A clear indication that people were also concerned with this issue in Norway is the ordinance sent out by the government in 1830 to the nation's doctors. In Section 2 of that ordinance, they request that the doctors present information in their medical reports about 'the influence that the climate is thought to have had on the state of health of the general population and on the propagation of certain indicated diseases in particular, if possible meteorological observations of the barometer, thermometer and hygrometer that were made for the year should be reported as well, and thereafter data on the wind, the number of clear, dark and damp days, etc…' (21, p. 163).

In Danish medicine they were increasingly concerned in the 1840s and 1850s with anchoring the study of the possible relationship between climate and disease in statistical analyses of large amounts of numerical data. In this case as well, the inspiration from French statistical theory had a strong impact. Through various surveys, however, the Danes concluded that it was highly doubtful that the weather could be used as a fundamental explanatory mechanism in epidemic theory. Simply stated, their conclusion was that there was no connection whatsoever between climate and epidemic disease. In the 1850s and 1860s, other theories of disease became more popular. Thus, medical meteorology was marginalised, said Skydsgaard. However, the interest in it has never disappeared, and in our age, there are many examples indicating that people are still concerned with this matter. For example, there are many studies that show a clear relationship between seasonal influenza and climate (cf. for example 22).

**Epidemiology – the early phase**

The science of epidemiology, which took shape in the wake of the great epidemics of the 1800s, developed analytical tools and statistical methods in order to trace the health and disease profile in small or large population groups. From a modest beginning in the 19th century, epidemiology developed in four phases, but phases that have overlapped one other: 1. *The Era of Sanitary Epidemiology* from the 1850s. 2. *The Era of Infectious Disease Epidemiology* from the 1880s. 3. *The Era of Chronic Disease Epidemiology* from the 1950s. 4. *The Era of Risk Factor Epidemiology* from the 1980s. This branch of medicine has likewise been strongly influenced by the society of which it has been a part at any given time.

From the middle of the nineteenth century, epidemiology has been a part of science-based medicine, equivalent to the experimental disciplines. It was Pierre Louis's pupils who first formulated the theoretical basis for the field. The pioneers were the Frenchman Louis René Villermé (1782-1863) and the Brit William Farr (1807-83), the latter of whom was a close partner of Florence Nightingale. In 1839, Farr was appointed the Compiler of Abstracts to the Registrar General, which was a landmark in the development of medical statistics and the preventative efforts in England (23). In detail he analysed the mortality data that was gathered and compared it with other mortality patterns for various diseases in different geographical areas. In this way, he called attention to the connection between socio-economic class and disease – the phenomenon that was so striking and urgent during the great waves of industrialisation and urbanisation in Great Britain and on the continent. The British historian Dorothy Porter wrote that what Farr and Villermé achieved was to expose that in urban, industrial societies, death had become a definitive 'social disease'. Both of them were concerned with preventing disease, but they were not proponents of the intervention of the state; for if the state should intervene with laws and regulations, it would undermine individual freedom and initiative. It was the manufacturers who were responsible for improving the health and living conditions of the workers, and to teach them 'moral habits' through 'Christian example' and instruction. However, Farr and Villermé argued that the poor also had a responsibility. Thus, they reflected the libertarian ideology that dominated large parts of Europe in the first half of the nineteenth century (24,25).

Farr and Villermé were concerned with prevention, but their approach contrasted with the view of those who founded The Epidemiological Society of London in 1850. These men were influenced by the wave of revolutions that swept over Europe in the middle of the nineteenth century, when workers and intellectuals demanded political, social and economic justice. In keeping with Florence Nightingale – who admittedly was no revolutionary – the new epidemiologists regarded the field as a political tool and an instrument for governing. It was not enough to just demonstrate and analyse the diseases and mortality, but attention should also be called to the unworthy conditions in which the underprivileged classes of society lived. Using this as a basis, they would fight to improve the health of the workers through government measures. That should be achieved through the development of remedies such as water supply, sanitation, control of food and beverages, healthy working environments, clean streets, hygienic dwellings, etc. Thus, the period from 1850 up to the last decades of the century was given the name of *The Era of Sanitary Statistics*. These efforts were based on the miasmatic theory of disease or the cause and effect paradigm, i.e. the idea that disease was caused by poor hygienic, climatic and geological conditions (24,26).

In Norway, we can date the first epidemiological study to the middle of the 1800s when a doctor analysed the distribution pattern of cholera during the great epidemic of 1848-49 in Bergen (4,9). Yet the actual breakthrough came in 1856 with the establishment of the
world’s first patient register, the National Leprosy Register (now listed in UNESCO’s Memory of the World Register). This too occurred in Bergen. The objective was to find the cause of the disease by mapping it out, making the data available for statistical analysis and monitoring the individual patient. In this way they would be able to obtain a better overview of a disease that was not only mysterious, but that also kept spreading to more and more new areas in Norway, and this would make it possible to prevent and easier to select the patients to whom they wanted to offer treatment. Finally, it would enable them to estimate the future demand for hospital beds. In this way, the interest in epidemiological examination and medical treatment were combined with the governing powers’ need for planning and control (27,28).

