James Lind Library

Illustrating the development of fair tests of treatments in health care

---

Tröhler U (2020) Probabilistic thinking and the evaluation of therapies, 1700-1900.

© Ulrich Tröhler, Institute of Social and Preventive Medicine (ISPM), University of Bern, Mittelstrasse 43, CH-3012 Bern, Switzerland, Email: ulrich.troehler@ispm.unibe.ch

TABLE OF CONTENTS

Introduction

The French road to Gavarret's clinical application of probabilistic thinking:

Development of clinical probabilistic practice in 18th and 19th century Britain:

Theory and clinical use of probabilities in Germany after Gavarret:

Conclusions and perspectives for the new century

Acknowledgements

References

Introduction

An overview of the topic
Collecting and comparing data

In this first pandemic year of an infectious disease (Covid-19) it seems particularly apt to recall that the foundations for controlling and eventually eradicating another devastating infectious disease – smallpox –

began during the 18th Century. I can draw on a vast secondary literature to briefly recount this history. It is relevant, for it will deploy this important early example of probabilistic thinking in the history of evaluation of a medical measure, and how this thinking was related to quantification.

Between 1715 and 1721 smallpox had killed one fourteenth of the population of London. Variolation – the inoculation of smallpox lymph into the skin of healthy people as a preventive measure against smallpox (Miller 1957; Rusnock 2002; Huth 2005) – was an oriental and north African practice (Boylston 2012). In Europe it was first used in Britain in the 1720s. Thomas Nettleton (b.1683; Boylston 2010), a physician in Halifax and one of the earliest to carry out mass smallpox inoculation calculated the outcomes in terms of death rates: the death rate of naturally acquired smallpox was "near one fifth" (636 out of 3405) whereas it was none out of 61 inoculated persons (Nettleton 1722). This was an unconsciously expressed probabilistic statement.

James Jurin (b.1684), Secretary of the (London) Royal Society, and a Cambridge MA and MD with a good mathematical education, was motivated by Nettleton's observations to solicit reports of personal and professional experiences with variolation from readers of the *Philosophical Transactions of the Royal Society.* From 1721 he received over sixty replies from physicians and surgeons and summarized them in a series of annual pamphlets (Jurin 1724; Bird 2018). Jurin's analysis concluded that the chance of death from variolation was roughly 1 in 50, while the chance of death from naturally contracted smallpox was 1 in 7 or 8 (Bird 2017; 2018). This was a further example of an unconsciously expressed informal probabilistic statement, implying a mode of probabilistic thinking.

After Jurin's death, the revolutionary technique of systematic collection and computation was continued in London by a Swiss, Johann Caspar Scheuchzer (b.1702). He presented his data in tabular form (Scheuchzer 1729). Similar tabular data were also produced by an American, Zabdiel Boylston (b.1679; Boylston 2008a; Boylston and Williams 2008), who, in his forties, had travelled from Boston to present them to the Royal Society in 1725! Such actuarial data were published, in the *Philosophical Transactions* widely circulated throughout Europe, but they did not end controversies over the propriety and efficacy of smallpox inoculation. Dependence on data collected was doubted: Could one trust in numbers? More data were needed. But there was also opposition of other kinds: Concerns about contagiousness of inoculated persons were raised; and religious fatalists saw inoculation a blasphemous attempt to escape God-sent providence (Rusnock 2002).

Eventually, however, inoculation became widely adopted during the 18th century (Huth 2005). By the end of the century, calculation had been used to evaluate the results of controlled clinical trials (Boylston 2008b), and even mathematics had been deployed to guide contact tracing and prevent spread of the disease (Haygarth 1784; 1793). Vaccination (inoculation with cowpox) had been identified as an even safer way of protecting people from the disease (Boylston 2012). Using these approaches developed in the 18th Century, smallpox was eventually eradicated 200 years later.

These 18th century numerical evaluations of healthcare interventions led to a fundamental debate on the applicability of a formal calculus of probabilities in decisions related to medical treatments.

Applying the calculus of probabilities

Probability had been a branch of mathematics before 1700 (Hacking 1975, 2006). The notions of 'opinion' and 'belief' had been used to express the meaning of certainty for centuries (and sometimes still are). However, these notions of *emotional certainty* of belief could be seen in reality as unconscious probabilistic reasoning. This became clear in the 17th century when mathematization began to deal with games of chance (Daston 1995) and probability became designated 'the doctrine of chance'.

In his book *Ars conjectandi* (The art of conjecturing, published posthumously in 1713), Jacob Bernoulli, professor of mathematics in Basel (b.1654), included works of mathematicians such as Christiaan Huygens, Gerolamo Cardano, Pierre de Fermat, Blaise Pascal and Gottfried Leibniz. As an additional motive for furthering the theory of probability, Bernoulli called for rational action at a time when passion, pride and

prejudice conditioned most political choices. But how could one arrive at a wise decision through a 'democratic process' when there were various loyalties and interests at play? Bernoulli suggested that the way out of this maze was a calculus of probabilities to estimate the errors in human judgment with a high degree of accuracy (Daston 1995). The calculus would be the basis of a science of decision-making (Matthews 2020a).

One of Jacob's nephews, Daniel Bernoulli, yet another member of the famous Basel family of mathematicians, physicists and physicians, attempted this by calculating the advantages provided by the inoculation of smallpox. He sent a *Mémoire* to the *Académie Royale des Sciences* in Paris, and an academic debate ensued.

An academic debate in 18th century Paris

Various historians have written about these deliberations. Their work allows me to summarise the story. Daniel, this younger Bernoulli (b.1700), had extended Jurin's work on "chance" (i.e. probability). Applying a calculus of probabilities to the life tables elaborated by Edmund Halley, his elder British contemporary, he had calculated a life expectancy at birth of 26 years and 7 months (Hald 1998, pp 131-141). This would be increased by three years if a population were inoculated systematically (taking account of the then current estimate of lethality of the procedure of 1 in 200). This result, he wrote, "*appeals to all reasonable (raisonnable) men*". Furthermore, it was in the interest of the State (Marks 2005). It illustrated how the calculus of probabilities was able to provide "certainty" (i.e. high probability) to medical practice by estimating its proximate risk. This practical example of his uncle's programme of applied probability in practice illustrates an early example of consciously used, formal probabilistic reasoning.

This sophisticated paper was read at a meeting of the *Académie* on 13 April 1760. It provoked a violent reaction from Jean Le Rond d'Alembert (b.1717), a younger yet already internationally known French mathematician. He was also the co-editor, with Denis Diderot (b. 1713), of the monumental Enlightenment work, the *Encyclopédie.*

D'Alembert, a longstanding anti-probabilist, reacted to Bernoulli's memoir in a lecture to the *Académie* on 12 November 1760. He pointed out that estimating an additional two years of life, on average, at an undetermined time in the future, would not tempt an individual to risk immediate death from inoculated smallpox. He stressed particularly that neither mothers nor the crowds would accept such a risk, for he considered both as irrational when he said: "*We know how heavily the proximity of feared danger, or of a hoped-for advantage weighs in influencing the crowds*" (Quoted by Rusnock 2002, p 86).

Contrary to Bernoulli's concern with the interests of the state, d'Alembert thus advanced that this did not at all persuade an individual who must risk death (Miller 1957, p 228). Finally, he held that the calculus of probabilities did not permit the assessment of chance (i.e. probability), since there existed no way of estimating future chance (Huber 1959). Indeed, he deemed the calculation of the probability of a probability an impossible task!

Thus, the debate turned about two fundamental kinds of issues, which we shall come across several times in this study:

> i. risk assessment using comparisons of groups; and
> ii. the controversial applicability to individuals of results derived from groups, the 'group-versus-single patient/case problem'.

When Bernoulli's memoir was eventually published by the *Académie* five years later, he defended his arguments by correspondence. He thought that rational actions, as defined by calculation, and actions chosen by individual citizens were synonymous, and that contrary opinions, as held by d'Alembert, were ridiculous and partly attributable to the latter's jealousy because he had not made the discovery himself (de la Harpe and Gabriel 2010).

Nevertheless, d'Alembert's critique drew attention to problems of psychological experience in the interpretation of data which do not seem to have been resolved mathematically even today (Daston 1995, pp 84-91, Marks 2005). By contrast, the data and their applicability were precisely Bernoulli's concern.

This debate was an intellectual highlight, now considered "a classic" in the history of probabilistic thinking (Gigerenzer et al. 1989).

The mathematical path and the clinical path

From the middle of the 18th century onwards, French mathematicians continued their efforts and established a tradition of formal mathematical treatment of probabilities. In 1840, this led Jules Gavarret – a young French physician and mathematician – to apply the calculus of probabilities to clinical practice. Meanwhile some clinicians had independently become involved in probabilistic thinking by informal quantification (Tröhler 2006).

Initially this consisted of nothing more than what had been known since Jurin's times: the systematic collection, counting, and tabulation of observations, and assembling them in groups, ideally for fair comparisons (avoiding bias), calculating averages (means), and then drawing inferences from them. Such calculations – actuarial *medical arithmetic* – implied probabilistic thinking, albeit unconsciously at first. It was also used in Geneva, a Swiss city with particular scientific links to Britain (Tröhler 2000; 2010; Bibliotheca Britannica 1824; Ruffieux 2020).

By the late 18th century a methodological toolbox was thus available for such unconscious probabilistic approaches to the evaluation of clinical practice and therapeutic innovations. And they were used, mainly in British medicine and surgery (Chalmers, Chalmers and Tröhler 2017). It amounted to 'Evidence-Based-Medicine *avant la lettre'.* These approaches were later also used in post-Napoleonic France. As many foreign students went to Paris at that time, they brought these ideas back to their home countries, particularly to Germany and the United States. All this entailed a new type of medical knowledge and was therefore disputable, prompting discussions about the new way of thinking (LaBerge 2005).

Clinics and mathematics merge

After 1840, the work of Jules Gavarret (b.1809; Huth 2006) influenced a group of young German clinicians who promoted discussions of the new methods, using arguments, requests and cautions about formal probabilistic reasoning in clinical medicine. They then started a process of mathematisation, which, by the end of the 19th century, led to the insight that evaluation should become a science in its own right. By contrast, contemporaneous British and French clinical thinking hardly evolved in these ways at that time.

In parallel, medical developments, especially in hygiene and surgery, led to calls for evaluation (Tröhler 2014), and these led to a resumption of discussions about methodological, evidence-based, probabilistic approaches, the *raison d'être* of such an evaluation science. Even so, the purpose of an evaluation science emerged only towards the end of the 20th century in the form of our contemporary, mathematized, probabilistic, Evidence-Based Medicine (EBM).

Lately, debates were resumed about the problems of the EBM approach. For example, modern genetics seem to promise the reality of a so-called 'personalized healthcare', apparently implying less relevance of mathematically sophisticated, probabilistic evaluation. This development reflects the eternal contrast between the empirical and the rationalist approaches for acquiring reliable medical knowledge. EBM is closer to empiricism than rationalism. Will the balance become more equalised (Howick 2016; Matthews 2020a)?

Introduction to the research
Background

The rise of EBM since the 1990s has brought many clinicians, and even medical students, to realize that current best knowledge is based on 'probable estimates' rather than on 'certain wisdom'. This holds particularly for therapeutics, which has witnessed substantial advances:

> These [advances] *require only an attention to probabilities, to leading principles, and to […] a quick discernment where the greatest probability of success lies, and habits of acting in*

consequence of this, with facility and vigour (Gregory 1772,, p 150).

This sentence was not written by a 21st century author. The words are those of a prominent 18th century Scottish professor, John Gregory (b.1724), contained in a work that was translated into French, German, Italian, and Spanish (Gregory 1772). The quotation signals a shift away from the search for an imagined absolute truth, derived from theories and from supportive (single) cases, towards estimates of probability resulting from evidence based on numerous facts, on comparisons, and on calculations. To explore the context of this quotation, and of its evolution over time, was one of the chief motivations of my research.

Scope of my research

Since EBM was thus 're-launched' in the 1990s, a variety of perspectives on it have emerged, including some from basic scientists, clinicians, and historians. For example, Rosser Matthews considered the rise of the RCT in the light of the debates about numerical thinking in the Parisian Academies in the 1830s (Matthews 1995), and Laura Bothwell et al (2016) studied *Lessons from the history of randomized clinical trials (RCT)* after World War II. Other related research has studied the history and sociology of quantification in medicine and health from various standpoints – philosophical, mathematical, epidemiological, clinical, social and political (Gillies 2000; Hacking 1975, Sheynin 1976, 1978, 1982, Stigler 1986; Rusnock 2002; Tröhler 2000, Magnello and Hardy eds. 2002, Jorland et al eds. 2005, Warner 1997, Schlich and Tröhler eds. 2006; Gigerenzer et al. 1989; Gigerenzer 2002, Porter 1986, 1995, 2005). Some of these studies have but marginally touched on the emerging use of probabilities in the clinical context. In my research I have endeavoured to address this gap.

Methods

The study I report here uses an historical perspective. Focussing on probabilistic issues, it is chiefly based on the examination of published original sources; some were also found in the secondary literature; authors were Swiss, French, British, and German, mostly academics of varying social status.

It has two distinct components. First, I consider how the beginnings of methodological approaches to the evaluation of medical interventions were linked to probabilistic underpinnings from the beginning of the 18th century. Second, I compare this evolution in the writings of Swiss, French, British and German authors. In brief, I have attempted a comparative European history of the long dawn of a veritable *evaluation science* in its own right, up to the first tests of statistical significance in late 19th century Germany. My account is thus based on the history of evaluation of medical interventions from simple quantification to mathematization. I find three modes of probabilistic thinking and note that they were used in parallel throughout this history.

- I have analysed the evolving and differing *status quo* in the evaluation of therapies over time, the arguments on which they have been based, and the criteria demanded.
- I have considered some aspects of the *sociology* of the promoters and detractors of these approaches and found them noteworthy; for they partly explain why farsighted mathematical sophistication has remained largely unrecognized by contemporary practitioners – and even by historians. This holds particularly for a group of British and German authors who have not hitherto been recognised.
- I demonstrate *national interdependence* by assessing the extent to which translations and reciprocal quotations have or have not been a feature of the development and application of probabilistic features of the publications examined.
- I speculate about whether there were *national differences* in the generation, reception and dissemination of a mindset across the francophone, anglophone and German speaking worlds.
- Considering the evidence after applying a *long-term, two-century perspective*, I draw attention to the *durability* of these criteria over time.

The French road to Gavarret's clinical application of probabilistic thinking

Introducing French dramatis personae

Let me begin with four outstanding French scientists, born in the 1740s, who knew each other and who all felt that an interest in the evaluation of a therapy would be best satisfied by calculating the probability of its success.

Antoine de Lavoisier (b.1743) was to become one of the most renowned scientists in Europe. Although primarily an experimental chemist (who discovered oxidation), he was also an economist, and an aristocratic high official (a toll collector and therefore beheaded during the French revolution). In the mid-1760s already, precisely when issues of probabilism were debated in the *Académie des Sciences*, he sought for membership of this eminent body. In his recorded works we find, undated but probably around 1784 a note that is worth quoting at length:

> The art of drawing conclusions from experiments and observations consists in evaluating the probabilities, and in judging whether they are large enough, or numerous enough, to amount to proof. This type of calculation is more complicated and more difficult than one thinks; it demands great sagacity and is, in general, beyond the powers of most men. It is upon their errors in this type of calculation that is founded the success of charlatans, sorcerers and alchemists … and, generally, of all those who deceive themselves or attempt to prey on the credulity of the public.

And he continued even more specifically, maybe in remembrance of the Bernoulli – d'Alembert debate of which he must have been aware:

> It is above all in medicine that the difficulty of evaluating the probabilities is greater … Nature, left to its own resources, cures a large number of maladies; when remedies are employed it is infinitely difficult to determine what is due to Nature and what to the remedy. Thus, for all that most people regard the cure of a disease as a proof of the efficacy of the remedy, in the eyes of a wise man this result is only a probability, more or less large, and this probability cannot be converted into certainty except by a large number of results of the same kind (Transl. by IML Donaldson from Lavoisier 1865, p 509. Donaldson 2016a).

And so it truly is. Handling probability is indispensable, but, as we shall see, doctors took a long time to understand this.

Scientists were more progressively minded, at least in their statements, the calculus of probabilities awaked the interest of French mathematicians, one of whom was also a clinician.

Marie-Jean-Antoine Nicolas de Condorcet (b.1743) was an aristocrat who had published several noteworthy mathematical contributions, including on the calculus of probabilities. He became an important official (*Inspecteur Général de la Monnaie*) and Perpetual Secretary of the *Académie des Sciences.*

Jacob Bernoulli's early call for a science of decision-making influenced him and in his *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Essay on the application of probability theory to decision making by majority vote, 1785) he argued how and why probability theory should also serve in political and social life. His 1785 *Essai* was Condorcet's most sophisticated mathematical undertaking. He attempted to set down the rules needed to calculate the veracity of decisions affecting a variety of civic values and matters of criminal justice.

Formerly, probability theory had proven its worth in estimating life annuities or rates of maritime insurance. Now, Condorcet argued, calculation could also serve in an entirely different domain, the operations of the human mind, "*where it weighs the grounds for belief and calculates the probable truth of testimony or decisions*", that is, the consequences of decisions. Events in pre-revolutionary France, prompted by passion and factionalism, proved the need for such a guide. Though never completed, this "social mathematics" remained a part of Condorcet's legacy to mathematicians, social theorists and it influenced some members of the medical profession (Daston 1995, pp 210-224).

Philippe Pinel (b.1745), the reformer of French psychiatry, a fine clinician and chief of psychiatric hospitals in Paris, was a mythical figure in his lifetime for allegedly liberating psychiatric patients from their chains (Weiner

2007). He also held a diploma in mathematics. In his *Traité medico-philosophique de l'aliénation mentale* (1800) (A treatise on insanity, also published in an incomplete English edition in 1806), he wrote:

> To be authentic and conclusive, an experiment must involve a large number of patients submitted to general rules and treated according to a determined order... Finally, favourable as well as unfavourable events have to be reported [from the experiment] so that we can learn from both. That is to say, if one wants to establish treatment methods for disease on solid foundations, they must use the theory of probabilities, which is already happily applied to various fields of civic life (Transl. from Pinel 1809, pp 402-403).

Pinel's reference to the application of probabilities to "various fields of civic life" was an allusion to Bernoulli's *Ars conjectandi* (which he erroneously attributed to Daniel Bernoulli). Finally, he specified that numerical data were needed to compare two competing methods of medical treatment (p 406-407) and concluded: "...it is necessary to apply the elementary notions of the calculus of probabilities" (p 424). As an illustration, he used data obtained at *La Salpétrière* psychiatric hospital in Paris and calculated the proportions of various groups of patients who had recovered. Put simply, he compared these probabilities, albeit not by defining the groups by the treatments they had received (Sheynin 1982, p 250).

Pinel repeated the same thoughts and data in a paper published in 1807:

> Medicine must be based on the theory of probabilities ... on which the methods of treating disease has henceforth to be founded if one wants to establish them on solid grounds.

In the same paper he complained about the habit of suppressing unsuccessful cases, thus *preventing medicine from acquiring the character of a true science, an achievement it will only attain through the application of the calculus of probabilities* (Transl. from quotes in Sheynin 1982, p 250. Dickersin and Chalmers 2010; Bird 2018; Bishop and Gill 2019).

Here we have further examples of conscious, yet pre-mathematical probabilistic reasoning. Others were soon to follow. Despite their reiteration, however, Pinel's references to probability were mere words, a loose allusion, implying no more than calculating proportions. His work contains no example of the application of the calculus of probabilities. Finally, as shown earlier in this manuscript, his method did not differ from that of Jurin (described above). Practical limitations, such as wide dissimilarity of his case histories, restricted the application of quantitative evaluations.

However, the contributions of French mathematicians, such as those of Laplace and Poisson, continued to assert the potential usefulness of probabilistic approaches to clinical medicine.

Pierre Simon de Laplace (b.1749) was the foremost French mathematician, physicist and astronomer of his day and remains one of the great scientists of all times; his pupil Poisson called him "the French Newton" (Stigler 1986, p 31). Laplace was one of the founding fathers of probability theory, and the Bayesian interpretation of probability was developed mainly by him (Stigler 1986, chapter 3). The mathematician-historian Robert Matthews will discuss Bayes's theorem, a particular approach of statistical inference, in a forthcoming paper.

In some lectures given in 1795, Laplace reaffirmed Condorcet's optimism concerning social mathematics: "*Let us apply to the political and moral sciences the method founded upon observation and calculation, which has served us so well in the natural sciences.*" Yet, echoing Jacob Bernoulli, he was sceptical about their usefulness because of passions and self-interests involved in decision-making in these fields. As the medical historian Terence Murphy has noted, Laplace sensed that *"...the more vital the issue, the more likely are vested interests to counter the voice of reason",* so it was useless to use the calculus of probabilities to determine the truth of such decisions. Yet he hoped for its valuable use in the future (Murphy 1981, p 305). He expressed this in two, albeit short, passages within two works.

Laplace explicitly mentioned therapeutics in his *Théorie analytique des probabilités* (analytical theory of probabilities 1812):

> *The calculus of probabilities can make us appreciate the advantages and inconveniencies of the methods employed in the conjectural sciences. Thus, in order to recognize the best treatment in healing an illness it suffices to try each of them on the same number of patients while keeping all the circumstances perfectly alike. The superiority of the most advantageous treatment will be manifested more and more as the number of cases increases; and the calculus [of probabilities] will make known the probability corresponding to its [the treatment's] advantage* (Transl. from de Laplace 1820, p LXXVII).