Yet what have they found, and what does such a register entail? For the period 1856-1970, the register provided extensive information about 8,231 patients in Norway. They found a concentration in Western Norway, and the hardest hit were the municipalities of Askvoll, Kinn and Naustdal in the Sunnfjord region. There was a rapid decline in new instances of disease in the last decades of the nineteenth century. The mortality was high and on average the patients lived 25 years less than their contemporaries. They could expect to live for five more years after being contaminated by the disease. All in all, they estimate that over 10,000 people died of leprosy (27). If we look behind the numbers we know that the disease was painful for both body and soul. It disfigured the patients and made them invalids; it stigmatised them and isolated them from family and friends; it deprived them of a free and open life. Yet the data both explains and transforms. The Leprosy Register called attention to the disease in a completely new way. You could say it transformed the disease into a national and public matter, and to an even greater extent into a medical issue. The Register was a powerful tool in the hands of physicians and the governing powers. In the struggle against the disease, it was easier now both to control and to mould the people whom it debilitated. That was accomplished through clinical and epidemiological research, hospitalisation, isolation, legislative regulation, authoritative orders and requirements (28,9).

**TOWARD A NEW EPIDEMIOLOGY**

In an international health exhibition in 1884, the president of The London Epidemiological Society formulated the organisation’s objectives as follows:

To watch pestilences; to study their mysterious ways, movements and changes, which are so often quite inscrutable even to the most experienced and learned; to become acquainted with their natural history; to track them step for step, as the hunter tracks the tiger and the wolf in all their concealments and devious lurkings, and thus to anticipate their attacks and discover means for their avoidance (cited from ref. 29, p. 199).

The historian Anne Hardy notes that this statement reflects much of the objective and the robustness of traditional British epidemiology all the way up to the turn of the 20th century. This was based on the major field research studies and the observational and deductive method. Likewise, they were interested in time, space and class – i.e. changes over a period of time, regional differences and social disparities in morbidity and mortality. In many ways, this contrasted with the developments in the USA and Germany, which were influenced to a greater extent by the microbiological revolution, by bacteriology, which came in the last two decades of the previous century. Bacteriology or germ theory introduced what was called The Era of Infectious Disease Epidemiology. With this change, the focus was both expanded and narrowed. Suddenly it was more important to examine identifiable micro-organisms associated with the contagious diseases via laboratory techniques. All diseases had to be examined separately in order to identify the bacteria (and later other micro-organisms as well) that caused the disease, and to investigate the organism’s pathways of infection and distribution. Only in this way would one be able to prevent the transmission of infection. Many argued that specific micro-organisms were not only necessary, but also sufficient causes of disease (25). The German serologist and immunologist Emil Behring (1854-1917) argued that with the discovery of micro-organisms, the study of infectious diseases could be carried out without having to take into consideration or reflect upon social and social-policy matters (30, p. 675). Thus, the new epidemiology was not based on the assumption that morbidity and mortality were related to structural conditions in the society and the material conditions of various groups in the population. Thus, external environmental factors were of little interest. This meant that epidemiology was also depoliticised.

Kristian Feyer Andvord (1855-1934) was a Norwegian doctor who was inspired by these ideas, but who also drew upon ideas from an earlier tradition. For nearly 40 years starting in 1889, he studied the spread of tuberculosis, and he was especially concerned with geographical variations and with monitoring various birth cohorts over a period of time, which he conceptualised as ‘generation studies’. In that way, he introduced and utilised the method that was later given the name of cohort analysis, an analytical method that has played a key role in epidemiology ever since the 1940s. For a long time, Andvord was an unrecognised pioneer in this field (31-33). The aim was to identify the origin and pattern of infection. He wanted to find out how the tubercle bacillus was transmitted from one individual to another and why and how the infection of individuals evolves into a manifest disease. In addition, he distinguished among various tuberculosis conditions and by age, but not by vocation, gender or class. Andvord’s hypothesis was that the infection by tubercle bacilli primarily occurs in early childhood, and that whether someone suffers an outbreak of the disease or develops immunity in adult-
hood will be subject to certain natural laws. In his conclusions, he referred to ‘the laws of nature governing the course of tuberculosis’ and thereby adopted a deterministic and evolutionary approach characterised by the genetic hygienic ideas that were common at that time. He followed up the tradition of looking behind the data established by Eilert Sundt – speculating about geographical variations and referring to processes that were governed by natural law. With broad experience both as a tuberculosis patient and as a tuberculosis doctor, he had good prospects for being able to interpret the data, but he gave minimal emphasis to distinctive social and cultural traits or the patients’ living conditions – a point of view that would seem astounding in today’s world. Otherwise, it is worth noting that Andvord’s statistical work was used as a basis in the formulation of the Tuberculosis Act of 1900 – the first of its kind in the world. This Act was based on a bacteriological interpretation of the disease, where the isolation of the patients and the disinfection of their possessions and of the dwellings and rooms in which sick and contagious persons had spent time, were the primary measures in the fight against tuberculosis in the period around the previous turn of the century (33).