Thus, the calculus would make known probability corresponding to a treatment's] advantage, as long as there were sufficient cases and an unvarying relationship between treatment and outcome exists – which did not reflect doctors' experiences/ expectations.

His *Essai philosophique sur les probabilités* (Philosophical essay on probabilities 1814) was the Introduction to the *Théorie analytique,* added to its later editions, but it was also published separately. As such it contains no mathematical formulae; it typically opens with a general statement:

> *One may even say, strictly speaking, that almost all our knowledge is only probable; and in the small number of things that we are able to know with certainty, is the mathematical sciences themselves. The principal means of arriving at the truth – induction and analogy – are based on probabilities* (de Laplace 1995, p 1).

Although he further wrote that "*the theory of probabilities is fundamentally only good sense reduced to calculation*", Laplace wanted to submit intuitive judgments to the rigors of analysis.

His clear, consciously formulated plea for the use of formal probability in therapeutic evaluation underestimated the difficulty resulting from his requirement for "a sufficient number" (Sheynin 1978, p 285), after realizing that, for the time being, the number was beyond a doctor's competence. Furthermore, the calculations assumed a constant relation between causes and observed putative effects. This inherent constant relation would have troubled physicians. In fact, whether numerical methods could actually influence the choice of a remedy was not raised, and Laplace himself never applied the calculus of probabilities to medical phenomena (Murphy 1981).

Laplace's *Théorie* was extremely influential. Not only did it have six contemporary editions, but it was translated into English (1820), German (1820), and Dutch (18…?); and it is still in print in many languages. Yet, mathematically speaking, theory of probability stagnated because new fields of application (physics, biology) had not yet appeared. Furthermore, his book was difficult to read. Almost the only mathematician elaborating on Laplace's work during this period was his pupil, Siméon Denis Poisson, 32 years his junior (Sheynin 1976, pp 179-180). (I shall discuss him below and here).

In fact, at#contents15 the end of the 18th and beginning of the 19th centuries, contemporaries agreed that observation and "experience" were the basis of sound therapeutics (Murphy 1981, p 309). But experience could have many meanings. In French medical journals and *Dictionaires* it might include, during these decades, subjective opinions, beliefs, based on single case descriptions or on extended follow-up, as well as hospital data. Furthermore, statistical work on social groups (which would become public health) was flourishing in France in the 1820s. Not until the mid-1830s, however, were methodological issues about quantitative comparisons in the evaluation of therapies debated – in Paris, of course.

Unconscious probabilists

I came across two remarkable French authors of very diverse social standing who remain unnoticed and unstudied in relation to their probabilistic thinking in medicine. Neither of them appeared to have been aware that he was an unconscious probabilist.

In the early 18th century the question whether an absolutely necessary amputation had to be performed as soon as possible after an injury or after a few days' delay had been preoccupying surgeons for some time.

Theoretically you could argue for or against both methods. An ordinary mid-18th century army surgeon, decided to solve the question using a trial. An army surgeon named JF Faure, whose Christian names and date of birth I was unable to find, described the experiment in detail:

> [This experiment was done] *in the hope that we would have a less equivocal success and also in order to affirm the principles by tests repeated sufficiently to overcome the disbelief of the most prejudiced. Ten English wounded, out of a number of about one thousand who had been taken to the hospitals of Douay after the battle of Fontenoy* [1745]*, were therefore set aside. Their wounds were such that amputation was essential in most of them... It was simply a question of whether the amputation was carried out sooner or later* (Transl. from Faure 1759, p 353).

All ten survived. Faure compared this outcome to the overall mortality of those who had received immediate amputation. Adducing further testimonies and numbers he calculated the chance of healing: It amounted to nine out of ten after delayed intervention, and to between one in ten and one in three after immediate amputation. He concluded that "*it seemed to me difficult not to be impressed by an experiment repeated ten times and always with the same success*" and he sent a "*Mémoire*" to the French *Académie Royale de Chirurgie* (Faure 1759).

The Academy's decision of to award its 1754 annual prize to Faure made delayed amputation respectable. But the issue continued to be debated on the same basis of retro-and prospective trials for another hundred years, particularly from German and British military statistics (Tröhler 2000).

My second example of an unconscious probabilist is the distinguished surgeon Baron Anthelme-Balthasar Richerand (b.1779). He was of provincial stock, but he became a *protégé* of Cabanis after moving to Paris. He rose to influence as author of a textbook of physiology (ten editions), then as chief surgeon of the *Hôpital St.Louis,* as a co-founder of the *Académie de Médecine* (1820)*,* and finally as a writer of popular texts on medicine.

Richerand proposed a solution for a seven-decade dispute about the treatment of cataract. Couching (i.e. displacement) of the opaque lens had been the standard therapy since Antiquity, and a problem arose when Jacques Daviel (b.1683 or 1693?) published his method for extracting the lens in 1753). Supporters of each method had fought on the basis of case series.

In his *Des progrès récens* [sic!] *de la chirurgie* (Recent advances in surgery, 1825) Richerand dealt with the uncertainty hovering above various approaches to treating illnesses. He observed that treatments for cataract in particular "*still divide the supporters of* extraction *and* couching *of the lens*" [author's italics]. But there was "*only one way open to provide an escape from this maze of contradictory opinions and to resolve this important point in surgical doctrine*". And this was, in today's terms, a prospective trial, comparing simultaneously "*a certain number of patients*" placed in the same circumstances, then operated on comparatively under the eyes of the Academy.

> [For] *an academic body alone, the sole interest of which is that of truth, is able to undertake and follow up such an experiment. Even the most able surgeon, and who in exercising his art aims at the truth with the greatest honesty and good faith, would be unable to defend himself against a multitude of prejudices, the existence and power of which he often ignores. Thence, what credibility can one attribute to those men of bad faith, for whom truth is nothing other than fashion acquired by misrepresentation? And what have we to understand by what they name their 'successes' by the use of this or that method?* (Transl. from Richerand 1825, p 27).

Richerand did not mention a single word about the probability involved as a representation of the much-maligned uncertainty. We would say the above was the outline of a protocol for a prospective controlled clinical trial, albeit, without randomization, specification of outcome or method of analysis. No trial appears to have taken place, but ten years later, some of Richerand's colleagues from the Royal Academy discussed it – without quoting him – and concluded that they had settled the debate.

Defence of the status quo

Pierre-Jean-Georges Cabanis (b.1757), a Paris physician and hospital administrator, was also a theoretician at the height of the currents of his time. He formulated earlier reactions to quantitative techniques in his *Du degré de certitude de la médecine* (On the degree of certainty in medicine, 1798), written in 1788. Cabanis referred to d'Alembert and Condorcet (a friend) when he wrote "*Each science has its own kind of proofs*": A "happy instinct", i.e. a non-quantifiable talent, a kind of "*sympathie morale*", allowed a doctor to choose. In effect, Cabanis was defending the anarchic *status quo* of clinical practice. He remained attached to the idea that medical practice had a specific nature. This ruled out any formal alliance with the natural sciences and justified claims for criteria of medical knowledge independent of any formal standards established in other fields of human enquiry (Murphy 1981, Canguilhem 1970). His book, translated into German (1799) and English, was widely distributed, including in America. It was an early expression of the later widely held feeling among doctors that there should be no 'strangers' (that is, mathematicians) at the bedside, encroaching on the 'sacred' field of medicine. This feeling found expression in two academic debates two decades later.

Debates about the applicability of numbers to clinical questions in mid-1830 Paris

The *Académie des Sciences*

In 1835, a statistical account of two treatments of bladder stone had been submitted for consideration by the *Académie des Sciences.* It compared the traditional extraction of the stone after cutting into the bladder to the innovative crushing of the stone by lithotripsy. A Commission of the Academy was charged with reporting to its members. It consisted of two elderly gentlemen, Napoleon's legendary surgeon Dominique-Jean Larrey (b.1766) and the physician François Double (b.1776), and two comparatively younger members, Siméon Denis Poisson (*1781), the mathematician, and the chemist Pierre-Louis Dulong (*1785), now the Academy's secretary.

The *Rapport* appeared on 5 October 1835 (Poisson, Dulong, Larrey and Double 1835). It disapproved of any application of calculations to medical problems. *Rapporteur* François Double's principal objection to numerical analysis was based on the suppression of individual differences required by the method: "*In statistical matters… the first care before all else is to ignore that a man is an isolated individual and only to consider him as a fraction of the species*". A second point was the practical unfeasibility: The need for a large number of facts [still, how many?] could never be met. The method was "*inappropriate to elevate the human spirit to that mathematical certainty found only in astronomy*" (Matthews 2020a).This *Rapport* was considered relevant enough to be reprinted and commented on in 2001 by the International Journal of Epidemiology (Poisson, Dulong, Larrey and Double 1835, repr.2001; Tröhler 2001).

Poisson – by now a renowned mathematician – was not convinced by these objections nor by the allegedly insuperable difficulties of mathematizing medicine. He was a probabilistic thinker. It was indeed only during these years that this versatile mathematician was dealing with the calculus of probabilities. In consequence, he was to publish two years later his respective contribution highlighting again the Law of Large Numbers, already described by Jacob Bernoulli and Laplace. (I will further point to it below on p 25). In essence it said that more data reduces uncertainty (Matthews 2020a).

The *Rapport* on bladder stone was the starting point for another long discussion, in 1837, in the *Académie Royale de Médecine*. Its context and history have been analysed (Murphy 1981; Matthews 1995; La Berge 2005). In what follows, I focus on the probability aspects.

The *Académie Royale de Médecine*

This time three advocates of the numerical approach, physicians of Poisson's generation and younger, were part of the Commission. They were Pierre-Charles-Alexandre Louis (b.1787), Auguste François Chomel (b.1788), and Jean Baptiste Bouillaud (*1796). This Commission reported on a study purporting to demonstrate the superiority of repeated purges over bleeding in the treatment of typhoid fever. The report cautioned against any premature application of numbers. This would distort results, yet properly applied, numbers could be decisive. Some members of the Academy did not share this Commission's enthusiasm and asked for an enquiry of the utility of statistics applied to medicine. Thereupon a debate started a month later, in April 1837.

The main speaker in favour of statistics was Louis, a solitary Paris hospital pathologist. He had authored papers of numerical anatomo-clinical descriptions of diseases (nosographies). The issue now was a study on various treatments of pneumonia previously submitted to the Academy. This had resulted in showing – was that not impossible? – the limited value of the then much favoured method of bloodletting. In this, Louis had used what came to be called "the numerical method", although he never introduced a formal definition of it (Sheynin 1982, p 250). In fact, it was nothing more than a statement of proportions of successes – or failures – out of total number of patients treated. He saw in this the only way to raise the epistemic state of medicine onto a par with that of other sciences. Unconscious, informal probabilistic thinking was behind it.

Louis's opponent, the younger clinician Benigno Risueño d'Amador (b.1802), represented the individualistic neo-hippocratic school of Montpellier. He had travelled to Paris especially for the occasion, where he droned on over seven consecutive sessions (up to July 1837) about the time-honoured *art médical*. D'Amador emphasised that it was being proposed that this art would be replaced by the counting method, "a uniform, blind and mechanical routine" yielding only probable rather than certain results. Was medicine to become a gambling place, a lottery? If, as a consequence, one followed treatment for the majority of patients, what would happen to the minority? "*What we need is certainty*", he insisted. Furthermore, biologic variability over time could not be fixed by a number (La Berge 2005). On the positive side, d'Amador pleaded for the use of induction based on similarity among cases, whereas, according to him, Louis's numerical method was based on haphazardly assembled groups.

Further criticism included again the impossibility of finding sufficient numbers of comparable cases and the fact that there was no reason to abandon one time-honoured treatment – bleeding – for another – purging – since patients had also died with the latter. And finally, Risueño d'Amador formulated the essence of his conviction:

> *Never, and in no instance, can a doctor judge the utility of his art by the results of large numbers. Nature preserves the species; art prolongs an individual's life as long as it can.*

Maybe he formulated this *as a way of articulating his fear of modernity*, when medicine would "no longer be an art, but a lottery", a warning to an utilitaristic science, a wrong science, as d'Amador argued (Quoted from La Berge 2005, pp 93-96).

Results of the Paris debates

With hindsight one sees that the main question was not "Which remedy do I prefer"? but rather "How is competent medical judgment to be achieved"? The central point was thus rather about the concept of medicine and its epistemic status compared to that of other sciences. There were also social issues at stake: mathematicians (and other scientists) were not considered to have sufficient knowledge to evaluate medical practice. These "strangers at the bedside" would destroy the unique nature of doctors' personal intervention in the life of an ailing patient: the doctor's prestige might be hampered. Quantification and probability were double-edged. As medical historian Andrea Rusnock summarised: "*Assigning numbers to people [...] de-individualized and de-humanized, and at the same time, it leveled an unequal and hierarchical society*" (Rusnock 2002, p 217). Finally, d'Alembert's old problem was a major point again: should, or how could one use results obtained from a group to gain a prediction for an individual? It was the 'group-versus-single patient/case problem' – an apparently irresolvable problem?

The advocates and opponents continued quarrelling during succeeding decades. Yet confusion and vendetta were resolved with respect to three essential points:

Firstly, henceforth the distinction was made between

i. medical statistics, seen as indispensable for the emerging field of hygiene (that is studies of populations, public health, which adopted them in its vocabulary); and
ii. the numerical method (*méthode numérique*) of clinical medicine, seen as dry calculations, often not based on discrete, homogenous data (an obvious fallacy of Louis's analyses).

Secondly, an old question had become apparent: How valuable was any of these methods? As one participant suggested:

> *There is one indispensable condition for the validity of statistical results, and that is the morality of the observer, his good faith, his intelligence. Good faith is necessary, for facts have been invented or falsified in the past* (quoted from Murphy 1981, p 315).

How true this sentence still is!

Finally, and for us: Louis's *méthode numérique* still corresponded to Jurin's mode of unconscious, informal probabilistic thinking, whereas Risueño d'Amador in challenging it, did sometimes so in consciously probabilistic terms. All three were pre-mathematical. This was not the case of Siméon Denis Poisson, the mathematician.

Jules Gavarret introduces the 'calculus of probabilities' to clinical medicine around 1840

Poisson was Laplace's heir in probability theory and therefore an interested attendant of the Paris debates., Concurrently he was working at extending Condorcet's and Laplace's work on probability theory. He had become a clear supporter of the calculus of probabilities, mainly for evaluating therapies – a formal probabilist. He developed relevant equations – specifically his Law of Large Numbers – and, after the summer break of 1837, he published his *Recherches sur les probabilités des jugements en matière civile* (Enquiries into the probabilities of judgments in civic matters, 1837). This book was to open the way to a new view of the calculus of probabilities which could be applied to clinical medicine. It was a notable step forward to applying scientific principles to the evaluation of therapies compared to the *méthode numérique* (Sheynin 1978).

Another follower of the 1835 discussion at the *Académie des Sciences* and of Poisson's work was **Louis-Denis-Jules Gavarret** (b.1809), who had studied under Poisson and graduated from the *Ecole Polytechnique* before turning to medicine (Sheynin 1982, [Huth 2006](#)). Born in the early 19th century he was also a first physician to be added as the next generation's link to the Bernoulli-Condorcet-Laplace-Poisson-chain. And, as a young clinician, he advanced the application of their intellectual heritage in practice. First, he distinguished descriptive and inferential statistics. Second, he published a first French textbook on the field of statistical inference, the *Principes de Statistique Médicale: ou dévelopement des règles qui doivent présider à son employ* (1840) ([Gavarret 1840](#)) in which he gave a concrete example of the application of the rules to clinical medicine.

Using Louis's data and Poisson's mathematics, he calculated the possible errors of averages or means. He called them "limites d' oscillation" (limits of possible errors; these are, however, not equivalent to today's confidence intervals, for the basis of calculation is different). He saw that

> *to be able to decide in favour of one treatment over another, it is not sufficient that the method yields better results, but that the difference found must also exceed a certain limit, the value of which is a function of the number of observations.* [Contrarywise he concluded, very harshly indeed, that] *Each difference between two results obtained which falls within this limit, which is the smaller, the greater the number of observations, may be disregarded and considered as null* (Transl. from Gavarret 1844, p 158).

Gavarret further set stringent requirements of basic comparability between groups when designing a trial that would yield reliable results. From mathematical assumptions he calculated that such a trial would need 300 (or at least some 200) cases per group. Then, if the resulting probability that a difference was not due to chance should amount to 99.5%. (In the 'odds-representation' of probability this would read: odds are 212 in 213, that amounts to a probability P=0.9953). This choice of probability went back to Poisson; it was based on mathematical convenience. It was a compromise to establish probability as near to certainty as possible: 212/213 is a fraction near to unity that cannot be reduced as both nominator and denominator are near to prime numbers. If one chose less cases the resulting probability would be lower, and that was not reliable in his view.

Gavarret's effort resulted in a new definition of probability, at least for medicine: determining the limits of error of two averages. Finally, someone had provided an answer to the long-standing question of "*how many* cases were needed". But of course, there were practical difficulties.

The responses to Gavarret's book varied widely. It was translated into German (1844). However, except in Germany, it elicited little attention among physicians and was no longer cited at the end of the 19th century (Matthews 1995).

This meant that discussions about the value of numbers, and different understandings about the notion of statistics, went on as trials continued. Some of these were well-designed, but many had obvious shortcomings – noticed by contemporaries. Gavarret's conscious, formally mathematical mode of probabilistic thinking surely was not mainstream.

Therapeutic reasoning in France after Gavarret

The multidisciplinary nature of the *Académie Royale de Médecine* should have provided a natural setting for deepening methodological insights in therapeutic reasoning, yet no methodological sophistication was elaborated. Numbers were used in medical debates about therapeutic innovations, but Louis's *méthode numérique* was obviously considered sufficient. These debates were characterized by disagreements and fights for exclusivity. A comparative sloppiness in observation was criticized as the weakness of this apparently precise method. Such statistics did not weigh greatly in the minds of the academicians when compared to results from morbid anatomy, experimental physiology and animal experiments, all of which were used as sources of evidence. The issue was often about the multiplicity of therapies rather than about which one was best. Furthermore, clinical statistics were not as easily collected as animal experiments, which could be repeated, or anatomical specimens, which could be demonstrated in the conference hall of the Academy. The emphasis was thus on therapeutic rationale rather than on therapeutic effects (Weisz 1993).

By mid-19th century, hospital statistics were being widely collected. In Paris, there was a statistical commission to co-ordinate the data procured. Medical historian George Weisz noted: "*…[c]ounting was not merely occurring… but was providing convincing data in a limited number of cases*". Even if simple counting was frequently insufficient to provide convincing evidence, this did not mean that it was opposed in principle: its mathematical limitations were not widely understood, so it had become incorporated as an element of individual clinical judgement rather than being an alternative to it (Weisz 1993, p 302).

This changed after the mid-1850s when, after the introduction of anaesthesia, surgical innovations became more and more frequent. One example was tracheotomy (Opinel and Gachelin 2010), on which two debates were held in the *Académie* in 1839 and 1859, respectively, the latter comparing this surgical intervention to intubation, i.e. the insertion of a metal tube into the trachea (Opinel and Gachelin 2010). In 1839, unilateral clinical statistics from various published sources were added up so that 18 "cures" of 60 operations could be claimed. Twenty years later, 27% of 446 tracheotomies performed during the previous nine years at the *Hôpital des Enfants Malades* had been successful. On the other hand, ulcerated dog larynxes after prolonged intubation were demonstrated to justify tracheotomy rather than intubation despite its low success rate. At least some children had been "saved" by surgery (Weisz, 1995, pp 169-172).

Major surgical procedures such as amputation and lithotomy had in fact been subjected to counting since the 18th century (Tröhler 2000; Sheynin 1982), because diagnosis, prognosis and therapeutic results seemed clear cut, namely survival or death. Now removal of tumours, treatment of ovarian cysts or infectious foci continued to be reported in this simple statistical way – and compared with the presumably fatal issue of conservative therapies. There were no recognized baselines permitting comparison with other treatments or with no intervention. Some dissatisfaction about such uncontrolled statistics of operative results was voiced by French surgeons in 1873 and still three decades later in 1908 (Weisz 1995, p 173; Verchère 1908).

In the case of systemic therapies, however, few things remained stable during the 19th century, neither the diagnostic category, nor often the constantly evolving procedures, nor the results. George Weisz concluded from

his historical study of the Paris *Académie de Médicine* that doctors attempting therapeutic evaluations

> *tended to rely on case descriptions. Frequently, they avoided the issue of evaluation altogether to concentrate on whether a particular therapy "made sense". And in doing so they sometimes made use of scientific techniques that were far more sophisticated than counting* (Weisz 1993, p 303).

Putting the numerical method into context

Medical science versus medical art

In 1828, the then young physician **Armand Trousseau** (b.1801) accompanied Louis to Gibraltar on a government commission to study an outbreak of yellow fever. Louis's manuscript was published a decade later in English. The translator, Dr. Cowan (whom we will meet again below), shows in his introduction the outdated conception of Louis's probabilistic thinking:

> *In the present state of science, we must often be content with probability. M. Louis acknowledges this, whilst he insists that there is a great difference between the probable and the true, for the probable may be false* (Louis 1839, p XV).

So, an everlasting "truth" was apparently in his mind and directing his research – and also Trousseau's, for he surely knew Louis's great work on the pathological anatomy of phthisis (tuberculosis) published in 1825, shortly before their common trip. Trousseau knew about the *méthode numérique* and he was certainly aware of Louis's later works, particularly on the lack of demonstrated beneficial effects of bloodletting (1835) (Hannaway 2007).

Thirty years later, by now a prominent Parisian clinician of the day, and a brilliant orator, Trousseau published his *Clinique médicale de l'Hôtel-Dieu* (Clinical lectures delivered at the Hôtel-Dieu, 1861). In the *Introduction* he devoted eight pages to the numerical method.

For him, this was no more than the replacement of expressions such as "sometimes", "frequently", and "often" by exact proportions. This might sometimes be useful, but only secondarily so, for example, when it would lead to new notions in the future. In that way, Trousseau recommended the method and admitted that he had used it himself.

As to the claimed veracity of the results, he asked, rhetorically:

> *Don't you think [...], messieurs, that if one wants to lie, one cannot do this as well with exact numbers as with approximates* [and without] *fabricating details much labor* [sic!] *and with less hypocrisy?* (transl. from Trousseau 1865, p XLIII).

Although Trousseau had never calculated any proportions, let alone any probabilities with all their claimed rigor, he stressed their limitations, saying that they could yield only

> *raw, unelaborated, elementary results [...that] are simply a pasture for the medical intelligence needed to elaborate them. I reproach [the method] to count only, [...] to stick to the rigorous result like a mathematician....*

And he continued the polemic in the style he had known from the *Academy* debate thirty years previously, sometimes quoting Risueño d'Amador:

> *This [numerical] method is the scourge of intelligence. It transforms the physician into a clerk, a passive servant of numbers which he has superposed; and the maximal reproach I raise against it is to suffocate medical intelligence* (Transl. from Trousseau 1865, pp XL-XLII).

Observing facts, systematizing them by counting, submitting an equal number of cases to two modes of treatment to decide a therapeutic question – these were the characteristics of medical science for Trousseau. But medical science was not to be confounded with medical art. He stressed

> *Medicine is more of an art, and the doctor truly worthy of his ministry must above all glorify*
> *himself not to be only a learned scientist. And even when the doctor, unfortunately, errs often,*
> *one nevertheless finds more charm, more attraction in the study of an art, and [... medicine] needs*
> *a bit more intervention of intelligence [understanding and knowing] than the sciences where we*
> *are directed by certain and invariable rules* (Transl. from Trousseau 1865, pp 4-5).

In conclusion, Trousseau thought statistics are

> *made so much noise of for such poor results that one ought not to support it to deceive young*
> *people by a kind of charlatanism of exactness and truth* (Transl. from Trousseau 1865, p XLIV).

Trousseau, too, was well aware of therapeutic trials and had done some himself, for example, to evaluate homeopathy in 1834 (Dean 2009). To be sure, his two homeopathy trials for various illnesses were single-blinded, yet he did not report whether symptomatic improvements, if any, were more than transient. Nevertheless, he believed that these were valid tests. Later he erred again in design and inference in his own sphere of orthodox medicine, dismissing one of its most effective specifics of the day, colchicine for gout, as a placebo in the same way he had done for homeopathy (Dean 2004, pp 142-144): Trousseau's understanding of the numerical method was absurd. Bearing this in mind we doubt his critical acumen when we read his review of Louis's book on venesection (1835):

> *... I confess that I have been one of the most violent, one of the most unjust detractors of this*
> *[numerical] method, I did not understand it; today, having studied it, I admit that it alone will*
> *enable science to make solid progresses, that it alone will allow in future centuries the use of the*
> *works of those who shall have lived before, and to raise slowly an edifice that the dreams of a*
> *Galen or of a Paracelsus will impossibly be able to throw down"* (Transl. from a quote in Bariéty 1972, p 182).

Clearly, Trousseau and therapeutic reasoning remained unconscious, pre-mathematical, complex, messy...and verbose.

Probability versus certainty

Trousseau's was one line of argument, medical art against medical science. Another line was probability against certainty. In France, this was articulated particularly forcefully by Claude Bernard**.** In 1865, when the second edition of Trousseau's *Leçons Cliniques* came out, his contemporary, Claude Bernard (b.1813), published his *Introduction à l'étude de la médecine expérimentale* (An introduction to the study of experimental medicine, 1865). At this time Bernard was already a member of the *Académie Française,* a world-famous physiologist. One of the principles underlying his work was that, in nature, every effect was due to a precise cause. This constant relationship of determinism could be discovered through animal experimentation. This was a typical line of arguments of physiologists, which we shall come across also in Germany (see below). Medical science was to look for certainty whereas statistics could only offer probabilities and was therefore inappropriate for physiology.

Less polemical than Trousseau, he saw empirical clinical medicine, based, as it was, on comparative experiments and statistics (in the sense of counting), as being in an intermediary stage between old "tact and intuition" and (future) "scientific medicine". This would be rooted in (animal and other preclinical) experimentation (Morabia 2007). In therapeutics – for the time being – one could not do without the probability of statistics; given constant progress, this was an unavoidable concession to pragmatism. Had it not been shown recently, by "*comparative experiment [...] "that treatment of pneumonia by bleeding, which was believed most efficacious, is a mere therapeutic illusion*" (Transl. from Bernard 1865, p 273). Of course, this was a hint to the shortcomings of Louis's research design and inferences.

Outlook

Despite his important contribution based on formal, mathematical probabilistic thinking, Gavarret seems to have had no followers in 19th century French clinical thinking. If seen at all as useful in clinical evaluation, Louis's *méthode numérique,* sensibly used, was the way to use numbers. This implied unconscious probabilistic thinking. There was much confusion about the notions of statistics, experiment and experience on one level. On other levels, there stood issues of medical science versus medical art and of probability versus deterministic certainty. Explicit probabilistic thinking was hardly considered.

Trousseau's lengthy *Introduction* containing his epistemological considerations was not included in the rapidly published translations of the *Leçons cliniques* into English (1st edition 1868), Spanish and German. There had been neither a contemporary English edition of Gavarret's book, nor was Bernard's book published in English until 1927. However, Gavarret was extensively reviewed in *The British and Foreign Medical Review,* probably by its editor, John Forbes, a very astute, critical thinker (Agnew 2008). And for various reasons, partly historical, partly out of intellectual curiosity, Gavarret became influential mainly in Germany (see below), the USA (Warner 2003, Bartlett 1844) and, to some extent also in Britain.

Building a clinical probabilistic tradition in Britain before Gavarret

The long 18th century
Theory

As noted in the introductory sketch of this study, there was some (unconscious) probabilistic thinking in British clinics during the 18th century. In this section, I will consider in more detail the motives for, and the modes of, this reasoning, and look further into the 19th century.

It is easily forgotten that the eighteenth century was a time of innovations in medical and surgical treatments (Tröhler 2000; 2003a; 2006; 2010; 2013). How were they presented? An important precondition for probabilistic thinking gaining ground in medicine, particularly in therapeutics, was – and still is – to step away from confidence in the absolute authority of doctors, whose opinions were too often based on selected, 'successful' (single) cases.

Two traditions were combining to build an indispensable basis for probabilistic reasoning during 18th century Britain, and these had their origins in the 17th century with Bacon, Sydenham and Locke (Dickersin and Chalmers 2010). This process involved drawing inferences, and even axioms, from carefully registered and sometimes comparative observations; and by using numbers to assign symptoms in order to differentiate disease categories and to evaluate interventions (Tröhler 2000; 2005; 2010). And there was a growing tendency to report all cases of new treatments observed during a given time period – whether successful or failures, a novelty in itself! Indeed, it has been suggested that this feature of James Lind's reporting was of more fundamental importance than his controlled trial of treatments for scurvy (Justman 2017).

Three modes of probabilistic reasoning by 18th century clinicians

In other words, there was a transition from (seemingly) certain knowledge to reliance on relative results based on many observations, successful or otherwise, results that were recognized as partial and evanescent as time went on.

The naval physician James Lind (b.1716) recognized this in 1772 after three decades of service:

> *A work more perfect and remedies more absolutely certain might perhaps have been expected from an inspection of several thousand … patients.* [Certainty was deceitful, he concluded, *for though they may for a little, flatter with hopes of greater success, yet more enlarged experience must ever evince the fallacy of positive assertions in the healing art* (Lind 1772, pp v-vi).

This and other statements by Lind (Tröhler 2003b; 2003c) illustrates the unconscious mode of probabilistic reasoning. I found a formulation of a conscious, pre-mathematical mode in the same year by the contemporary British physician, John Gregory, and an application of the conscious, mathematical mode by a clinician of the following generation, John Haygarth.

At the outset of this report I quoted **John Gregory** (b.1724), a celebrated professor of medicine in Edinburgh, for his explicit use of the term 'probability' in 1772 (see above). He had also had a mathematical education (and had taught mathematics): In his view, rather than getting stuck in endless argumentations any

> ...advancement of the sciences and the successful management of business in private life... require[d] only an attention to probabilities, to leading principles, and to [...] a quick discernment where the greatest probability of success lies, and habits of acting in consequence of this, with facility and vigour (Gregory 1772/1805, p 150).

He repeated this argument three more times in his book (pp 15, 132, 193). And describing the psychological hindrances to doing so he concluded:

> It is, indeed, difficult and painful for men to give up favourite opinions, and to sink from a state of security and confidence into one of suspence [sic!] and scepticism... Accordingly, we find that physicians do not easily change the principles they first set out with (Gregory 1772/1805, p 186).

These insights were published in Gregory's *Lectures on the Duties and Qualifications of a Physician* (Gregory 1772), still often quoted today as the first textbook of medical ethics written in English. They were re-edited in 1805 and in 1817 (in Philadelphia) and translated into French, German, Italian and Spanish. From the 19th century, this book became influential in these cultures for its concept of 'the sympathetic physician' (Baker and Mc Cullough 2009). However, with hindsight, it seems that the passages on probability, linked to the need to acknowledge one's limitations, encountered difficulties in being acted upon with "facility and vigour".

As mentioned above, new data were needed to show the effectiveness of variolation, this all-embracing innovation in 18th century medical practice In turn, this novel concept generated a field of application for the equally novel concept of probability (de la Condamine 1754; Franklin 1759; Watson 1768).

In 1784, **John Haygarth** (b.1740, Booth 2000) of Chester, yet another physician with a mathematical education, wrote in his *Inquiry how to prevent the smallpox:*

> It occurred to me that it might be computed arithmetically by the doctrine of chances, according to the data, if *one, if two, or if three persons were exposed, for the first time, to the variolous infection, what degree of probability there was that one or more of them would catch the distemper. At my request a mathematical friend made the following computation, on each suppos[it]ion* (Haygarth 1784, pp 25-26).

After two lengthy sets of suppositions and calculations, added in small print, Haygarth concluded that "*when three or more persons together, at the same place, at the same time, have all escaped the small pox, [...] they were not exposed to the variolous infection,"* and he confirmed this in an enquiry with 31 doctors. His "mathematical friend", Mr. Dawson, "*a truly mathematical genious"*, had indeed applied a calculus of probabilities (Haygarth 1784, pp 25-26).

Quantification in clinical experience

James Lind, John Gregory and John Haygarth were significant representatives of the 18th century British movement of 'arithmetic observation': Quantified empirical observations were used to challenge therapeutic dogma (Donaldson 2016b) and to monitor the introduction of new therapies, both based on what were believed to be 'rational' theories (Tröhler 2000, 2010). These efforts carried with them the notion of probability of success of a therapy rather than certainty – and a host of new problems. For instance, could averages derived from documenting outcomes in groups be applied to an individual? (This question lacked a satisfactory response

(Tröhler 2005). And how could comparable data be assembled? As implicit in the variolation story, record keeping was the answer and it was well underway in the 1720s (Tröhler 2013).

Indeed, exact day-to-day record keeping in tabular form was repeatedly propagated and practised in Britain by many doctors throughout the 18th century. The resulting returns, from hospital and military registers and public dispensaries were discovered as a new source for research (Clifton 1732); Fordyce 1793; Haygarth 1805; Tröhler 2007). Quantification of data derived from them was a new research tool.

Building up traditions

This 18th century mental bent was shared by many clinical investigators. They participated in the endeavours and communicated results in an informal network between various cities. It cumulated in a textbook by **William Black** (b.1749), a London dispensary physician, entitled *Arithmetical and medical analysis of the diseases and mortality of the human species* (Black 1789). Yet, the new numerical approach was also criticized (Niebyl 1977): by 1800 this sketchy probabilistic *Evaluation Science* – the 18th century 'Evidence-Based Medicine (EBM)' which I have mentioned above – had to compete, in research methods, with clinical and pathological observation and description (*Description Science*), laboratory experiments (*Explanation Science*), the study of medical classics (still!), and in practice – as always – with dogmatic routine and fashions (Tröhler 2005).

Clinical arithmetic continued in the early 19th century. In 1819, Sir **Gilbert Blane** (b.1749), a former naval physician, now a distinguished Fellow of the Royal Society, Gulstonian Lecturer, Baronet, (and later *Physician in Ordinary* to two kings), summarized the insights he had gained during his experience in various walks of life. At the age of seventy he published an account of it in his *Elements of medical logick* in terms of a typical British compromise between two epistemic camps, the rationalists and the empiricists. It was a plea for rational empiricism as we would define it:

> [Rationalism and empiricism] *ought not to be regarded as adversaries, but as allies, and… good sense will consist in… fairly appreciating which is due to each. This is a compromise.* And further: *It is only by a sort of arithmetical computation, founded upon large averages, that truth can be ascertained; and hence the danger of founding a general practice on the experience of a single case, or a few cases [be avoided]* (Blane 1819, pp 200, 208).

This was unconscious, informal probabilistic thinking at its best as it would henceforth be characteristic of British epistemological literature. In 1823 – Blane's book was in its third edition – when **Thomas Alcock** (b.1784), an apprenticed surgeon turned London practitioner and workhouse surgeon, continued the tradition with a 61-page "*Essay on the education and duties of the general practitioner…containing suggestions relating to the investigation of disease, and the registration of practical results",* 1823).

And a decade later, **Tweedy J Todd** (b.1789) an Edinburgh MD and former naval surgeon, made the tradition explicit, although he entitled his book *The book of analysis or a* new (my italics) *method of experience…* (Todd 1831) to encourage physicians and scientists to apply the Baconian experimental method. Both authors propagated various complicated tables for clinical signs and results of different treatments to be compared as "*an easier, a surer method*" for obtaining a better use of experience (Alcock 1823, pp 85, 93, 95, 99, 100; Todd 1831, pp 86-104, 162-163, 184). Alcock repeated 18th century positions when writing in 1823 that tables

> *may be constructed so as to exhibit the general result of all the cases of diseases which have fallen under the student's observation […]. The advantages thus obtained, by enabling the student to generalize the facts, to compare the result of various modes of treatment, […] are too obvious to be dwelt upon* (Alcock 1823, p 99)

Todd, a former naval physician that is, hierarchically a subordinate of Blane, saw the dawn of a future science founded upon this practice, provided students were "*Thoroughly disciplined in* classical[!] *and mathematical*

learning" (my italics) ([Todd 1831](#), p 121, 159). Alcock mentioned probability *en passant* in a footnote (p 78). Neither dealt explicitly with quantification. Yet their endeavour implied unconscious probabilistic reasoning…

By 1830 the movement found a second textbook-like summary in **Francis Bisset Hawkins**'s *Elements of Medical Statistics* (1829). In this updated book in English on medical statistics, the young London physician (b.1796) had delivered its contents as the Royal College of Physicians' Gulstonian Lecturer of the previous year. He summarized the status quo as follows:

> *Statistics has become the key to several sciences. […] And there is reason to believe that a careful cultivation of it, in reference to the natural history of man in health and disease, would materially assist the completion of a philosophy [science] of medicine […] Medical statistics affords the most convincing proofs of the efficacy of medicine.* [And he specified:] *If we form a statistical comparison of fever treated by art, with the results of fever consigned to the care of nature, we shall derive an indisputable conclusion in favour of our profession* ([Hawkins 1829](#), pp 2-3).

If the word "statistical" were replaced by "arithmetic, or "numerical" in the above sentences, it might well have been written fifty years earlier by a prolific and militant writer on the subject, **John Millar** (b.1733), another Scottish physician at the London Dispensary (Tröhler 2005).

So, clearly, there was methodological awareness in Britain tied to probabilistic thinking throughout all these decades, mostly in an unconscious mode. It is therefore not surprising that Louis's work was scrutinized in Britain.

British perspectives on prominent French researchers

Louis seen through British eyes

As we know, Louis contributed notably to two fields – to anatomo-clinical research, and to the notion of numerical evaluation of therapies. He was influential in Britain in both areas – although British writers did not refrain from pointing to earlier British authors. When "*the uncertainty of medicine, and…the numerical method of Louis*" were the topic of a paper read at the London University Medical Society on 16th November 1836, a discussant reminded the audience that this course of action "*did not originate with Louis. The late Mr. Alcock reported his cases numerically*" (The Lancet 1836-37). He was aware of the fact, presumably also because [Alcock](#) had repeated his 1823 plan in his published *Lectures* (1830, pp 61,115). Another example was Thomas Hodgkin.

Quantification in anatomo-clinical research

When, in the 1820s, **Thomas Hodgkin** (b.1798), a Quaker, interrupted his studies in Edinburgh and spent over a year in Paris, he became so influenced by the exacting pathological and numerical approach of Louis that, back in London and in charge of the Department of Pathology at Guy's Hospital, he also started carrying out autopsies and collecting the reports. However, Louis's methodical numerical system was "*not altogether new or singular in his hands",* as Hodgkin rightly pointed out with reference to seemingly independent British work (Todd among others) (pp 1092-93).

In 1834 Hodgkin lectured to the Physical Society of Guy's Hospital on the *Numerical method of conducting medical inquiries* (Hodgkin 1854 [sic!]). He deplored the conjectural state of medicine. This could only be overcome by strict adherence to precise descriptions of many clinical cases followed by equally precise inspections at autopsy among those ending fatally; to grouping and presenting them in tabular form for comparison and statistically analysing them. This was Louis's extolled anatomo-clinical- research method for the description of disease (nosography) which was unquestionably useful, as in "*therapeutics, in which the assistance of the numerical system [was also] required*". Foreign students understood this, while the French principally followed the ideology of Louis's elderly enemy, **François-Joseph-Victor Broussais** (b.1772): The much younger Hodgkin deployed the advantage of Louis's rigorous, objectivating approach at great length, culminating in caustically comparing it to Broussais's doctrinaire system:

> *The unquestionable talent and powerful sarcasms of [this] author…under which many other systems and authors have seemed to give way, the physiological views upon which he so arrogantly plumes himself, and the authority of his name… make unitedly but a miserable figure when confronted with the counted facts of his accurate and statistical opponent* (Hodgkin 1854, p 1093).

Remarkably, he revealed his conscious mode of probabilistic thinking when concluding that, with respect to the past, the numerical system was,

> *indeed, an invaluable method…; but with respect to the future, it appears to me to be rather an application of the doctrine of probabilities: this however, I do not advance as any objection to its application.*

Indeed, he reproduced his lecture twenty years later with the motivation that this would not only serve to exhibit what the numerical system was, but also what it was not, *"and to what it cannot effect, and thus counteract a* cacoethes numerandi [Hodgkin's italics], *or abuse of statistics, to which, as a statistician, I should object"* (p 1090).

This was a British compromise, a warning, lest pure numerists should despise inquiries in which the *méthode* need not to be employed. For there were many cases, if "*carefully observed and detailed, possess very considerable value even when they stand alone"* (p 1093).

This held for descriptions of pathology, Hodgkin's main subject, for which he referred to Louis before the Paris debates about evaluation of therapies took place, and Louis's respective work was published.

Evaluation of therapy

Louis's evaluative work and the Parisian debates were promptly received in Britain. As early as 1835, *The Lancet* reviewed his *Recherches sur les effets de la saignée* (Researches on the effects of bloodletting, 1835):

> *The method adopted by M. Louis may be easily stated. If the [natural] mortality in pneumonia were known to be 25 in 100, and its mean duration twenty-one days, it would only be necessary to subject a considerable number of similar cases to a particular treatment, to count the deaths or recoveries, and to take the mean duration, in order to state in precise terms the modifying power of the treatment. The same process would serve to compare or contrast the effects of two systems of treatment …From data thus furnished the results could be calculated, thrown into tables, and readily compared* (Lancet [1834-35]. Essay review p 84-85).

To make this plain, the reviewer adduced "*a few examples, selected from several others now before us*": There was the report by **Sir James McGrigor**, Director (Inspector) General of the British Army's Medical Service, about different treatments for syphilis. Indeed, McGrigor (b.1771), when still head of Wellington's Army Medical Corps during the Napoleonic wars, had developed a statistical bent, drawn up regular reports, and had stimulated others to do so. Since 1815 McGrigor had drawn on thousands of cases treated by his subordinates ([Tröhler 2000](#), pp 100, 104, 108-110). So, it was easy for *The Lancet*'s reviewer to criticize Louis's data compared to McGrigor's arrangements. The number of patients studied by Louis was too small, he held, and "*this distinguished pathologist falls into some errors*" by not taking account of patients' ages: by reference to British life-tables, it had become clear that mortality varied by age, and this might have accounted for the differences observed by Louis.

Furthermore, Louis had not presented his data in tabular form – and *The Lancet*'s reviewer made up for that by dressing them up in two tables! Of course, he aimed "*to separate facts from opinions, dissipate the scepticism* (sic!) *which some have entertained concerning the utility of medicine, and raise it to the rank of an exact science*" – just as had been formulated in Paris (Lancet [1834-35], pp 85, 87).