Andvord’s main interest lay in statistics, and he collaborated closely with experts in the field. Internationally this was also a distinguishing trait of the new epidemiology, which began to develop in the period around the previous turn of the century. In the USA and in Europe, doctors worked in partnerships with statisticians, and they developed increasingly sophisticated mathematical-statistical methods. Among other things, biometrics emerged at this time, a mathematical method for studying biological patterns, genetic processes, evolution and variation. All and all and as a result of new developments in medicine, statistics became more and more firmly integrated into medical research (34,5).

After long preparation and further development of the epidemiological methods, The London School of Hygiene and Tropical Medicine was officially opened in 1929 – later described as ‘a Mecca of learning’. It was here that the science of epidemiology consolidated its interest in diseases other than the clearly infectious ones. Those were now well on their way to coming under control, so attention was being focused to a greater extent on occupation diseases and stress-related diseases and gradually also on chronic diseases (5). In the inter-war period – a time of powerful radical social currents – we likewise see that ideas from the first period of epidemiological history about the ways in which social environmental factors such as poverty and destitution play an important role and merge, as it were, with points of view that were developed under bacteriology. An ‘ecological point of view’ would increasingly dominate epidemiology (35, pp. 315-16). Some experts thought that even bacteriology was subordinate: ‘Singly and solely, the social structure is […] in reality crucial to tuberculosis as a disease of the masses […].’ Bacteriologically oriented measures to combat disease are powerless alone against these laws’, wrote an anonymous author in Journal of the Norwegian Medical Association (Tidsskrift for Den norske lægeforening, 36, p. 1081).

Interest in the chronic diseases really came into its own after the Second World War when we entered The Era of Chronic Disease Epidemiology. The great breakthrough came in Great Britain in 1950 with the studies of the statistician Austin Bradford Hill (1897-1991) and the physician Richard Doll (1912-2005) on the relationship between smoking and lung cancer. But that is another story.

The ‘objective’ epidemiology

It may seem striking that epidemiology, with its ‘objective’ statistical methods, has been so strongly coloured by currents of a particular time, by ideology and by politics. Others might say that all science is coloured by time and space, and epidemiology is no exception. It has to be specified now that without a doubt there has been an interaction and interplay between this academic discipline and politics, that epidemiology has shaped public health activities to a great extent, but that the ideas that politicians have demanded or have accepted as relevant knowledge have by no means been arbitrary.

The economist and historian Einar Lie calls attention to the dual role that statistics has always played. It produces facts in a way that is often complicated, but that ideally can be checked. Thus, statistics is difficult and scientific, and numbers often speak louder than words. At the same time, we can interpret and use data in various ways. ‘Few things can be used as flexibly and populistically as statistics. By choosing the right series of numbers, it is easy to give the impression that there are particular cause and effect relationships even when we cannot directly state that such relationships exist’, writes Lie (2, pp. 129-30).

So far in the 21st century, epidemiology – which currently deals to a great extent with Risk Factor Epidemiology – is extremely specialised and requires insight into complicated mathematical methods and statistical analyses. As such, it is as a rule inaccessible to the ordinary citizen, and furthermore, many of us demand broader analyses of epidemiological data related to qualitative assessments. Indeed, Eilert Sundt may very well still have a thing or two to teach us about finding out ‘what is going on behind our numbers’? What’s more, the challenge of the first director of Statistics Norway Anders Nicolai Kiær is certainly still relevant; in 1903 he argued, as already mentioned, that if you are unaware of the conditions that lie behind the data, you end up with tables that are of little worth. ‘For that which lies behind the numbers, the pulsating life and the driving forces, is as a rule unintelligible to anyone who does not know and understand the development at close hand so that they are able to explain it’ (2, p. 69).
REFERENCES


27. Irgens LM. Registerbasert epidemiologisk forskning og forvaltning – et resultat av tilfeldighetenes spill eller en unngåelig historisk utvikling? (Register-based epidemiological research and administration – a result of the workings of chance or an unavoidable historical trend?). *Norsk Epidemiologi* 2007; 17 (2): 199-200.