A few weeks later another review of an earlier work by Louis, was also published in *The Lancet*.   It was on the *Pathological Researches on Phthisis,* Louis's first field of activity so much admired by young Hodgkin. It had meanwhile been translated into English by Charles Cowan. The editor was

> *glad of a new opportunity of drawing attention to the subject* [and hoped] *to prosecute and promulgate the numerical doctrine with effect in this country.* [It was] the *method of weighing medical facts and establishing medical principles, by counting and comparing them, and registering quantities, majorities, and minorities* (Lancet [1834-35], Review of Louis, p 292).

And it implied probabilistic thinking, albeit in the unconscious mode.

The rest of the article consisted of an extensive quotation from the British translator's introduction. **Charles Cowan** (b.1806) was a well- educated young physician. He held an Edinburgh MD and a bachelor *ès-lettres* of the Sorbonne in Paris. He had spent nearly four years in the hospitals of Paris; and he was personally acquainted with Louis, having followed his clinical rounds and assisted in his post-mortem room. First, he drew attention to the principles of observation according to that famous 17th century Englishman, **Thomas Sydenham** (b.1624). Cowan then went on once again to emphasize the need for large numbers of cases, to be presented in tables. He also stressed the need to consider patients' ages, sex, the severity of disease, its natural course, and the characteristics of the epidemic at a given time when submitting "*all* [...] *facts to the unerring test of arithmetical analysis*". All these efforts were to be made in order not to overlook fallacies, particularly in therapeutics: "*No part of medical knowledge is more in want of a rigorous method of investigation [...]*". One had to overcome preconceived ideas and selection bias to obtain the necessary fair comparisons. Put in one sentence:

> *It is not our intention, in advocating the numerical method, to conceal for a moment its difficulties; these are great and numerous, but at the same time they can never form any solid argument against its utility, though they will necessarily curtail the number of its disciples* (pp 295- 296).

What sober, yet sensible and farsighted words! McGrigor's and Cowan's reviews illustrate the culture of numerical accountability, unconsciously linked to probabilistic reasoning, that *The Lancet* propagandized. It can further be illustrated by yet another 'home grown' British contribution a few years later.

Probabilistic reasoning in 19th century British clinics

William A. Guy (b.1810), a young London-educated physician with a Cambridge BM, described the state-of-the-art of medical epistemology from a British point of view in a 22-page article entitled *About the value of the numerical method as applied to science but especially to physiology and medicine* (Guy 1839). Like Hodgkin, he had studied on the continent, in Heidelberg and especially in Paris (whence he was well aware of Louis's work). He had become interested in statistics, and he wrote for one of the first issues of the *Journal of the Statistical Society of London.*

Guy's analysis was original in that it based its strong advocacy of the numerical method on the history of science. "*A very obvious and most important application of statistical investigation is as a test of the truth of theories*". The idea of hypothesis-testing with the help of quantification was an immense epistemological step. Unconsciously it implied probabilistic reasoning. With reference to astronomy and chemistry he continued: "*The certainty of a science is exactly proportional to the extent to which it admits of the application of numbers*". Numerical probability became the natural substitute when numerical certainty was not available (Guy 1839, pp 30-31, 34, 37). Now, this was obviously the case in medicine, with its variable quantities and events. And so, medicine was also amenable to perfection through adoption of the numerical method as had been the case in astronomy and chemistry. Indeed, he diagnosed *"…a growing disposition to apply calculation to the phenomena of life, …as one of the characteristics of the age in which we live*" (p 35).

To start with, Guy referred to Bisset Hawkins's *Elements of Medical Statistics* (1829). It was the only reference to a statistical work he gave. From there he developed a methodological hint, namely "*a best rule [...] for*

ascertaining whether the observations which we have collected are sufficiently numerous to yield a true average" (pp 32-33). For medicine this was obvious in vital statistics and nosography (description of diseases). And then he continued:

> *...as to the action of remedies, and the relative advantage of different modes of treatment – nothing can determine these but an accurate numerical comparison of their fatality and duration under the several methods of treatment proposed* (p 40).

This statement was comparative, quantifying and consciously, yet informally, probabilistic. But as a current example of formal probabilistic thinking, Guy referred to the trustworthy estimate of the benefit of vaccination against smallpox, and of the extent of protection it gave (p 46). This was the example in which a calculus of probabilities had been used 60 years earlier to show the utility of inoculation of smallpox. Numbers added precision to words of doubtful meaning. Yet Guy admitted that the application of this method to individual cases was limited. It was again the group-versus single-case-issue: Medical practice needed "tact" – what the French referred to as *art médical*. But, on a closer look, this "tact" was nothing more than "*a rough calculation of chances in which all the elements of calculation are rapidly seized and accurately estimated*". This was conscious probabilistic reasoning. Guy saw this as "*common sense to men in ordinary affairs of life*" (p 44) and we may recall here Laplace's "common sense reduced to numbers".

Although Guy could have taken these arguments from 18th century Scottish doctors, or literally – yet improbably – from John Gregory (see p 1above), his qualification of the numerical method was new. In his words it was

> *to supply the want of tact by furnishing the inexperienced with accurate calculations of the probable event of different diseases, and the* probable [my italics] *consequences of different modes of treatment. These calculations supply but one element for the solution of the problem, for they apply only to cases of average severity* (p 42)

In other words, they applied to the average of a group. And the important practical consequence of this apparent limit was entirely new, too: these calculations "*leave to the physician the task of ascertaining all the circumstances in which any particular case departs from the average severity*": This was clearly probabilistic thinking (p 42-43). And it was not obstinately doctrinaire, but pragmatically realistic.

Thus, Guy's definition of the numerical method differed from Louis's or Trousseau's. His was the method of averages (p 32). These essentials being specified, he had neither the intention to discuss the numerous errors with which a statistician could be charged, nor to "*defend him and his method against the objections as well as the ridicule of his opponents*". For the errors were precisely those to which the results of common observation, expressed in common language, were equally exposed (p 43). Indeed, at the outset of his reflections, he had enumerated the necessity, the difficulties and pitfalls of precise observation, of grouping of comparable facts in tables, of large numbers etc. (pp 31-35).

As the French clinicians had done, Guy distinguished medicine considered as a *science* from medicine considered as an *art*. Yet the latter was medicine's disadvantage, for as a practical art it must necessarily remain imperfect, whereas nothing could hamper it from attaining a high degree of perfection as a science. Calculation was the necessary feature of a science, or inversely: "*Rob science of calculations and we degrade it to an art*". He concluded optimistically "*...we may rest assured that every addition made to our science will be [also] a gain to our art*" (pp 45-46).

All these prevarications about probabilism condensed into a clear-cut statement in a 21-page review of Gavarret's *Principes Généraux de Statistique Médicale* (1840) published immediately in 1841 by *The British and Foreign Medical Review*. Because the reviewer mixed the contents of Gavarret's book and his own input, very much taken from Guy, I believe he wrote it.

A British view on Gavarret

Considering the independent British tradition of unconscious mode of probabilistic thinking, it is instructive to consider how Gavarret's book was reviewed. The text provides a kind of theoretical standard of probabilistic reasoning after 1840 (British and Foreign Medical Review 1841; Agnew 2008).

The long introduction advocating the NUMERICAL METHOD (author's upper case) repeated at length Guy's arguments about the status of medicine within the sciences, and their history. It also referred to Todd's tabular analysis (described above) as a development of Baconian precepts. But then the review developed two novelties: the clarification of definitions, and the inclusion of the calculus of probabilities:

> The numerical method is sometimes erroneously regarded as a mere substitution of figures for words. Against this mistake Gavarret strongly protests, and with good reason, though th[is] mere substitution is [already] a great improvement in our scientific methods, seeing that figures admit of strict comparison which words do not (p 13). [According to Gavarret] medical statistics, or as we prefer to call it, the numerical method, is "la théorie des grands nombres," the application of the calculus of probabilities to the science of the physician, "le complément le plus indispensable de la méthode expérimentale" (pp 16-17).

This terminological clarification was deemed necessary to specify that the use of numbers as a research tool differed from the historical meaning of statistics as the science of the state (p 12). The calculus of probabilities supplied a method to determine the limits of error of our observations, a method, for instance, to specify the limits of confidence of a difference between two treatments (p 19). For "the medical brethren as are conversant with the mathematics", the formula was included in the review (p 20). Examples followed, amongst others some from Louis, with the reviewer's appropriate critique of his having studied too few cases (pp 19-20). And he continued:

> If even Louis [...] lies open to censure, what shall we say of the majority of his followers, and in what terms shall we speak of those who still persist [...] in drawing important conclusions from one or two scattered and not comparable facts (p 18).

What was to be done? One could just repeat:

> Once more, then, the SCIENCE OF MEDICINE wants facts – comparable facts – numerous facts: well observed, carefully arranged, minutely classified, and acutely analyzed. Her language must be the language of figures; her test, the calculus of probabilities; her example, the most perfect and exact among the sciences of observation and experimentation [i.e. astronomy and chemistry, respectively] (p 21).

This was certainly a truly remarkable passage, an example of the proverbial British pragmatism re-uniting the perennial quest for dogmatic certainty with the proposed yet reasonably unattainable model of probability under review to form practical modus vivendi.

Traditions persist

If only these words had been taken to heart! Yet the application of formal probabilistic procedures was hampered, within medicine, by ignorance, many inherent difficulties, (for example in collecting comparable cases), confidence in presupposed certainty of technological innovations, and socio-culturally, by hierarchical authoritarian structures within medicine. Despite the methodological insights into the possible sources of errors described, many but not all doctors continued to compare small groups, often incomparable, to select cases, to fail to take account of the natural course of diseases, to fall into the post hoc-ergo-propter-hoc fallacy, (for instance in the appreciation of homeopathy), even to cheat. Other doctors were more satisfactory (Tröhler 2014; Holmes 1861; Barclay 1864; Andrew 1891).

Interestingly enough, (as I have shown above), Hodgkin saw no necessity to change anything in his text, when re-producing his lecture twenty years later (Hodgkin 1854), rather he had a new motive. He was in fact no clinician, but he had become a leading British pathologist by then: Hodgkin's disease is still known to-day.

Likewise, Guy's Croonian Lectures of 1860 bore a title similar to his 1839 paper (described above), namely, *The numerical method and its applications to the science and art of medicine* (Guy 1860). By now engaged in forensic medicine and public health statistics, Guy had noticed the current of the times: He had read Gavarret as I concluded from its review in the *British and Foreign Medical Review* and therefore now emphasized the importance of calculations in the "theory of probabilities [his italics], *and in that numerical method which may be said [to be] one of its principal branches, and certainly its most important one.*" (p 331). He drew on Gavarret's "*mathematical formulae for calculating the limits of possible error attaching to any given number of facts*" – criticizing again Louis for his insufficient numbers. But if in practice one stuck to Gavarret's

> *...ordeal of a mathematical formula, we shall be driven, if not for state purposes where we can command almost unlimited supplies of facts, [...] to forego the use of the numerical method altogether."* So, he asked:

> "*Is there not, in the absence of certainties, at least a fair probability, that the average results of even a small number of facts may be entitled to confidence?[...] is there not some escape from the very disagreeable dilemma of being obliged to reject all average results except those derived from one thousand, two thousand, three thousand facts?* (p 469)?

Yes, there was one. Namely when the difference of averages was either none or uniformly very marked, or "*when we have to do with the apostles [...] of wild medical heresies.* Then "*a heavy battery of facts and figures*" was not always required to demolish a hypothesis (p 554). He offers as an example four sentences quoted from his friend **Thomas Balfour** (*1813). These describe a controlled trial to test the claim that belladonna reduced the likelihood of developing scarlet fever. "To prevent the implication of selection", Balfour took 151 boys in a military orphanage and assigned them alternately either to receive (76) or not to receive (75) belladonna. Two boys in each group developed scarlet fever. Balfour properly cautioned against inferring on the basis of such small numbers that belladonna had no prophylactic effect (Balfour 1854), whereas for Guy this evidence was sufficient!

Guy concluded his lecture series by recounting Ambroise Paré's amputation trial of 1536 (Paré 1575) as "*an unconscious illustration of the value of that numerical method which I have endeavoured to set before you to explain, and to vindicate ...*" (p 597).

Outlook

Guy's special position becomes clear if we see him in the context of a well-known clinician of his age, **Andrew Whyte Barclay** (b.1817). This Edinburgh and later Cambridge MD lectured to students for twenty years, till 1882, on pathology and therapeutics at St. George's Hospital in London. Thus, he was a man of the establishment. His attitude towards the methodological aspect of medical research may be derived from his Lumleian Lectures for 1864 which bore the title *Medical errors...Fallacies...of the inductive method of reasoning to the science of medicine* (Barclay 1864). They were all about the ways to find laws of cause-effect by experimentation in the Baconian tradition. Yes, large case collections, tabulated in statistical form, were better than to trust in memory, and there might even be a way to induce a causal relationship (p 17). But they were simply used for calculating averages, and thence for deducing erroneous assertions "*of the curative power of a remedy*", which was anyway not applicable to individual cases (pp 57-58).

Barclay shrewdly analysed the therapeutic inquiries issued lately by the British Medical Association on therapies of pneumonia (with and without bleeding), non-syphilitic psoriasis, tapeworm, and scarlatina. He quoted huge accumulated Parisian and London statistics (not deemed necessary) Dr. Balfour's controlled trial and even William Guy, without any details. He did this from the standpoint of principles of logic without entering in the ways of thinking behind them. No kind of probability was mentioned, let alone the calculus. Rather he repeated a typically paternalistic hackneyed saying:

> It is much to be hoped that scientific medicine may ere long be delivered from this, the oldest, the most obstinate, the most universal fallacy...the most constant theme of logicians of all times-

[namely] *the* post hoc ergo propter hoc [Barclay's italics].

But how? Here he quietened down. Barclay's were state-of-the art lectures, overall critical, full of warnings, but without a vision for advancement:

> *The numerical method has not yet been applied to any great extent in therapeutical* [sic!]
> *inquiries. The difficulties attending its employment are so great, and the method itself so open to*
> *fallacy, that the results are not likely to be very available for scientific purposes* (pp 116, 119-120).

This attitude, which he repeated, less outspokenly, 17 years later in his Harveian Oration to the Royal College of Physicians (Barclay 1881), was not constructive, yet it may very well have been representative of the views of a great majority of physicians:

> [The] *curative power of a remedy* [was asserted] *because in ten, twenty or even a hundred cases*
> *recovery followed its administration; and* y*et this is what is commonly meant when experience is*
> *appealed to* (Barclay 1864, p 119).

Under these circumstances, for the few who thought about methodological issues, a numerical method was, for the time being, the solution of the complicated problems of day-to-day clinical practice. Insofar as Guy, for instance, had found some well-designed and cautiously interpreted trials, (as judged by contemporary insights), he confirmed a genuine British tradition of enlightened pragmatism (Pickstone 2000), whereas the French mathematical tradition, sophisticated formally by Gavarret, was only theoretically valuable.

The general situation in practical therapeutics was one of *laisser-faire*. But German authors would not leave things there. As soon as Gavarret's book had been edited in German in 1844, young clinicians faced its theoretical and practical difficulties. They developed his mathematical concepts and applied them in practice.

Theory and practice of probabilities in Germany after Gavarret:

Introducing German dramatis personae
Theory

Books by Lind, Gregory, Haygarth and Black quoted above had all been translated into German by the end of the 18th century – but not into French (with the exception of Lind). I have been unable so far to find any of their methodological probabilistic passages referred to in the wider contemporary and early 19th century German literature. (German medicine was trapped for a time by the speculative philosophical systems of romantic medicine (Wiesing 1995). As we have seen, this had changed by the mid-1830s, Paris's new hospital medicine attracted open-minded, frustrated German students after the end of the Napoleonic period. The emphasis was on clinical examination including the ultra-modern auscultation, and Louis's anatomo-clinical outlook based on great numbers of patients and bodies, respectively. Naturally, from the 1830s onwards they came across the *méthode numérique* in one way or another, and some were also aware of the associated academic debates. That explains why we find references to these issues by German doctors from this latter period onwards, for instance by Jacob Henle.

Jacob Henle (b.1809) was one of the typical, keen young German doctors who visited Paris (1837). When he recalled this some years later as young professor of anatomy at the university of Zürich, he was reflecting on methodologies, and on the right way to acquire knowledge in medicine: His *Medizinische Wissenschaft und Empirie* [Medical science and empiricism] ended in a plea for (British) rational empiricism. I seem to hear Gilbert Blane – his work had been edited in German in 1819 – when I read

> *But not only to fill the deficiencies of both parts should empiric and rational medicine be linked to*
> *each other, but to foster one another where both can be applied simultaneously* (Transl. from
> Henle 1844, p 34).

Of course, Henle's warning of the danger of falling into of the *post hoc-ergo-propter-hoc* fallacy by basing one's practice on successful single cases had also been raised by Blane (1819, p 226) and Guy (1860, p 554).

Henle moved to Heidelberg in 1844, the year that Gavarret's landmark book came out in German. Two years later, he was the first German I have been able to trace so far who referred to it. And this was the only precisely quoted reference in Henle's 19-page text "*On doctors' methods*" at the beginning of the introduction to the first volume of his *Handbuch der rationellen Therapie* (Handbook of rational therapy, 1846). In admirably worded sentences he summarized the contemporary epistemological basis of therapeutics and, in a farsighted way, looked ahead. Of course, he also came to speak of Louis, whose

> *numerical or statistical method [...was] the only one the application of which might let us expect some advantage of empirical medicine, for [*and here he paraphrased Gavarret*] claims derived from experience never feature logical certainty, but only a major or minor grade of probability, [...]and even the so-called laws of nature have only the highest grade of probability".* As to therapeutics, Henle pointed out that "*numbers only determine the grade of probability with which we can deduct a given effect from a given cause and which may entitle us to prophesy the same effect from the same cause in the future* (Transl. from Henle 1846, pp 12,15).

Henle then gave precise methodological guidance: The number of cures obtained after a particular therapy had always to be given in relation to the untreated or otherwise treated patients, i.e. compared to the natural course of disease or to a control group; adherence to a therapeutic regimen by patients (and physicians) had to be supervised etc. Therefore

> *men, who are as familiar with the value as with the shortcomings of medical statistics, want[ed] to base their calculations on nothing else than upon hospital practice* (p 17).

Henle had obviously read about the Paris debates and was aware of the problem of the 'group-versus-case/individuals- issue'. The solution lay in "*tact*", and "*tact cannot be taught, nor is it inherited, what is inherited is only the talent to acquire it*" (p 18), and this acquisition needs time, otherwise practice is thoughtless routine.

Finally, he did not eschew

> *...the cliffs that lay in empiric medicine...The less control a doctor is to be afraid of and the more splendid the rewards are in this world...the nearer is the danger that not only the superficiality of self-deception, but also true, mean fraud obfuscate facts so that the course of the successors is lead astray* (p 17).

Bias and vested interests had already been acknowledged by the "fathers of probability", such as Jacob Bernoulli and Laplace, as possible implications. They are still huge problems today.

Henle's entire methodological introduction was written against the obviously prevailing strict separation of the empirical from the rational method – another age-old issue (Matthews 2020a). It was one of the rare pleas for rational empiricism as it had been propagated by 18th century British medical arithmetical observationists: "*Both were made ready to amble henceforth friendly close side by side*" (Henle, p 19).

Henle did not follow this track further. After all, he was a professor of anatomy and not a clinician. He later worked in Göttingen, and he was soon to acquire a world-wide reputation. Henle's loops in the kidneys are just one example of his many contributions. But younger German clinicians (who might have read this early book of his during their studies), took up Gavarret. In Tübingen particularly, a network established itself from the mid-1840s around Carl Wunderlich.

Methodology for evaluation: a first Tübingen circle

Carl Theodor August Wunderlich (b.1815) spent two postgraduate terms in Paris – in the winter of 1837/38 and the summer of 1839 – that is, precisely when the deliberations of the *Académie Royale de Médecine* were still

very much in the air. According to him, the numerical method was practised sloppily. Its usefulness was anyhow very restricted:

> *If the numerical method, provided it is correctly used, may have some value … for the diagnostic and prognostic significance of some phenomena, it certainly is devoid of any use/profit, a drawback even, for the decision of pathological and therapeutic problems. … How can one altogether dare to determine a therapy with it?* (Transl. from Wunderlich 1844, pp 41-42).

The reason behind this rhetorical question was once more the 'group-versus-case/individuals issue'. But above all, he considered this method to be inhumane, when implying human experimentation. As an experience in support he mentioned, in remarkably sarcastic words, an experiment for which a French physician divided patients with typhus into three groups (bloodletting, laxatives or nothing

> *…explicitly without any selection* [he did not say how] *and with undaunted tenacity until death. I could not help the impression that we live in times more barbaric than when criminals sentenced to death were used for [testing] operations or physiological experiments. Medicine's first duty is indeed scientific research; however, all his objects should be more holy to a doctor than to an entomologist, who transfixes his beetles without mercy* (Transl. from Wunderlich 1841, pp 41-43).

Yet by 1851, Wunderlich, now professor of internal medicine in Tübingen, had made a complete *volte-face*. He realized the classic confusion of a method as such with its incorrect execution, both scientifically and ethically. He now also recognized that therapeutics was in a crisis. To him it was like a basket filled with a mixture of personal beliefs, authorities' reminiscences, and a variety of systems; in a word – therapeutics was in a state of "systematic charlatanism". Thus, it needed a strong, reliable basis, and only mass-observations and statistics could and should provide this: "*Every doctor should be a statistician,*" he wrote (Transl. from Wunderlich 1851, pp 107, 110).

Of course, he referred to Louis and criticized him for often having asked the wrong questions, for example concentrating, crudely, only on the final result, on "cure or death". This also made clear that statistics had so far not achieved much. Wunderlich also criticized Gavarret for his quest for 400 cases, because this made "*any application of statistics impossible*" (Wunderlich 1851, p 111). He might have understood all this from discussing the methods to be followed in evaluation research with an already potent former collaborator in his Tübingen clinic, Wilhelm Griesinger.

Wilhelm Griesinger (b.1817) had been a slightly younger friend of Wunderlich' s since their college days. He too had been in Paris twice (1838, 1842), and in Vienna (as had Wunderlich). When the latter became professor at Tübingen in 1843, he engaged Griesinger as his assistant. Subsequently, Griesinger became *Privatdozent,* and extra-ordinary professor there, before leaving for Kiel in 1849. Like his friend, he engaged in reforming the ossified medical system in Tübingen. Very early on, he grappled with the methodological issues that had been discussed in France. His article *Zur Revision der heutigen Arzneimittellehre* (On the reform of today's pharmacy, 1848) was published in *Archiv für Physiologische Heilkunde* (Archives for Physiological Medicine)*,* which had been co-founded by Wunderlich (Griesinger was actually its editor at this time). He described the *status quo,* knew the literature, had come across Louis, read Guy and Gavarret in detail, and obviously knew all about Wunderlich's experiences in Vienna and Paris.

Griesinger was all in favour of numerical and statistical methods. Quoting Guy (Griesinger 1848, p 39), he noted:

> *… the "sometimes" of the prudent – […] – is the "often" of the sanguine, the "always" of the empiric and the "never" of the sceptic; the numbers 10, 100, 1000 [however] have the same meaning for everybody* (p 6).

Provided the method was used correctly! Lamentably, doctors were still unfamiliar with the notions of precise observation, note-taking, comparability of cases, comparison without selection of cases etc. Their "common experience" was nothing else than "mere conceit" (p 5-8). To reform all this one needed to bring together

rational theory and empiric facts, both based on accurate observations, not on the philosophical speculations of German romantic medicine. This was the rational empiricism propagated in Britain since the 18th century arithmetic observationists. Mathematisation would be the next step, as in every true science. This needed time and confidence – and a method for *a posteriori* calculus of probabilities. Gavarret's was impracticable. In the whole world one would not find an institution allowing for the assembly of 200 similar cases, at least, per group, that is 400 for two groups, to be compared! Adding up cases from the literature did not work because of their heterogeneity. But "*an association of many hospital and civil doctors working together according to a predetermined plan*" might get around this difficulty (p 8-11). Therefore, seen from Gavarret's standpoint, Louis's results obtained from relatively small numbers were valueless – quite apart from the fact that his cases had been selected.

But anyhow, figures, even mathematically calculated valid differences between groups *à la* Gavarret, were not everything; they needed *interpretation*: Who died, what of and when did he die? In that sense, smaller groups could also be valuable. And there was the problem of the relevance of mean values the individual cases. Much remained to be done (p 22)!

Although approved, Louis was also criticized by yet another friend of Wunderlich's and Griesinger's, Friedrich Oesterlen. All three had studied together in Tübingen and had visited Paris and Vienna in the 1830s.

Friedrich Oesterlen (b.1812) had become *Privatdozent* in Tübingen together with Griesinger in 1843. Three years later he left for a full professorship at Dorpat, whence he had returned to Heidelberg as a *Privatdozent* in 1848. With the hope of an academic re-start at home he started to publish extensively, for instance, *Medicinische Logik* (1852), which was published in English as Medical Logic, by the Sydenham Society (1855).

Oesterlen had also read Gavarret. He now aimed to set the issue of statistics in a theoretical context by applying in medicine (pVI) the teaching of J. Stuart Mill's *System of Logic* (1846) (Bailey and Howick 2018). Oesterlen's book dealt with medical observation, the concepts of induction, deduction, generalization, experiment, experience *and* statistics. But valid *scientific* results consisted in the discovery of causation, not just in the discovery of statistical correlations. He wrote about the contribution of statistics in general terms as "*the essential step in our research on truth based on experience*". This held as long as one kept to the rules of extremely precise observation, compared comparables, considered the natural course of diseases, collected large numbers to establish high grades of probability. He was very cautious about generalizations (Oesterlen 1852, pp 129-140) and hasty conclusions, as had been the case with Louis. These could lead to nonsense, and the general error of internal medicine was the *post-hoc-propter-hoc fallacy*. Comparison was needed (pp 129- 140).

This work would prove to be a major contribution to the methodological discussion in Germany (Rothschuh 1968. However, neither Oesterlen nor Wunderlich mentioned the calculus of probabilities in this context, in contrast to their friend Griesinger. It was beyond their horizon at this time.

Having thus initiated probabilistic thinking into therapeutic evaluation early in their careers, these three Tübingen friends moved on to further responsibilities: Wunderlich went as chief of internal medicine to the university of Leipzig. Griesinger succeeded him in the Tübingen chair; he turned more and more to reform of psychiatry. Oesterlen did not succeed academically in Germany. After much publishing on various issues he retired to private practice, eventually in Switzerland.

But the *Archiv für Physiologische Heilkunde*, edited after Wunderlich by Griesinger and now by another Tübinger, the physiologist Karl Vierordt, continued to open its pages for a major contribution to the field from Georg Schweig (b.1806), a little-known doctor turned civil servant in the Grand-Duchy of Baden. He wrote a fifty-page paper entitled *Auseinandersetzung der statistischen Methode in besonderem Hinblick auf das medicinische Bedürfniss* (Deliberation about the statistical method with a special view on medical needs, Schweig 1854). This continuation and expansion of Oesterlen's work was a contribution designed to explain the bases of statistics to doctors, given that, in Schweig's opinion, their statistical works were usually unusable because their authors were insufficiently knowledgeable about the principles and methods requested. The field was still in its infancy (Schweig 1854, pp 305, 349).

Schweig started his article by clarifying definitions: medical statistics were for him a special method for drawing conclusions (Schlussziehung) (p 307-309). He had clearly read Jacob Bernouilli, Poisson, and Gavarret (p 322-323), and he wrote at length on the establishment of arithmetic averages (means) of groups of cases. Such averages were only of any significance if compared with other averages. And for a valid comparison, it was necessary to know their "their limits of oscillation" (according to Gavarret). These had to be as small as possible. But the exact determination of the sufficient number of cases or groups to achieve this (by the method of least-squares) was too complicated. Therefore "*the probability is to approach certainty [simply] by further observations or experiments*" (pp 330-331). Thus finally, he set up the following rules:

- Know the state-of-the-art and ask a precise and sharply limited question
- Collect well observed cases according to a plan defined by the question, yet do not select them in ways that are biased by a preconceived idea
- Form groups and calculate averages (means)
- Draw conclusions based on calculations that accord with clearly stated conditions
- As to the validity of a conclusion, be aware that it always depends only on a probability, and that it is provisional until other works performed under similar conditions achieve the same result (replication), wherewith it rapidly approaches certainty (pp 351-355).

These rules were certainly clarifying, but they were not acknowledged by the medical world. Schweig was not quoted by a group of contemporaneous, yet somewhat younger mathematician-physicist-physiologist-physicians who advanced the methodology by developing tests of significance for assessing the meaning of differences between groups.

Testing the validity of comparisons, 1858-1877

All these theoretical contributions of the 1840s and early 1850s reflected probabilistic thinking in the unconscious (Wunderlich) and conscious, pre-mathematical modes (Henle, Griesinger, Oesterlen, Schweig). In the next two decades two generations of younger men acted in compliance with the formal, mathematical mode.

Gustav Radicke (b.1810), the first of this group to publish in the field, was its oldest member. He was only a professor *extraordinarius* of physics, that is without any strong institutional ties. In 1858, he published a very original paper in the *Archiv für physiologische Heilkunde* (Archives for Physiological Medicine), now edited by Wunderlich. It had a 30-word title which, when abridged, reads *Die Bedeutung und Werth arithmetischer Mittel …und Regeln zur exacten Beurtheilung…* (On the value of arithmetical means… and rules for the exact assessment…). This article contained a unique novelty for its time, namely *"a simple significance test that might render reasoning more assured and conclusions more persuasive*" (Coleman 1990, p 201). This method was designed not only for physiological experiments, but also for enquiries dealing with purportedly effective therapeutic measures. Radicke rejected conclusions derived from (Louis's) numerical method because, in his view, it only consisted in comparing arithmetical means derived from two groups, but this said nothing about the meaning of any difference between these means. Instead, he proposed comparing the differences between the means including their standard errors. This would show the degree of confidence that could be attributed to such a difference. When Radicke applied his test to some physiological and therapeutic examples it suggested that no effects had been produced (Coleman 1990). This was deemed impossible.

A storm in a teacup ensued over the next few years. The opposition was led by physiologists of Radicke's generation such as Karl Vierordt (*1818), the newly appointed professor of physiology at Tübingen), and Friedrich Wilhelm Beneke (*1824, who acted also as *Kurarzt* (spa doctor)*,* was still a *Privatdozent* with vested interests, which added confusion.

They argued that effects were due to a "*logic of determining facts"* and that probabilistic mathematics was a valid, but purely formalistic form of medical statistics without substance in the real world. This line of argument was later used also by Claude Bernard (as I have outlined above) The participants in this debate did not understand what Radicke's approach was about. In the end, determinism prevailed; Radicke and his test disappeared from the German literature (Coleman 1990).

Adolf Fick (b.1829) had read mathematics before turning to medicine, and he was to become a founding father of medical physics. He was ordinary professor of physiology in Zurich when he published *Anwendung der Wahrscheinlichkeitsrechnung auf medicinische Statistik* (On the application of the calculus of probabilities to medical statistics) as an *Anhang* (appendix) to the second edition of this textbook of *Medicinische Physik* (Medical physics, Fick 1866). He was familiar with the writings of Jacob Bernoulli, Laplace and Louis, and he "re-discovered" Gavarret. For, above all, Fick noted, it was thanks to Gavarret that

> *mathematical aids are handily presented to medical researchers ready for use. [...] Yet they still don't make comprehensive use of them.* And, after quoting from puzzling Beneke at length, he added: *Yes, even quite frequently, weighty voices have risen against them in principle* (Transl. from Fick 1866, p 430).

Fick took the application of numbers for granted. That was quite something. He agreed with Gavarret's critique of Louis. But it was "*clear that the interpretation of a statistical compilation is only and exclusively a matter of … [Laplace's] healthy common sense, that is …particularly of the calculus of probabilities*" (p 434). The next problem to solve was the elaboration of

> *…a covenant about the degree of probability one wishes to require. A certain measure is naturally to be observed. Since probability is more or less to replace certainty one must not be satisfied with too scanty a probability, e.g. it would be completely senseless to ask for a probability of only ½* (pp 430,434 and 440).

Then Fick said that one should neither go too far in the opposite direction. Poisson's choice of a probability of 212/213, or 99.53%, had had the rationale that it was based on a pragmatic compromise between either an unimpressive probability or an overly demanding sample size. So, this value was still near unity, the symbol of certainty. Fick now developed a formula and calculated a logarithmic table which permitted determination of the limits within which a probability was included. And this "*in less than five minutes*". It functioned for a rather large number of cases, at least not fewer than a hundred (pp 441-447).

That was a methodological advance, yet Fick did not contribute to solving the practical difficulty of the computation of hundreds of comparable cases. Consequently, the large number of patients required according to Gavarret (and Fick) continued to be criticized. Several ways to solve the problem were suggested. Wunderlich had, irrelevantly, proposed concentrating on the effects of a given remedy rather than on a disease because of the diagnostic uncertainties (Wunderlich 1851, p 111).

So, questions remained open. But new inputs were soon to be propounded by three physicians of an even younger generation born in the 1840s and then elaborated by an older, remarkably versatile colleague.

Theodor Jürgensen (b.1840), when still a *Privatdozent* at Kiel -he later became a professor of internal medicine at Tübingen-, applied the Poisson-Gavarret calculus in the methodical assessment of a historical comparison: There were 330 cases of abdominal typhus treated with the usual, purely dietetic measures (between 1851 to 1861), and 160 later cases treated with cold-water bathing (since November 1863). This study (1866) fulfilled many of the methodological conditions established so far: Jürgensen demonstrated the similarity of two populations in terms of age, sex, duration before hospitalization, and hospital conditions, meaning that they were more likely to differ only in the treatments they received. The crude death rates were 15.4% in the traditional group and only 3.1% in cold water group (they were further differentiated by the gravity of their condition). Jürgensen then applied the calculus of probabilities as proof that this difference was due to the different treatments. This yielded a probability of above the 99.5% required according to Poisson-Gavarret (Jürgensen 1866).

> *Therefore, [he said] it is also permitted to choose this stricter form of calculation, although the absolute numbers are not very large. … Fick's formula is insufficient, for* [the number of cases] *is too few.* [And he concluded] *maybe it is time just now to open the doorway for analytic statistics*

for non-specialists. This science, albeit hardly existing today... will in the future solve problems of which we now have not the slightest idea (Transl. from Jürgensen, pp 65-76, 129).

Willers Jessen (b.1840s?) was a young doctor in Kiel when he published an article *Zur analytischen Statistik* (About analytical statistics, Jessen 1867). He used this term meaning mathematical probability, for he had obviously read Poisson and Gavarret in German translations. He understood that "*Gavarret considers a result as valid when, and only when one can bet 212 to 1 that it is true*" (p 128). Accordingly, he had devised tables for the application of the calculus of probabilities, Fick had simplified them, and Jessen now provided one even more easily to use (p 130). He concluded, with foresight:

> ...perhaps it is timely just now to popularise analytical statistics more generally ... This science, albeit hardly existing nowadays, ... would in the future solve problems of which we had not yet a clue (Transl. from Jessen 1867, pp 128, 136).

He became a clinician in his father's psychiatric asylum and consequently changed his field of interest. This was also the case of Julius Hirschberg.

Julius Hirschberg (b.1843) – later a world-famous ophthalmologist, world traveller and historian of ophthalmology – had also studied higher mathematics and physics. In 1874, when still a *Privatdozent* in Berlin, he wrote a book with the enticing title *Die mathematischen Grundlagen der medizinischen Statistik elementar dargestellt* (The mathematical bases of medical statistics elementarily presented, Hirschberg 1874). There were tables reducing the probability that a difference was not due to chance from Poisson's and Gavarret's 212:213 (99.5%) to 11:1 (91.6%) – a quirk of a mathematical formulation of probabilities, the so-called odds formulation (Matthews 2017). This allowed the comparison of much smaller groups. The multi-talented Carl von Liebermeister considered this a great step forward.

Tübingen again

Since his youth, **Carl Liebermeister** (b.1833) had developed a deep knowledge and ability in mathematics: At 29 he had published an article on their *application in the physical sciences*. This was during his five years as assistant, *Privatdozent* and extra-ordinary professor of internal medicine at Tübingen (1860-1865) where Griesinger was his chairman. Then recently appointed professor at Basel, he became aware of young Jürgensen's work on cold water fever therapy, had it repeated by an assistant and reported the results statistically. The two books were reviewed in the *Edinburgh Medical Journal* (Edinburgh Medical Journal 1869). In 1871 Liebermeister returned to Tübingen as chairman of internal medicine, and re-entered the clinical statistics scene after 1873, when Jürgensen had also arrived there. Obviously the two colleagues met.

Now Liebermeister started working like a professional on the methodological issue that Jürgensen had dealt with a decade before in an amateurish way. In Basel he had wondered why the lethality of typhus patients was higher in his clinic when compared to Jürgensen's. This had to be explained. But he also had in mind to find a mathematical solution to the meaning of a statistical difference between two therapies. For this he developed a test of significance for such a difference (1877). He lectured on the issue, and sent a manuscript to two professors of mathematics, former colleagues from the University of Basel, for critical examination. They approved. The ensuing publication bore a similar, yet more specific, title than Fick had chosen, namely *Ueber Wahrscheinlichkeitsrechnung in Anwendung auf therapeutische Statistik.* (On the calculus of probabilities applied to therapeutic statistics (Liebermeister 1877). He named this mathematical solution a 'four-table-test', practicable for use in analysing very small groups [n<10], which however led to extensive calculations as larger groups were analysed. He included examples of both situations (Ineichen 1994, Senata et al 2004).

Assessments of the state-of-the-evaluative-art
Two comments from outside and inside Germany

These contributions had not passed unnoticed. A conscientious young medical historian doctor sensed that issues of evaluation and probability were in the air: **Julius Petersen** (b.1840) of Copenhagen gave a substantiated description of the situation in his *Hauptmomente in der geschichtlichen Entwicklung der medicinischen*

Therapie (Key moments in the historical development of medical therapy, 1877). In 29 pages he dwelled on Poisson, Gavarret, Louis, Wunderlich, and others, and on the numerical method. He quoted, rightly (and at length), Gavarret saying that there was much loose verbiage about probability whereas only the calculus of probabilities could really help to estimate the worth of mean values (averages). Albeit still far from being perfect, this method was important for future developments (p 179). Petersen qualified the confusion between Benecke and Vierordt: the former explained the effect of a cure, the latter thought that it was being demonstrated statistically. Again, this was the old debate between rationalism and empiricism. I think we can trust him when he said in the mid-1870s that in France, England and Germany polypharmacy and the *post hoc ergo propter hoc* fallacy prevailed, Louis's principles were lost sight of, but some British followers of Bacon were still eclectics and indulged in common sense (Petersen 1877).

A young German insider engaged in an overview of the methods available in clinical research, particularly in therapeutics. Friedrich Martius (b.1850), while a military doctor and later an assistant physician at the Berlin Charité University Clinics, published two lengthy articles on the subject: *Die Principien der wissenschaftlichen Forschung in der Therapie* (The principles of scientific research in therapy, 21 pages, Martius 1878) and the even more erudite *Die numerische Methode (Statistik und Wahrscheinlichkeitsrechnung) mit besonderer Berücksichtigung ihrer Anwendung auf die Medicin* (The numerical method [statistics and calculus of probabilities] with special reference to its application in medicine, 41 pages, Martius 1881). Later he became professor of internal medicine at Rostock and, typically, did not publish any longer on the subject.

As Oesterlen and Schweig had done some thirty years previously, Martius begun by clarifying the confused terminology. He analysed the French and German works (Radicke was not mentioned!) by putting them in the wider historical context of the theories of cognition: the term 'induction', he wrote, was often used without distinguishing whether 'logical', 'numerical' or 'experimental' induction was meant. Numerical induction he understood as being based on statistics and probability calculus. For, although they had been developed separately, statistics and the calculus of probabilities could be summarized under the common term "numerical method". Indeed, they complemented each other

> *necessarily and happily [...]. The calculus of probabilities needs for its application materials collected according the strict rules of statistics and the latter, without the calculus, would not [...] always find its critical utilization and the elaboration of which it is capable* (Transl. from Martius 1881, p 349).

Consequently, aware of the Paris deliberations of 1835, "*which have since acquired fame*", Martius regretted that Gavarret, in his enthusiasm, had disparaged statistics in favour of the calculus of probabilities (Martius 1878, p 1185; Martius 1881, p 243). Gavarret was following Laplace, who had declared that all knowledge was based upon probability (Martius 1881, p 347-348). This was obviously not true. One had just to think of anatomy. Of course, one had to be familiar with basic mathematical principles to be able to discuss the appropriateness of conclusions arrived at by the "numerical method", for it was often used simply to prove what one wanted to prove (Martius 1881, p 338-339). But doctors' continuing aversion to the mathematical approach stemmed from their "*mathematical incapability*" (Transl. from Martius 1881, p 346).

As to the fundamentals on which the true "numerical method" rested, Martius identified some open questions. He wrote, unfoundedly as we have seen, that Gavarret's famous probability ratio of 212/213 (99.5%) had been chosen arbitrarily on the basis of Poisson's formulas, as had Hirschberg's simplification by fixing a ratio of 9/10. This showed the arbitrariness and unreliability of such ratios and of the whole process: which of these haphazardly proposed probabilities excluded the hazard? Here was Martius's answer:

> *To remedy this undeniable drawback, Liebermeister now intends – by dropping Poisson's formulas completely and, departing from other preconditions – to develop new formulas, that can serve to calculate, with certitude and precision, the degree of probability with which the hazard is excluded. And this for any [...] observational material, be it ever so small [...provided] the*

comparability of the cases, this eternal crux of all statistical data collections can be demonstrated (Transl. from Martius 1881, p 375-376).

If Liebermeister's formulas were more easily applicable, they were also "more unscientific" than Poisson's, for

> *...they completely neglect the law of large numbers, and they offer nothing but the reflection, expressed in numbers of probability, that when one ignores the nature of the process in course the best thing to do is,* faute de mieux [my italics]*, to stick to true, existing successes* (Trans. from Martius 1881, p 376).

Like Claude Bernard (see above) had done, Martius made clear that progress in identifying constant, determined causal relations required induction through laboratory experiments, not the numerical method. This did not mean that he proposed neglecting statistics. On the contrary, through mass observation and reliable assessment of treatment successes, the probability of obtaining important indications for practical action increased (Martius 1881). It was not, as Gavarret had deemed, the ripest fruit of modern thought, or

> *the highest and most consummate level of all research methods usable in therapeutics. Rather it is and remains a makeshift, albeit a very important one [...], that is undoubtedly worth an even deeper foundation and more extensive application* (Transl. from Martius 1878, p 1185).

These were clever insights, and such efforts would effectively be made in the 20th century. But before, new difficulties, and consequently, new desiderata were recognized by two practitioners from Breslau (now Wrocklaw, Poland), **Alfred Ephraim** and **Ottomar Rosenbach**.

When reading these two testimonies one realises that formal, mathematical probabilistic reasoning had clearly made an impact on its authors. But on the practical side the consequences were limited, whilst on the theoretical new requirements for scientific evaluation were identified – for the new century.

Towards the *fin-de-siècle*

Unsystematically compiled statistics continued to be worked up and interpreted in the manner of shopkeepers, and without additionally calculating probabilities despite what Martius had called for. So, **Alfred Ephraim** (b.1863) felt once more – just as Wunderlich had fifty years earlier – that therapeutics were chaotic. In his *Über die Bedeutung der statistischen Methode für die Medicin* (On the significance of the statistical method for medicine, Ephraim 1893) he saw the reason for this desolate state in the oblivion of the provisions stipulated time and again since the Paris discussions of years ago. The methodology of clinical evaluation was eclipsed by new technical methods of examination. Ephraim noted that a recent discussion between two eminent German physicians made clear that the numerical method continued to have both detractors and supporters. He claimed that the *reversal* of previously statistically founded claims did not help to convince the medical world of the value of such work (Ephraim 1893, pp 695-696). That is why Ephraim answered the two eternal questions of the statistical endeavour – (i) what was to be counted? and (ii) how many cases should be counted – by recalling the precepts established by Gavarret (p 712). But while these theoretical difficulties could be dealt with, one should not overlook the practical ones; and here he enumerated three new criteria for solid comparisons:

- diagnoses must have been made using the same diagnostic methods, which is particularly difficult when cases are assembled from various sources.
- adherence to treatments must be strictly observed:
- trials of treatment should be conducted over sufficient duration.

The untreated cases were necessary for comparison, but difficult to find. If lack of treatment seemed inhuman, it could be justified because most treatments had actually not been demonstrated to be useful.

"*Non-adherence [to these precepts] was being seen every day and lead to delusive therapeutic-statistical conclusions*" (p 715), but they were as difficult to fulfil as they were indispensable.

Ephraim concluded that those who deemed these requirements insurmountable must be aware that they are renouncing trustworthy therapeutic knowledge. He noted that "*to substantiate the efficacy of mercury in syphilis, of quinine in malaria…, one might perhaps not need statistics*" (p 711). Yet, reliable identification of less dramatic treatment effects could only be assured by the results of statistical research. However, he did not mention the calculation of probabilities as a complementary method of evaluation, thus once more ignoring terminological precision.

Ottomar Rosenbach (b.1851) had worked since 1874 as a hospital physician at Breslau. By1896 he had resigned his position as chief of the medical department and his associate-professorship and retired to private practice in Berlin, but continued publishing. It is probable that he knew young Ephraim since they had lived in Breslau at the same time. Certainly, he knew the latter's methodological work for he extended it in two publications, a lengthy ten-page one in three parts on *Serumtherapie und Statistik* (Serum therapy and statistics, Rosenbach 1896) and a shorter paper on *Der Kampf um die Zahl in der medicinischen Wissenschaft* (The fight about numbers in medical science, Rosenbach 1899). Both were conspicuously published in the *Münchener Medicinische Wochenschrift* (The Munich Medical Weekly).

Like Ephraim, Martius and many others before them, Rosenbach criticized the misuse of therapeutic statistics:

> *Although everybody now knows that small numbers prove absolutely nothing, although everyone knows […Poisson's] law of large numbers, yet people preferentially use small numbers, and even many of those who with aplomb only exploit large numbers are in error about their bearing in that it is not the large numbers as such that matter, but the circumstances [over time] in which they are generated (*Transl. from Rosenbach 1896, p 913).

And he further emphasized, the arbitrarily defined, often inadequate duration of trials, the failure to use modern diagnostic criteria (for example, the use of clinical signs *and* bacteriology), and the differences among cases (an issue that had already been proclaimed innumerable times).

As a consequence, statisticians' concentration on the *Genesungsquotient* (recovery rate) was misleading, since both numerator (number of cures) and denominator (number of diseased) were often based on variable and subjective criteria. In short, these statistics served only to reinforce preconceived opinions and frequently, when there was no comparison group, to fall into the trap of *the post hoc-ergo-propter hoc* fallacy (pp 912-913).

As new elements, he drew attention to bias mechanisms, namely:

- The selection of cases by enthusiasts who, with what they refer to cynically as *'scientific thoroughness', eliminate all unsuitable cases so that, under the new method, deaths must, in reality, no longer occur. (That they still happen is, by the way, … always the fault of unhappy external circumstances, never imputable to the procedure …. Or it is the impossibility of using the panacea sufficiently promptly).*
- The historical insight that *this procedure of unevenly distributing light* [on successful cases] *and shadow* [on failures] *has repeated itself in the history of medicine countless times, and it never loses its impression on credulous minds who do not want, or are unable to understand that highly astounding results can be brought about by the simple 'slight-of-hand' (legerdemain) of a new scientific definition"* (pp 912-913).
- A historical comparison was only admissible if the forms of an epidemic remained essentially unchanged over the years. In the case of diphtheria, for example, where Behring's serum-therapy had been introduced since 1893, he demanded "*that one should try once again to obtain a large series of observations [of patients treated] without serum-therapy*" over many years (Transl. from Rosenbach 1899, p 256).

Of course, many people did not understand or like Rosenbach's method-based objections. And maybe they did not like him: had he not, while still an aspiring *Privatdozent*, written quite aggressively in a book on the *Foundations, Duties and Limits of Therapeutics* (Grundlagen, Aufgaben und Grenzen der Therapie, Rosenbach 1891):

Statistics – what would they not have sanctioned in the hands of able arrangers. (And later:) *The history of medicine furnishes enough examples of friends and foes fighting with equal obstinacy and equal certainty for a dogma established on the basis of such contradictory* [statistical] *results* (Transl. from Rosenbach 1891, pp 66,183).

Even when in private practice, Rosenbach indefatigably continued responding to his detractors, writing critically in the *Zeitschrift für klinische Medicin* (Journal of Clinical Medicine) on methodological problems, right up to his death in 1907. In his last paper on this issue – *Die Diagnose als ätiologischer Factor* (Diagnosis as aetiological factor, Rosenbach 1905) – Rosenbach adduced yet another new criterion for a valuable experiment: the *method of alternation.* Returning to the serum-therapy of diphtheria, he repeated the need for the *experimentum crucis (the decisive experiment),* namely,

> *always to treat one case with and the next without the promoted medicine, whether the medicine is tested in all places at the same time, or in different places one after the other. And of course, this holds not only for the treatment of diphtheria* (Transl. from Rosenbach 1905, p 233).

 If one did not want to, or could not perform this process of evaluation, which Rosenbach felt was easy to carry out, one deprived oneself straightaway of the possibility of doing scientific research (Rosenbach 1905; Chalmers et al. 2011).

Germany by 1900

According to Petersen and many of his contemporaries (and later historians), the overall situation in the 19th century may have been similar in practice to that in France and Britain: if considered at all, the actuarial method of counting and statistical analysis prevailed. In particular, however, Louis's numerical method as applied to the evaluation of therapies was also constructively criticized right from the beginning, and appropriate efforts to ameliorate it were made. In the end, selection bias, making results positive by changing criteria, and quantification of preconceived ideas were decried as misuses (Rosenbach 1896). The notion of probability was understood by many clinicians, and a few of them actually struggled to apply formal mathematical probability and its consequences throughout the whole second half of the 19th century. The influence of Poisson and particularly of his medical pupil, Gavarret, was pivotal.

Conclusions and perspectives for the new century

200 years of discussion
Modes of probabilistic thinking

Arguably the first outflow of probabilistic thinking in medicine was evaluation of the inoculation of smallpox in 18th century England. Numerical comparisons of death rates of inoculated and uninoculated groups were made by mathematically- inclined clinicians such as James Jurin. From then, probability became a problem of numbers, of quantification (Daston 1995, p 4). This implied probabilism in that proportions of average mortalities of groups were calculated and *compared.* The probabilistic reasoning behind these approaches was unconscious (Mode I in Table 1). Neither was it made explicit by most 18th / early 19th century British "arithmetic observationsts", nor by Louis in Paris and those who followed his *méthode numérique* from the 1830s. Quantification remained informal: simple counting, summation and calculating averages, rates, proportions and frequencies. In other words, it was pre-mathematical in the strict sense of the word. This mode of practice was widespread. It prevailed with different intensities throughout the two centuries covered in my research.

Another mode spoke of probability explicitly, but in practice still used informal, pre-mathematical quantification, i.e. without a calculus of probabilities (Mode II in Table 1). I found scattered examples of this authored by clinicians from 1772 onwards, and increasingly after 1800.

It was in the 1770s, and in the issue of inoculation of smallpox in particular, that reasoning became mathematically probabilistic, and therefore conscious (Mode III in Table 1). This emerged in a violent debate in the Paris *Académie des Sciences* between a French mathematician who had contradicted a Swiss colleague. This mode of thinking was present among subsequent generations of French mathematicians, until 1840, when it became practical again with the young French mathematician-clinician Jules Gavarret. It was received wholeheartedly in Germany during the 1860s and 1870s, as manifested in the overlooked mathematical contributions of at least half a dozen young German physicians (Fick, Jürgensen, Jessen, Hirschberg, Liebermeister, Martius). Mode III was had definitely been launched, but on a small scale.

Context

A list of those who unconsciously propagated probability by fostering informal, pre-mathematical numerical evaluation of therapy (Mode I in Table 1) could easily be compiled. It would be endless (see for example Tröhler 2000; Tröhler 2010). It became common practice from the second half of the 19th century onwards. I have selected authors for their motives, insights, arguments – and/or flaws.

In the same way I identified authors who consciously propagated pre-mathematical probabilistic aspects (Mode II), or even evolved formal, mathematical probability (Mode III) in clinical medicine. Both groups referred to Gavarret from the 1840s onwards (Henle 1844; Griesinger 1848; Wunderlich 1851; Oesterlen 1852; Schweig 1854; Guy 1860; Fick 1866; Jürgensen 1866; Jessen 1867; Hirschberg 1874; Liebermeister 1877; Petersen 1877; Martius 1881; Ephraim 1893).

After extensive research, this list seems to me fairly exhaustive. It suggests the relative rarity of Mode III probabilistic thinking in clinical medicine.

Typically for their profession, the prominent French and German physiologists Claude Bernard (Paris) and Karl von Vierordt (Tübingen), rejected probability in favour of the certainty of determinism, that is, a constant relation between a cause and its effect (Matthews 1995, p 15). In other words, they were looking for laws of nature.

In this context, Bernard was counted both among the supporters and the detractors as he approved consciously of numerical probabilistic comparative evaluation (Mode II in Table 1) in therapeutics, but not for physiological phenomena.

18th century British arithmetic observationism and its younger French sister, the "numerical method", became quite outspoken in their requests for statistics. Both preached the need for straightforward quantification – for "statistics" – albeit not for the abstract calculus of probabilities. Such quantifications fitted neatly into the contemporary statistical movements that became so active in Europe and North America (Porter 1986, p 396). They can be seen as a reaction against the arbitrary exercise of authoritarian personal powers characteristic of the *Ancien Régime.* In a more democratic society, trustable action would be called for, and, for many, numbers seemed trustworthy. The fight against superstition, fixed ideas, prejudices, and newly the church, also played a role.

Besides this general societal trend there were certainly individual psychological stimuli. I can only speculate about these. Rather, let me enquire about the intellectual motives behind the phenomenon I have observed. To fathom this quantifying "probabilistic turn", it is helpful to consider the arguments invoked by its proponents and those who hampered its acceptance.

Incentives

In clinical medicine, thinking became probabilistic when new interventions and therapies were invented. Enlightened doctors wanted to compare them with older ones to find which one was to be preferred: was inoculation valuable in preventing smallpox compared to leaving the disease to take its course? Was Peruvian bark or bloodletting the better therapy of "fevers"? In other words what were the risks of medical innovation? This led also to questioning the value of long-established standard therapies *per se*, like bloodletting, when

checked against the natural course of disease, as reported, for instance, by Hamilton (1816) ([Milne and Chalmers 2014](#); [Tröhler 2000](#); Tröhler 2006).

The response lay in collecting cases and quantifying the harvest, "to improve the evidence of medicine", as the saying went in the 18th century, or to "raise medicine to the level of other sciences", as it was repeated throughout the 19th century.

Another motive came from young men bluntly recognising, time and again, that therapeutics was chaotic. Polypharmacy reigned, fashions and 'systems' came and went; prescriptions were built on unrecorded experience, arguments, and reasoning; but those were mere speculations! The truth lay in observed facts, and in many of them, assembled in groups for calculation. Fostering this philosophy motivated some clinicians. In turn it meant probabilistic thinking would be involved, as we now know, in counting, comparing and, in the end, mathematical analysis.

Obstacles

Objectivation and scientification brought new problems, however. Was probability not afflicted with errors, whereas old 'Certainty' was – well – certain? This I name the 'certainty-versus-probability' problem. Another important issue was uncertainties about the appropriateness of results and risks calculated from groups for evaluating a medical measure, for they might not apply to an individual patient: I may call it the 'group-versus-single patient/case problem'. It had been identified and addressed in the 1760s in the Paris inoculation debate, later, for instance, in the Paris disputes in the 1830s, and then by Henle (1846), Griesinger (1848), and Trousseau (1862). And the conundrum remains and seems likely to remain a bone of contention.

And then, questions also arose concerning the moral level: probability was 'less good than the truth' (how then were both defined?); or, did composition of groups and calculations not facilitate arrangements according to one's preferences, beliefs or vested interests? This I call the 'easy-to-cheat problem'. It was pointed out by British arithmetic observationists in the 18th century ([Tröhler 2000](#)) and throughout the 19th century by Laplace (1995, cf. above p 16), Oesterlen (1852, p 135), Trousseau (1865, p XLIII), Hirschberg (1874, p IX), Rosenbach (1891, pp 66, 183).

In mid-1830s Paris (and later by Henle) the necessity of "good faith" when working with statistics was stressed (Murphy 1981, p 315; Henle 1844, p 17). On the other hand, numerical work was also decried on moral grounds as inhumane because it stubbornly adhered to a research protocol instead of a true treatment plan (Wunderlich 1841). Furthermore, there was the '*post-hoc-ergo-propter-hoc*-fallacy': it had already been identified in the 17th century ([Matthews 2020a](#)) and was evoked time and again during 19th century, a British reminder being the 1891 Harveian Oration given by James Andrew (b.1829) ([Andrew 1891](#)). The issue seems eternal to me.

For their criticism of quantification, denigrators of probabilistic reasoning – such as Le Rond d'Alembert in the 18[th]century, Risueño d'Amador in the first half of the 19th century, and Trousseau in the second half of the 19th century – in one combination or another, all used the 'art-versus-science argument', as well as the 'certainty-versus-probability' issue, the 'group-versus-single patient/case', the 'post-hoc-ergo-propter hoc' and the 'easy-cheating' arguments.

Another controversial issue concerned the essential interpretation of data. Since it could imply value judgements, inferences and generalizations could be considered correct or injudicious. It is true that the meaning of average differences between comparison groups was eventually quantified by 'sophisticated' statistical significance tests, with the probability of a difference being judged using the now widely applied concept of confidence intervals. But these tests were too complicated to be used routinely; and anyway, in the end they might only raise false hopes of "moral certitude" ([Matthews 2020b](#)).

There were still other contentions: statistical work could be seen methodologically, as misuse, that is, as quantification of preconceived ideas, or falling into the trap of the '*post hoc-ergo propter hoc* fallacy', or both. And there was the phenomenon of apparently contradictory results of, say, two or more successive clinical

trials. This was referred to as 'medical reversal' and it was repeatedly mentioned (Richerand 1825; Lancet 1834/35; Martius 1881; Rosenbach 1891; Ephraim 1893). This was a misapprehension, because these 19th century clinical authors did not take into account Gavarret's demonstration that there is uncertainty associated with every estimate of a difference. Nowadays the term 'medical reversal' (Prasad and Cifu 2011; 2019) has sometimes been used when a medical intervention had been introduced enthusiastically but without adequate evidence, and abandoned when better evidence revealed that not only did it not help, but might even harm, patients.

Then, the still ongoing confusion of the value of a method as such with the difficulties of its application and/or its potential for misuse, was also identified in the 19th century.

Another line of argument concerned the question of the risks of medical innovations. As I have shown at the beginning of this essay, the onset of probabilism in the 18th century was triggered by new measures (variolation) and therapies, for instance in surgery (Tröhler 1987, 2000, 2006). While traditional treatments such as bleeding and purging were just there, unquestioned from time immemorial, innovations met not only with approval or repudiation, but also with scepticism and uncertainty. A specific strategy for dealing with uncertainty was the new notion of risk. It was based on the calculation of probabilities. As medical historian Thomas Schlich states (quoting sociologist Renée Fox):

> Probability-based logic has been employed "to approach the uncertainties of diagnosis, therapy, and prognosis, and in the clinical judgement that lie at their heart" since the eighteenth century (Schlich 2006, p 1).

The term 'risk', derived from the French "risqué", seems to have appeared in medicine in its anglicised form only in the early 19th century. So, while uncertainty was felt, its handling as 'a risk', was still only unconsciously probabilistic. It produced a new kind of knowledge in order to reduce uncertainty (possibly to certainty…), namely numerical data

Finally, there were psychological impediments, well recognised since the 18th century, when d'Alembert and Haygarth had acknowledged the influence of human feelings and intuitions. And there was a fact that we know with hindsight: the human brain does not recognise probability; *prima vista* it is neither perceptible, discernible, nor evident; it must either be believed or calculated; and calculations are barriers.

On top of these intellectual and psychological difficulties there were continuing practical obstacles: the elaboration of statistics was a cumbersome and time-consuming enterprise. In fact, the prerequisites for meaningful statistical comparisons increased over time. There was the number of cases theoretically deemed necessary, their comparability, and difficulties of concurrent comparisons. However, those who were convinced of the need to use probabilistic thinking underestimated these practical difficulties and thereby marginalised themselves, whilst deeming clinicians to be mathematically incapable (Martius). Researchers did not even apply mathematical aids "prepared ready for mechanical use" (Fick 1866, p 430).

It is important to realise that all these problems and impediments were Janus-headed: They were challenges on the one side and reasons for criticism on the other. Clinical medicine as 'Science' – implying numbers and probabilism – was the perspective of the progressively minded; traditionalists saw clinical medicine as an "Art".

In the 19th century, however, the preoccupations developed in a new direction. Students became thrilled by the discoveries of the new disease conception, Virchow's cellular pathology, anaesthesia and its sudden consequence, modern surgery, diagnostic innovations (such as the stethoscope and ophthalmoscope, laboratory methods, and radiology). They could not be bothered with complicated epistemic issues. Yet they complained that therapeutic chaos could no longer be ignored; it should be vanquished by exactly the methods maligned by those who simply muddled through.

But who were they, these propagators, doubters, critics and opponents of methods involving probabilistic thinking? To answer this question, it is helpful to consider their social status within their respective communities.

Social, national, and long-term perspectives
Who was concerned about probabilism?

The initial 18th century theoretical advocates of formal probabilistic thinking (Mode III in Table 1) in medicine were Swiss and French – the Bernoullis, Condorcet, Laplace, and Poisson. They had high standing as established professors and/or scientists. But as mathematicians they were only marginally interested in practical real-world issues, and published their results in learned books, journals and societies. They saw themselves, or they were regarded from the point of view of clinicians, as "strangers at the bedside", and doctors probably hardly took any notice of them. Daniel Bernoulli was in fact a first medically qualified adopter. He was followed two generations later by Condorcet, Pinel, and Poisson (in words and formulae); and finally, after a further generation, by the pivotal young physician-mathematician Jules Gavarret (in practice). Yet, Gavarret's 1840 book remained his only contribution to the field. He was to become prominent in the Paris medical world as a physiologist and was at one time President of the *Académie Royale de Médecine*.

In Britain, more than a century before Gavarret, James Jurin and others – for instance, his young associate, the Swiss Johann Kaspar Scheuchzer (b.1702) – initiated pre-mathematical quantified evaluation (Mode I) in practice. Jurin, the secretary, and Scheuchzer, a fellow of the Royal Society, were both part of the British scientific establishment. Both held Cambridge MDs, Jurin after having also read mathematics. They were learned physicians.

British arithmetic observationists such as Lind, Gregory, Haygarth, Black, Millar (a particularly militant author), McGrigor and many others emulated them. They acted upon Modes I and II probabilistic thinking. They were practical clinicians. I have characterized them as 'marginal men' in that they were typically from provincial origin, not Oxbridge graduates, but Edinburgh-trained Scots, naval or army doctors, dispensary practitioners, and dissenters, decried as "democrats and levellers", who challenged authoritarian traditionalism, unqualified opinions, prejudices, and they fought for transparency (Tröhler 2000; 2014).

In the 19th century such features of a 'marginal man' continue to apply e.g. for Alcock, Todd and Hodgkin, *a fortiori* as they were all born in the heydays of arithmetic observationism and also made no academic career (in Hodgkin's case, this was despite remarkable anatomo-clinical research such as a cancer of the lymph nodes still known as Hodgkin's disease). It fits also for Blane and Balfour, for after active service, they rose to high posts in naval and army medical administration, respectively.

Young Francis Bisset Hawkins and William Guy, however, were Oxford and Cambridge graduates, respectively. Yet they both abandoned medical practice early in their lives and became distinguished London figures, FRCPs (Croonian, Gulstonian and Lumleian Lecturer, and Harveian Orator in the case of Guy) and FRSs (vice-president), acknowledged for their commissioned public work and active in public health statistics. Clinicians may therefore also have seen them as "strangers at the bedside".

In Paris, Louis set out to propagate unconsciously informal probabilistic thinking (Mode I) with his *méthode numérique.* He taught pathology at two Parisian Hospitals; yet in a way he was a loner, who occupied no important posts. His efforts had waned by the end of the 1840s, after his son had died. He was less influential with local than with foreign students, for instance with Swiss, German, British and US-American (Lancet 1834-35; Müllener 1967; Hannaway 2007).

In the second half of the 19th century, after Gavarret, formal probabilistic evaluation became further fostered and sophisticated mathematically up to tests of statistical significance; authors were chiefly German. Yet Schweig (1854), a physician turned civil servant, and Radicke (1858), a physicist, were hardly mentioned in the medical literature. Fick (1866) was already a professor of physiology and may have had some temporary impact. The others were young clinicians with a mathematical bent like their 18th century forerunners had been

(Griesinger 1848; Oesterlen 1852; Jürgensen 1866; Jessen 1867; Hirschberg 1874; Liebermeister 1877; Martius 1881; Ephraim 1893; Rosenbach 1896; 1905). Interestingly, six of these nine clinicians were of Jewish origin . This religious peculiarity was a feature of 19th century German probabilists, just as non-conformism was a feature among18[th] century British arithmetic observationists. Whether anti-semitism was a hampering reason is an open question that could only be addressed using detailed individual biographical studies (Hammerstein 1995; Weber 2003).

Two paths were open for these ambitious men: (i) either they became professors and chiefs of university departments and then, indirectly at most, continued their probabilistic work (Griesinger, Jürgensen, Martius; Liebermeister was the exception); or (ii) they were academically unsuccessful, abandoned that career, and became "marginal men" in private practice, with the consequence that their work, lacking institutional authority, was ignored (Jessen, Oesterlen, Hirschberg, Ephraim, Rosenbach). The small medical faculty of Tübingen was a kind of centre of interest in probabilistic thinking over three decades (Griesinger 1848; Oesterlen 1852, Wunderlich 1853, Liebermeister 1861,1877, Jürgensen as of 1873).

I reckon that the (temporally limited) marginal social positions of most clinicians when publishing on epistemological questions, thereby thinking probabilistically and actively trying to foster it, was fairly typical of the 18th century British and the 19th century French, British, and German authors I have studied. Antagonists on the contrary were well established members of the academic community, interested in maintaining the *status quo* and their personal prestige.

Were there national differences?

The question now arises whether there were nevertheless national differences in the emergence, reception and dissemination of probabilistic thinking. Does the evidence suggest different national models of emergence of a science of therapeutic evaluation, and were there differences in the communication of ideas? I am inclined to answer "yes" to both.

In *France* the issues of evaluation and of risks were first treated theoretically by scientists interested in probability, who saw this notion as applicable to the real world of clinical medicine. This developed in Paris, first among colleagues, over four generations in a *master-to-pupil-chain*. Probability stayed mathematical even when Gavarret finally meant to apply it formally in 1840 – in a practically unusable mode. This state-of-the-art remained for the rest of the century.

The French ignored the pragmatic mode of *British* arithmetical observationism, as well as the later German thinking. A remarkable phenomenon was the openmindedness to Louis' work among young foreign students. German and English translations of Louis' relevant publications appeared quickly. As results from Table 2, this had already been the case for German editions of the 18th /19th century British works by Lind, Gregory, Haygarth, Black and Blane, for example. The French themselves seemed not to be eager to learn either from abroad, or even from the locally generated novelty of Louis's *méthode numérique.* Had they been interested in these matters they would have taken notice of the prior and concurrent British pre-mathematical probabilistic evaluations and quantified nosography, of which I have not so far found any translations into French (see Table 2).

British authors unconsciously propagated probabilistic thinking in multi-centred *networks over time* (Tröhler 2000). When some realized by the mid-1830s that their type of informal probabilistic thinking was also being practised in Paris, it was favourably but also critically reviewed. Guy then combined it with Gavarret's formal French achievements and thereby made it practically applicable, although he realized that it was not ideal. This pragmatism can be seen as a genuine British tradition (Pickstone 2000). Guy's 1860 state-of-the-art seems to have remained the British standard for the rest of the century.

True, there was some interest in epistemology after the Sydenham Society's publication of an English translation of Oesterlen's *Medical logic* (Oesterlen 1855) and of Radicke's paper on the value of arithmetical means (Radicke 1861), but they remained the only translated German works on probabilistic issues I have been

able to find. Two lectures by Jürgensen and Liebermeister on cold water used for fever therapy were also translated and published by the Sydenham Society in 1877 (Liebermeister and Jürgensen 1877). Yet, although they included some of their statistics, the study design and the methods underlying them, let alone Liebermeister's work on formal probabilistic issues and his 'four-table-test (described above), were neither mentioned nor translated. Trousseau's diatribe against the numerical method continued to be hidden in the English edition of his long-winded *Leçons cliniques* (Trousseau 1868). Clearly, the (editor's) interest was on clinical issues.

When German authors became involved in the 1840s, they could – and actually did – draw from French and British sources. Local *contemporaneous networks* established themselves in Tübingen, Kiel and Breslau through *personal collegiality.* They tackled some pending problems in applying formal probabilistic techniques to clinical practice. Over time, a solution was presented, deemed too complicated, a new answer was proposed, and so on, in a dynamic, indefatigable manner.

Thus far, I do not know for sure whether any of those German solutions were taken into account in Britain or France, but I doubt it. Authors with an international outlook were rare among those I have studied. As Table 2 suggests, if they did not know the languages, they were unable to keep abreast of the reported developments in probabilistic thinking. Maybe also, they simply did not care about this specific topic.

The value of a long-term perspective

When surveying historical evidence covering 200 years one expects today to identify what has changed. And indeed, I have mentioned many changes between 1700 and 1900. On the other hand, there were also remarkably constant features.

During the two centuries, and in all three of the countries studied, there was increasing awareness among clinicians of a need to publish dependable information about medical achievements. The simple fact that evolving probabilistic modes of thinking and clinical action had encroached on minds over 200 years makes a difference, particularly in the long run. The meaning of probability had changed. A new kind of knowledge was being generated, and this new situation created new problems. As there were more and more innovations, the epistemic issues seem likely to have concerned more people.

However, the typology of those tackling these issues remained the same. The ways the questions were considered or not, why and by whom, suggest ten enduring features.

First, it reveals that most French, British and German clinicians throughout the two centuries considered were not aware of the underlying probabilistic nature of their thinking and action when counting and analysing their cases using Mode I quantification.

Second, it is clear that relatively few of them consciously mentioned probability, according to Mode II, let alone to apply the formal mathematical techniques of Mode III to estimate the value of a medical measure.

Third, it turns out that, with few exceptions, the probabilistically orientated clinical authors whom I have studied were young when publishing references to this topic.

Fourth, I have noted a consistently forward-looking, future-oriented attitude among these authors, explicitly expressed, for example, by Todd (1831); Hodgkin (1834/54); Griesinger (1848); Wunderlich (1851); Schweig (1854); Guy (1860); Jessen (1867); Petersen (1877); and Rosenbach (1891).

Fifth, most of the clinicians were (still) in marginal social positions when publishing on mathematically orientated probabilistic evaluation. For ever one who later achieved a recognized academic position, another dropped the manifested interest – for various reasons. Some may have resigned themselves to the conservative influence of the established clinical community.

Sixth, others remained marginal men – again for various reasons. At any rate, neither the field of epistemology nor the clinicians who tackled it ever received academic recognition for this endeavour.

Seventh, by ignoring or underestimating the practical difficulties in propagating their ideas among clinicians, authors marginalized themselves and the calculus of probabilities intellectually over the 200 years covered in this study.

Eighth, it became clear that fundamental problems with the evaluation of therapies were unlikely to be solved to everyone's satisfaction. Indeed, this is a manifestation of discussions about the right "method of conducting medical enquiries", a debate that had existed for nearly 2000 years (Matthews 2020a). As the mathematician-historian Robert Matthews states, one continues to ask today – as my 18th and 19th century authors did – whether medical practice is best guided by the rationalists' approach informed by their understanding of fundamental mechanisms of disease and treatment or, by contrast, by the empiricists' argument that reliable knowledge comes from simply observing large numbers of cases (Matthews 2020a). Is therapy based on old-fashioned rational certainty, experimentally ascertained laws determined by nature, or on modern probable results of empirical observations, or, differently expressed, whether medicine is a 'Science' or a (healing) 'Art' – both requiring definitions? Questions and answers remain.

Ninth, the long-term perspective makes clear that the arguments advanced for or against probabilistic solutions persisted over time, and they were used in whatever way supported the analyses and perspectives, in other words the interpretations of the disputatious parties. In fact, interpreting the results of calculations is an essential step in the circle of probabilistic reasoning [see graph]

Tenth, the uphill battle fought by the supporters of probabilistic thinking engaged in research suggests that only very rarely before 1900 was clinical practice guided by conscious probabilistic reasoning pursued through informal quantification, let alone through formal methods.

The above ten persistent features may help to explain why, on the whole, clinical medicine had never really adopted even informal probabilistic thinking by the end of the 19th century, let alone the rigorous conditions of statistical testing as required by the calculus. Certainly, the active, published opposition, and, above all, the passive daily resistance or simple lack of interest in teaching and practice, were major contributory reasons.

Above all there is a timeless constancy, a reason which also emerges from the present historical research: clinical research and practice "have a mind of their own" as clinical epidemiologist Jan Vanderbroucke has suggested (Vanderbroucke 1998). In other words, it is the insight that medicine is neither theoretically dogmatic-rational, nor empirically knowledgeable. It is both – a *rational-empiric unity*. Perhaps by accident, but typical anyway, it was a British clinician who arrived at such a pragmatic conclusion. The 18th century **William Cullen** (b.1710), a leading light of Edinburgh medicine and a physician of European reputation who wrote between 1768 and 1789 in his *Practice of Physic*:

> *…for two thousand years past there have been* [these] *two plans proposed…and…it is extremely necessary to know that both have their imperfections, … and that, in the present state of science, either of them is by itself unsufficient* [sic!] (Quoted from Vandenbroucke, p 14).

My study has now extended the timespan over another 150 years converging to a century prior to our present days.

Outlook for the new century

At the end of the 19th century, doctoring continued to be seen as an Art requiring "tact" and intuition in the treatment of an individual patient. This could be learned by long-term experience. It was also understood to be the application of results provided by Science – a quantifying science producing and evaluating average data that were useful for everyone, albeit requiring specialist knowledge. In that restricted sense clinical practice was probabilistic. Both aspects represented dogmas and had flaws: authoritatively proclaimed certainty was more easily believed by practitioners (and patients) than calculated probable estimates; but the scientifically minded considered such certainty a phantom.

The second half of the 19th century, had witnessed a rapid rise in operative surgery thanks to pathological anatomy, anaesthesia and anti-sepsis and asepsis. New therapeutic possibilities had been introduced on a much larger scale than during the 18th century. This increasingly opened up important therapeutic options. Surgeons used simple methods to show the success of their innovative ventures, mainly uncontrolled case series of operated cases or by using inadequate historical controls (Tröhler 1984, pp 97-120; 2014). There were also advances in non-operative medical disciplines. The era was optimistic. Optimism overshadowed epistemic questions, and probabilistic reasoning "did not rule the world".

It always required exceptional insight (and courage, and/or stubbornness, sometimes idealism, and candour) to publicly express that therapeutics was in a chaotic state, as some clinicians had already felt it to be since the second half of the 18th century. This conclusion was characteristic motive for scientifying evaluation procedures which had emerged since. Two Germans, Ephraim and Rosenbach, stated this again towards the end of the 19th century. They formulated a host of indispensable conditions, old and new, to improve the relevance of quantitative evaluations of therapies: comparison to an untreated or differently treated group of patients in the same conditions; using uniformity of diagnostic methods; documented adherence to treatment; trials extended over sufficient time; the concepts of placebo and blinding. (In fact, the latter had been in the minds of some people since the 18th century (Booth 2002; Donaldson 2016a; Stolberg 2006; Kaptchuk 1998; Jütte 2013). Some saw these as possibilities to potentially improve the approach to impartial statistics aiming at objective probabilities: a science for the future. This idea was conceptualised early in the 19th century in France (Laplace 1810th), later in England (Todd 1831; Hodgkin 1834/54; Guy 1860), and finally and thoroughly, in Germany (Griesinger 1848; Wunderlich 1851; Schweig 1854, p 349; Jürgensen 1866; Jessen 1867; Petersen 1877; Martius 1881; Rosenbach 1891).

And it came. With hindsight, the 18th and 19th centuries were the long dawn of this science of probabilistic testing. By 1900, the time seemed ripe for it, and it started soon after in Britain with statisticians – Karl Pearson (b.1857), William S. Gosset (b.1876; alias "Student"), Ronald A. Fisher (b.1890), Major Greenwood (b.1880), and Austin Bradford Hill (b.1897), who developed hypothesis testing with resulting p-values, or with the confidence intervals favoured by Jerzy Newman (b.1894). They had all forgotten their 19th century forerunners. During the second half of the 20th century, evaluation science developed into a recognized discipline, widely accepted by doctors, whether researchers or practitioners, and young idealists from a variety of backgrounds, for example, those who initiated the Cochrane Collaboration.

In every generation, the question bridging the gap, namely "what are the reasons for not adopting the majority result in a particular case?" was formulated in 1839 (Guy) and revived in the 1890s. What about the answer? Would it eventually be found by further mathematization and/or by adoption of Bayes' theorem?

Propagators of a quantifying probabilistic approach did not foresee the emergence of difficulties in quantifying medicine (Matthews 2020a; Sheynin 1978, p 285): when one intricacy was recognized and a solution proposed, another one emerged, like Hydra's new heads. Science never comes to an end – except when it is mathematised.

Are we happier now – with tests of statistical significance being accused of having deleterious consequences (Amrhein et al. 2019; Matthews 2020b)? How should one deal confidently with uncertainty when the replication of trials addressing the same question yield apparently incompatible results? Will the relevance of mathematically sophisticated, probabilistic evaluation decrease in view of 'personalized healthcare', with which newest genetics dazzles us (Senn 2004)?

History goes on, and one thing is sure: people are inevitably constrained by the times they live in, but they must always strive for the best available solutions. But what are these? The discussions about probabilities that started three centuries ago have continued, and they may well continue on an even more complex level. Who knows whether these developments will have an impact on clinical practice? As Lavoisier realized 300 years ago,

the integration of a probabilistic approach, albeit essential in trustworthy interpretation of the results of experiments, observations and calculations, has proved particularly difficult "above all in (clinical) medicine".

Acknowledgements

Editorial remarks

Except when otherwise mentioned, translations into English are my own. When not specifically referenced, biographical details stem from:

- Bynum WF, Bynum H eds. (2007). Dictionary of medical biography. Westport Conn., London: Greenwood.
- DSB, Dictionary of Scientific Biography. New York: Charles Scribner's Sons.
- Hirsch A ed (1929). Biographisches Lexikon der hervorragenden Ärzte aller Zeiten und Völker, 2nd reviewed and completed edition. Berlin, Wien: Urban & Schwarzenberg.

References

Agnew RAL (2008). John Forbes FRS (1787-1861). JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/john-forbes-frs-1787-1861).

Alcock T (1823). An essay of the education and duties of the general practitioner in medicine and surgery containing suggestions relating to the investigation of disease, and the registration of practical results. Transactions of the associated Apothecaries and Surgeon-Apothecaries of England and Wales 1: 75-135.

Alcock T (1830). Lectures on practical and medical surgery...London: Burgess and Hill.

Amrhein V et al (2019). Retire statistical significance. Nature 567: 305-307.

Andrew J (1891). The Harveian oration. London: Allard.

Bailey R, Howick J (2018). Did John Stuart Mill influence the design of controlled clinical trials? JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/john-stuart-mill-influence-design-controlled-clinical-trials/).

Baker RB, McCullough L (2009). Discourses of philosophical medical ethics. In: Baker RB, McCullough L eds. The Cambridge world history of medical ethics. New York, NY, Cambridge University Press, 284-290.

Balfour TG (1854). Quoted in West C. Lectures on the Diseases of Infancy and Childhood. London: Longman, Brown, Green and Longmans.

Barclay AW (1864). Medical errors: fallacies connected with the application of the inductive method of reasoning to the science of medicine. London: Churchill.

Barclay AW (1881). The Harveian oration. Royal College of Physicians. London: Harrison.

Bariéty M (1972). Louis et la méthode numérique. Clio medica 7: 177-183.

Bartlett E (1844). An essay on the philosophy of medical science. Philadelphia: Lea and Blanchard.

Bernard C (1865). Introduction à l'étude de la médicine expérimentale. Paris: Flammarion.

Bisset Hawkins F (1829). Elements of Medical Statistics. London: Longman et al.

Bird A (2017). Systematicity, knowledge, and bias. How systematicity made clinical medicine a science. Synthese. DOI 10.1007/s11229-017-1342-y.

Bird A (2018). James Jurin and the avoidance of bias in collecting and assessing evidence on the effects of variolation. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/james-jurin-and-the-avoidance-of-bias-in-collecting-and-assessing-evidence-on-the-effects-of-variolation/).

Bishop D, Gill E (2019). Robert Boyle on the importance of reporting and replicating experiments. JLL Bulletin: Commentaries on the history of treatment evaluation. (https://jameslindlibrary.org/articles/robert-boyle-on-the-importance-of-reporting-and-replicating-experiments/).

Blane G (1819). Elements of medical logick, or philosophical principles of the practice of physick, 3rd. edition (1823). London, Underwood.

Black W (1789). An arithmetical and medical analysis of the diseases and mortality of the human species. London: Dilly. (2nd ed. reprinted. Farnborough, Hants: Gregg 1973.

Booth CC† (2002). John Haygarth FRS (1740-1827). JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/john-haygarth-frs-1740-1827/).

Bothwell LE, Greene JA, Podolsky SH, Jones DS (2016). Assessing the gold standard – lessons from the history of RCTs. New England Journal of Medicine 374: 2175-2181.

Boylston AW (2008a). Zabdiel Boylston (1679/80-1766). JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/zabdiel-boylston-167980-1766/).

Boylston AW (2008b). William Watson's use of controlled clinical experiments in 1768. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/william-watsons-use-of-controlled-clinical-experiments-in-1767/).

Boylston AW (2010). Thomas Nettleton and the dawn of quantitative assessments of the effects of medical interventions. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/thomas-nettleton-and-the-dawn-of-quantitative-assessments-of-the-effects-of-medical-interventions/).

Boylston AW (2012). The origins of inoculation. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/the-origins-of-inoculation/).

Boylston AW, Williams AE (2008). Zabdiel Boylston's evaluation of inoculation against smallpox. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/zabdiel-boylstons-evaluation-of-inoculation-against-smallpox/).

British and Foreign Medical Review (1841) 6 (2): 1-21.

Cabanis PJG (1798). Du degré de certitude de la médecine. Paris: Didot.

Canguilhem, G (1970–1980). Cabanis, Pierre-Jean-Georges. In: Dictionary of Scientific Biography. New York: Charles Scribner's Sons, vol 3: 1–3.

Chalmers I, Dukan E, Podolsky SH, Davey Smith G (2011). The advent of fair treatment allocation schedules in clinical trials during the 19th and early 20th centuries. JLL Bulletin: Commentaries on the history of treatment

evaluation (https://www.jameslindlibrary.org/articles/the-advent-of-fair-treatment-allocation-schedules-in-clinical-trials-during-the-19th-and-early-20th-centuries/).

Chalmers J, Chalmers I, Tröhler U (2017). Helping physicians to keep abreast of the medical literature: Andrew Duncan and Medical and Philosophical Commentaries, 1773-1795. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/helping-physicians-keep-abreast-medical-literature-andrew-duncan-medical-philosophical-commentaries-1773-1795/).

Clifton F (1732). The state of physick, ancient and modern, briefly considered: with a plan for the improvement of it. London: Nourse.

Coleman W (1990). Experimental physiology and statistical inference: The therapeutic trial in nineteenth century Germany. In: Krüger L, Gigerenzer G, and Morgan MS (eds). The probabilistic revolution. Cambridge MA–London: MIT Press, vol 2, pp 201-226.

Daston L (1995). Classical probability in the enlightenment. Princeton NJ: Princeton University Press.

Daviel J (1753): Sur une nouvelle méthode de guérir la cataracte par l'extraction du cristalin. Paris: Mémoires de l' Académie Royale de Chirurgie 2: 337–354.

Dean ME (2004). The trials of homeopathy. Essen: KVC Verlag.

Dean ME (2009). Comparative evaluation of homeopathy and allopathy within the Parisian hospital system, 1849-51. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/comparative-evaluation-of-homeopathy-and-allopathy-within-the-parisian-hospital-system-1849-51/)

de la Condamine CM (1754). Mémoire sur l'inoculation de la petite vérole. Lu a l'assemblé Condamine e publique de l'Académie royale des sciences, le mercredi 24 avril 1754. Paris: Durand.

de la Harpe P, Gabriel J-P (2010). Daniel Bernoulli, pionnier de modèles mathématiques en médecine. Images des mathématiques. CNRS, Janvier 12, n.p.).

de Laplace PS (1820). Théorie analytique des probabilités, 3rd ed. In: Oeuvres complètes 7. Paris: Gauthier-Villars 1886.

de Laplace, PS (1995). Philosophical essay on probabilities. New York: Springer.

de Lavoisier de AL (1865). Oeuvres. Paris: Imprimerie Impériale, 3 (mémoires).

Dickersin K, Chalmers I (2010). Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organisation. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/recognising-investigating-and-dealing-with-incomplete-and-biased-reporting-of-clinical-research-from-francis-bacon-to-the-world-health-organisation/)

Donaldson IML (2016a). Antoine de Lavoisier's role in designing a single-blind trial to assess whether 'Animal Magnetism' exists. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/antoine-de-lavoisiers-role-in-designing-a-single-blind-trial-to-assess-whether-animal-magnetism-exists/).

Donaldson IML (2016b). George Starkey's 1658 challenge to Galenists to compare their treatment results with his. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/george-starkeys-1658-challenge-to-galenists-to-compare-their-treatment-results-with-his/).

Ephraim A (1893). Über die Bedeutung der statistischen Methode für die Medicin. [Volkmann's Sammlung Klinischer Vorträge N.F. Innere Medicin 24: pp 695 -716. Leipzig: Breitkopf und Härtel.

Edinburgh Medical Journal (1869). Review.14: 842-846.

Faure JF (1759). Receuil des pièces qui ont concouru pour le prix de l'Académie Royale de Chirurgie. Paris: Le Prieur, 8: 489-520.

Fick A (1866). Anwendung der Wahrscheinlichkeitsrechnung auf medicinische Statistik. In: The same. Die medicinische Physik, 2nd ed, Anhang, pp 416-439.

Fordyce G (1793). An attempt to improve the evidence of medicine. Transactions of a Society for the Improvement of Medical and Chirurgical Knowledge. London: Johnson.

Franklin B (1759). Some account on the success of inoculation for the small-pox in England and America. London: W Strahan.

Gavarret LDJ (1840). Principes généraux de statistique médicale: ou développement des régles qui doivent présider à son emploi. Paris: Bechet jeune & Labé.

Gavarret J (1844). Allgemeine Grundsätze der medicinischen Statistik, oder Entwicklung der für die numerische Methode gültigen Regeln. Erlangen: Enke.

Gillies D (2000). Philosophical theories of probability. London, New York: Routledge.

Gigerenzer G, Swijtink Z, Porter T, Daston L, Beatty J, Kruger L (1989). The empire of chance. How probability changed science and everyday life. Cambridge: Cambridge University Press.

Gigerenzer G (2002). Reckoning with risk. Learning to live with uncertainty. London: Allen Lane, Penguin.

Gregory J (1772/1805). Lectures on the duties and qualifications of a physician (new ed.). Edinburgh: Creech.

Griesinger W (1848). Zur Revision der heutigen Arzneimittellehre. Archiv für Physiologische Heilkunde 7: 1-24.

Guy WA (1839). On the value of the numerical method as applied to science but especially to physiology and medicine. Journal of the (Royal) Statistical Society 2: 25-47.

Guy W (1860). Croonian Lectures. The numerical method, and its application to the science and art of medicine. British Medical Journal, Part I, (i)331-334; Part IV (i) 467-469; Part VI (ii) 593-597.

Hacking I (1975, repr.2006). The emergence of probability. Cambridge: Cambridge University Press.

Hald A (1998). A history of mathematical statistics from 1750 to 1930. New York: Wiley-Interscience.

Hammerstein N (1995). Antisemitismus und die deutschen Universitäten. Frankfurt/M: Campus.

Hamilton L (1816). De synocho castrensi. Edinburg: Ballantyne.

Hannaway C (2007). Louis, Pierre-Charles -Alexandre. In: Bynum WF and Bynum H (eds). Dictionary of medical biography. Westport, London: Greenwood, vol 3, 814-815.

Hawkins FB (1829). The elements of medical statistics. London: Longman, Rees, Orme, Brown & Green.

Haygarth J (1784). An inquiry into how to prevent smallpox. Chester and London: Monk & Johnson.

Haygarth J (1793). Sketch of a plan to exterminate the casual smallpox from Great Britain; and to introduce general inoculation. London: J Johnson.

Haygarth J (1805). A clinical history of diseases 1. Acute rheumatism. 2. Nodosity of the joints. London: Cadell and Davies.

Henle J (1844). Medicinische Wissenschaft und Empirie. Zeitschrift für rationelle Medicin 1:1-35.

Henle J (1846). Handbuch der rationellen Pathologie. Braunschweig, Vieweg, Vol. 1.

Hirschberg J (1874). Die mathematischen Grundlagen der medizinischen Statistik. Leipzig: Veit.

Hodgkin T (1854). Numerical method of conducting medical inquiries. Association Medical Journal, New series 2: 1090-1094.

Holmes OW (1861). Currents and countercurrents in medical science with other addresses and essays. Boston: Ticknor and Fields, 1861.

Howick J (2016). Aulus Cornelius Celsus and 'empirical' and 'dogmatic' medicine JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/aulus-cornelius-celsus-and-empirical-and-dogmatic-medicine/).

Huber T (1959). Daniel Bernoulli, Physiologie und Statistik. Stuttgart, Schwabe.

Huth EJ (2005). Quantitative evidence for judgments on the efficacy of inoculation for the prevention of smallpox: England and New England in the 1700s. JLL Bulletin: Commentaries on the history of treatment evaluation (http://www.jameslindlibrary.org/articles/quantitative-evidence-for-judgments-on-the-efficacy-of-inoculation-for-the-prevention-of-smallpox-england-and-new-england-in-the-1700s/).

Huth EJ (2006). Jules Gavarret's Principes Généraux de Statistique Médicale: a pioneering text on the statistical analysis of the results of treatments. JLL Bulletin: Commentaries on the history of treatment evaluation (http://www.jameslindlibrary.org/articles/jules-gavarrets-principes-generaux-de-statistique-medicale-a-pioneering-text-on-the-statistical-analysis-of-the-results-of-treatments/).

Ineichen R (1994). Der „Viererfelddtest" von Carl Liebermeister (Bemerkudengen zur Entwicklung der medizinischen Statistik im 19.Jahrhundert). Historia Mathematica 21: 28-38.

Jessen W (1867). Zur analytischen Statistik. Zeitschrift für Biologie 3:128-136.

Jorland G, Opinel A and Weisz G eds (2005). Body counts. Medical quantification in historical and sociological perspectives. Montreal, Kingston, London, Ithaca, McGill-Queen's University Press.

Jurin J (1724). A letter to the learned Dr. Caleb Cotesworth, F.R.S. of the College of Physicians, London, and physician to St. Thomas's Hospital; containing a comparison between the danger of the natural small-pox, and that given by inoculation. Philosophical Transactions of the Royal Society of London, 32:213-227.

Jürgensen T (1866). Klinische Studien über die Behandlung des Abdominaltyphus mittelst des kalten Wassers. Leipzig, Vogel.

Justman S (2017). James Lind and the disclosure of failure. Journal of the Royal College of Physicians Edinburgh 47: 384–387.

Jütte R (2013). The early history of placebo. Complementary Therapies in Medicine 21: 94-97.

Kaptchuk TJ (1998). Intentional ignorance: a history of blind assessment and placebo controls in medicine. Bulletin of the History of Medicine 72: 389-433.

La Berge AF (2005). Medical statistics at the Paris school. What was at stake? In: Jorland G, Opinel A and Weisz G (eds). Body counts. Medical quantification in historical and sociological perspectives. Montreal, Kingston, London, Ithaca, McGill-Queen's University Press, pp 89-108.

Lancet (1834-35). Essay Review of state of therapeutics. Vol 2: 81-87.

Lancet (1834-35). Essay review of PCA Louis. Pathological researches of phthisis. Introduction, notes, additions and an essy by Charles Cowan. Vol 2: 292-97.

Lancet 1836-37, vol 1:413-14.

Liebermeister, von C (1861). Bemerkungen über die Anwendung der Mathematik auf die physikalischen Wissenschaften. Journal für praktische Chemie 84: 416-419.

Liebermeister, von C (1877). Ueber Wahrscheinlichkeitsrechnung in Anwendung auf therapeutische Statistik. In: Volkmann R (ed). Sammlung klinischer Vorträge. Lepizig: Breitkopf und Härtel, 110, Innere Medicin, No. 39, pp 935-962.

Liebermeister C, Jürgensen T (1877). Treatment of fever; Principles of treatment of croupous pneumonia. London: The New Sydenham Society 71: pp 275-300; 314-348.Lilienfeld DE (1978). ”The greening of epidemiology”: Sanitary physicians and the London Epidemiological Society (1830-1870). Bulletin of the History of Medicine 52:503-528.

Lind J (1772). A treatise of the scurvy, 3rd ed. London: Crowder.

Louis PCA (1839). Anatomical, pathological and therapeutic researches on the yellow fever of Gibraltar of 1828 from observations by himself and M. Trousseau. Boston: Little and Brown.

Magnello E, Hardy A eds (2002). The road to medical statistics. Amsterdam-New York: Rodopi.

Marks HM (2005). When the state quotes lives: Eighteenth century quarrels over inoculation. In: Jorland G et al eds (2005), pp 51-64.

Martius F (1878). Die Principien der wissenschaftlichen Forschung in der Therapie. In: Volkmann R (ed). Sammlung klinischer Vorträge. Leipzig: Breitkopf und Härtel, 139, Innere Medicin No 47, 1169-1188.

Martius F (1881). Die numerische Methode (Statistik und Wahrscheinlichkeitsrechnung) mit besonderer Berücksichtigung ihrer Anwendung auf die Medicin. Virchow's Archiv für pathologische Anatomie…83:336-377.

Matthews JR (1995). Quantification and the quest for medical certainty. Princeton NY: Princeton University Press.

Matthews R (2017). Chancing it. The laws of chance and how they can work for you. London: Profile Books.

Matthews RAJ (2020a). The origins of the treatment of uncertainty in clinical medicine. Part 1: Ancient roots, familiar disputes. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/the-origins-of-the-treatment-of-uncertainty-in-clinical-medicine-part-1-ancient-roots-familiar-disputes/).

Matthews RAJ (2020b). The origins of the treatment of uncertainty in clinical medicine. Part 2: the emergence of probability theory and its limitations. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/the-origins-of-the-treatment-of-uncertainty-in-clinical-medicine-part-2-the-emergence-of-probability-theory-and-its-limitations/).

Miller G (1957). The adoption of inoculation for smallpox in England and France. Philadelphia: The University of Pennsylvania Press.

Milne I, Chalmers I (2014). Alexander Lesassier Hamilton's 1816 report of a controlled trial of bloodletting. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/alexander-lesassier-hamiltons-1816-report-of-a-controlled-trial-of-bloodletting/)

Morabia A (2007). Claude Bernard, statistics, and comparative trials. JLL Bulletin: Commentaries on the history of treatment evaluation (http://www.jameslindlibrary.org/articles/claude-bernard-statistics-and-comparative-trials/).

Müllener ER (1967). Pierre-Charles-Alexandre Louis' Genfer Schüler und die «Méthode numérique». Gesnerus 24:46-76.

Murphy TD (1981). Medical knowledge and statistical methods in early nineteenth century France. Medical History 25: 301-319.

Nettleton T (1722). Part of a letter from Dr. Nettleton, physician at Halifax, to Dr. Jurin, R. S. Sec concerning the inoculation of the small-pox, and the mortality of that distemper in the natural way. Philosophical Transactions of the Royal Society of London 32:209-212.

Niebyl PC (1977). The English bloodletting revolution, or modern medicine before 185o. Bulletin of the History of Medicine 51: 464-483.

Oesterlen F (1852). Medicinische Logik. Tübingen: Laupp.

Oesterlen F (1855). Medical logic. London: New Sydenham Society.

Opinel A, Gachelin G (2010). French 19th century contributions to the development of treatments for diphtheria. JLL Bulletin: Commentaries on the history of treatment evaluation (http://www.jameslindlibrary.org/articles/french-19th-century-contributions-to-the-development-of-treatments-for-diphtheria/).

Paré A (1575). 1575 Les oeuvres de M. Ambroise Paré conseiller, et premier chirurgien du Roy avec les figures & portraicts tant de l'Anatomie que des instruments de Chirurgie, & de plusieurs Monstres. Paris: Gabriel Buon.

Petersen J (1877). Hauptmomente in der geschichtlichen Entwickung der medicischen Therapie. Kopenhagen: Höst.

Pickstone JV (2000). Ways of Knowing: A New History of Science, Technology and Medicine. Manchester: Manchester University Press.

Pinel P (1809). Traité médico-philosophique de l'aliénation mentale, 2nd ed. Paris:

Poisson SD (1837). Recherches sur la probabilité des jugements en matière criminelle et en matière civile,

précédées des règles générales du calcul des probabilités. Paris, Bachelier.

Poisson, Dulong, Larrey and Double (1835). Recherches de statistique sur l'affection calculeuse, par M. le Docteur Civiale. Comptes rendus hébdomadaires des séances de l'Académie de Sciences [Statistical research on (urinary) stone, by M. Dr Civiale]. Paris: Bachelier, 167-177. (English translation reprinted 2001: Statistical research on conditions caused by calculi by Doctor Civiale. International Journal of Epidemiology 30:1246-1249).

Porter T (1986). The rise of statistical thinking 1820-1900. Princeton NY, Princeton University Press.

Porter T (1995). Trust in Numbers. The Pursuit of Objectivity in Science and Public Life. Princeton NJ: Princeton University Press.

Porter TM (2005). Medical quantification: Science, regulation and the state. In: Jorland et al (2005) p 394-401.

Prasad V, Cifu A (2011). Medical reversal: Why we must raise the bar before adopting new technologies. Yale J Biol Med. 84: 471–478.

Prasad V, Cifu A (2019). Ending medical reversal. Baltimore: Johns Hopkins University Press.

Radicke G (1858). Die Bedeutung und Werth arithmetischer Mittel mit besonderer Beziehung. Archiv für physiologische Heilkunde NF 2: 145-219.

Radicke G (1861). On the importance and value of arithmetic means. London: New Sydenham Society.

Richerand, de AB (1825). Histoire des progrès récens de la chirurgie. Paris: Béchet le Jeune.

Rothschuh KE (1968). Friedrich Oesterlen (1812-1877) und die Methodologie der Medizin. Sudhoffs Archiv 52: 97-129.

Rosenbach O (1891). Grundlagen, Aufgaben und Grenzen der Therapie. Wien-Leipzig: Urban und Schwarzenberg.

Rosenbach O (1896). Serumtherapie und Statistik. Münchener Medicinische Wochenschrift 43: 911-915,

Rosenbach O (1899). Der Kampf um die Zahl in der medicinischen Wissenschaft. Münchener Medicinische Wochenschrift 46:256-257.

Rosenbach O (1905). Die Diagnose als ätiologischer Factor. Zeitschrift für klinische Medicin 56:219-240.

Ruffieux C (2020). Les méthodes d'évaluation de nouveaux remèdes dans la première moitié du 19ème siècle: l'exemple des médecins genevois. Gesnerus 77: in press.

Rusnock A (2002). Vital accounts. Quantifying health and population in Eighteenh-century England and France. Cambridge: Cambridge University Press.

Scheuchzer JG (1729). An account of the success of inoculating the small-pox in Great Britain, for the years 1727 and 1728: With a comparison between the mortality of the natural small-pox, and the miscarriages in that practice; as also some general remarks on its progress. London: Peele.

Schlich T (2006). Risk and medical innovation. A historical perspective. In: Schlich T, Tröhler U eds. (2006): The risks of medical innovation. Risk perception and assessment in historical context. London-New York: Routledge, pp 1-19.

Schlich T, Tröhler U eds (2006). The risks of medical innovation. Risk perception and assessment in historical context. London- New York: Routledge.

Schweig G (1854). Auseinandersetzung der statistischen Methode in besonderem Hinblick auf das medicinische Bedürfniss. Archiv für physiologische Heilkunde 13: 305-355.

Seneta E, Seif F, Liebermeister H, Dietz K (2004). Carl Liebermeister (1833-1901): a pioneer of the investigation and treatment of fever and the developer of a statistical test. Journal of Medical Biography 12:215-221.

Senn S (2004). Individual response to treatment: is it a valid assumption? British Medical Journal 329:966-68.

Sheynin OB (1976). P.S. Laplace's work on probability. Archive for History of Exact Sciences 16: 137-187.

Sheynin OB (1978). S.D. Poissons's work in probability. Archive for History of Exact Sciences 18: 245-300.

Sheynin OB (1982). On the history of medical statistics. Archive for History of Exact Sciences 26: 241-286.

Stigler S:M (1986). The history of statistics. Cambridge–London: Belknap Press of Harvard University Press, pp 62-70.

Stolberg M (2006). Inventing the randomized double-blind trial: The Nürnberg salt test of 1835. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/inventing-the-randomized-double-blind-trial-the-nurnberg-salt-test-of-1835/)

Todd TJ (1831). The book of analysis or a new method of experience whereby the induction of the Novum Organon [of Bacon] is made easy of application to medicine, physiology...London: Murray.

Tröhler U (1984). Der Nobelpreisträger Theodor Kocher 1841-1917. Basel-Boston-Stuttgart: Birkhäuser.

Tröhler U (1987). Die Gewissheit der Chirurgie: Grundlagen klinisch-therapeutischer Bewertung um 1750. Praxis (Schweiz. Rundschau Medizin) 76: 958-961.

Tröhler U (2000). The 18th century British origins of a critical approach. Edinburg: Royal College of Physicians. (This book is freely available here)

Tröhler U (2001). Commentary: 'Medical art' versus ,medical science': J. Civiale's statistical research on conditions caused by calculi at the Paris Academy of Sciences in 1835. International Journal of Epidemiology 30: 1152-1253.

Tröhler U (2003a). Edward Alanson 1782: responsibility in surgical innovation. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/edward-alanson-1782-responsibility-in-surgical-innovation/)

Tröhler U (2003b). James Lind and the evaluation of clinical practice. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/james-lind-and-the-evaluation-of-clinical-practice/).

Tröhler U (2003c). James Lind at Haslar Hospital 1758-1774: a methodological theorist. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/james-lind-at-haslar-hospital-1758-1774-a-methodological-theorist/).

Tröhler U (2005). Quantifying experience and beating biases: A new culture in Eighteenth-century British Clinical medicine. In: Jorland G et al eds. (2005) pp 19-50.

Tröhler U (2006). To assess and to improve: Practitioners' approaches to doubts linked with medical innovations 1720-1920. In: Schlich T, Tröhler U eds (2006) pp 20-37.

Tröhler U (2007). An early 18th century proposal for improving medicine by tabulating and analysing practice. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/an-early-18th-century-proposal-for-improving-medicine-by-tabulating-and-analysing-practice/)

Tröhler U (2010). The introduction of numerical methods to assess the effects of medical interventions during the 18th century: a brief history. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/the-introduction-of-numerical-methods-to-assess-the-effects-of-medical-interventions-during-the-18th-century-a-brief-history/).

Tröhler U (2013). William Cheselden's 1740 presentation of data on age-specific mortality after lithotomy. JLL Bulletin: Commentaries on the history of treatment evaluation. (https://www.jameslindlibrary.org/articles/william-cheseldens-1740-presentation-of-data-on-age-specific-mortality-after-lithotomy/).

Tröhler U (2014). Statistics and the British controversy about the effects of Joseph Lister's system of antisepsis for surgery, 1867-1890. Republished in the Journal of the Royal Society of Medicine 2015;108:280-287. JLL Bulletin: Commentaries on the history of treatment evaluation (https://www.jameslindlibrary.org/articles/statistics-and-the-british-controversy-about-the-effects-of-joseph-listers-system-of-antisepsis-for-surgery-1867-1890/).

Trousseau A (1862) Conférences sur l'empirisme faites à la Faculté de médecine de Paris…Paris: Delahaye.

Trousseau A (1865). Clinique médicale de l'Hôtel-Dieu, 2e éd. Paris: Baillière.

Trousseau A (1868). Lectures on clinical medicine delivered at the Hôtel Dieu, Paris. London: The New Sydenham Society, XLII

Vandenbroucke JP (1998). Clinical investigation in the 20th century: the ascendancy of numerical reasoning. The Lancet 352: suppl. 12-16.

Verchère F (1908). Discussion sur le traitement opératoire du cancer. In: 2e Congrès de la Société Internationale de Chirurgie. Bruxelles: Hayez, Vol. I, 557-560.

Warner JH (1997). The therapeutic perspective: Medical practice, knowledge, and identity in America, 1820-1885. Princeton NJ: Princeton University Press.

Warner JH (2003). Against the spirit of system: The French impulse in nineteenth-century American medicine. Baltimore and London: Johns Hopkins Univ. Press.

Watson W (1768). An account of a series of experiments, instituted with a view of ascertaining the most successful method of inoculating the smallpox. London: Nourse.

Weber T (2003). Anti-semitism and philo-semitism among the British and German elite in Oxford and Heidelberg before the first World War. English Historical Review 118:86-119.

Weiner DB (2007). Pinel, Philippe. In: Bynum WF and Bynum H (eds). Dictionary of Medical Biography. Westport, Conn. – London, Greenwood Press 4; 1008-1013).

Weisz G (1993). Academic debate and therapeutic reasoning in mid-19th century France. In: Löwy I et al (ed). Medicine and change: Historical and sociological studies of medical innovation. Paris: INSERM – J Libbey Eurotext, 287-315.

Weisz G (1995). The medical mandarins. NewYork-Oxford: Oxford University Press.

Wiesing U (1995). Kunst oder Wissenschaft. Konzeptionen der Medizin in der deutschen Romantik. Stuttgart-Bad Cannstadt: Fromann-Holzboog.

Wunderlich CRA (1841). Wien und Paris: ein Beitrag zur Geschichte und Beurtheilung der gegenwärtigen Heilkunde in Deutschland und Frankreich. Stuttgart: Ebner & Seubert; repr. 2012, USA, Nabu Press.

Wunderlich CRA (1851). Ein Plan zur festeren Begründung der therapeutischen Erfahrung. CTA Schmidt's Jahrbücher der in-und ausländischen gesammten Medicin. 41:106-111.