The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2020, Vol. 113(12) 497–503
DOI: 10.1177/0141076820968486

Theory and clinical use of probabilities in Germany after Gavarret. Part I: introducing German dramatis personae

Ulrich Tröhler
Institute of Social and Preventive Medicine, University of Bern, CH-3012 Bern, Switzerland
Corresponding author: Ulrich Tröhler. Email: ulrich.troehler@ispm.unibe.ch

This article continues our series on probabilistic thinking and the evaluation of therapies, 1700–1900
The series will continue to appear as separate articles in the forthcoming issues of JRSM

Books by Lind, Gregory, Haygarth and Black quoted earlier in this series 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).[1] This had changed by the mid-1830s and 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 large 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 a 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).

Henle's warning of the danger of falling into the *post-hoc-ergo-propter-hoc* fallacy by basing one's practice on successful single cases was also 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, Henle 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 summarised 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 deduce 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 their 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 worldwide 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 about 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 example 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 for 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 realised the classic confusion of a method as such with its incorrect execution, both scientifically and ethically. He now also recognised 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 criticised 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 criticised 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 of 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), 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 other than 'mere conceit' (pp. 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 (pp. 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 for the individual case. Much remained to be done (p. 22)!

Although approved of, Louis was also criticised 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 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).[2] Oesterlen's book dealt with medical observation, the concepts of induction, deduction, generalisation, 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 generalisations (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.[3] 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 reforming psychiatric care. 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 50-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) (pp. 307–309). He had clearly read Jacob Bernouilli, Poisson, and Gavarret (pp. 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,[4] 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.[4] This was deemed to be 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. 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.[4]

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 criticised. 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 – became a professor of internal medicine at Tübingen – and 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 and 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 hospitalisation, 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 the 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 at Kiel, with Jürgensen, 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 used (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.[5] 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'. It was practicable for analysing very small groups ($n = <10$), but led to extensive calculations when larger groups were analysed. He included examples of both situations.[6,7]

Declarations

Competing interests: None declared.

Funding: None declared.

Ethics approval: Not applicable.

Guarantor: UT.

Contributorship: Sole authorship.

Acknowledgements: My heartfelt thanks to: Iain Chalmers, without whose unflinching encouragement, gentle whip, intellectual and unrenounceable practical help over the years, I would neither have begun nor ever terminated this work; Thomas Schlich, who critically and helpfully read all previous versions; Robert Matthews, whose help with mathematical matters was very welcome; Brigitte Wanner and Christian Wyniger of the Institute of Social and Preventive Medicine, Bern, who helped me, together with Patricia Atkinson, Oxford, with ever so many IT technicalities; my wife, Marie Claude, whose patient love is not probably, but absolutely true.

Provenance: Invited contribution from the James Lind Library.

Supplementary file: The references listed below are chosen as essential to the reading of the article. However, the full list of primary and secondary references is available online both on the

Journal's website as supplementary material, and with the original publication at https://www.jameslindlibrary.org/articles/probabilistic-thinking-and-the-evaluation-of-therapies-1700-1900/. Except when otherwise mentioned, translations into English are the author's own

Author's note: When not specifically referenced, biographical details stem from:

- Bynum WF and Bynum H, eds. *Dictionary of Medical Biography*. Westport, CT and London: Greenwood, 2007.
- *Dictionary of Scientific Biography*. New York: Charles Scribner's Sons.
- Hirsch A, ed. *Biographisches Lexikon der hervorragenden Ärzte aller Zeiten und Völker*, 2nd ed. Berlin, Wien: Urban & Schwarzenberg, 1929.

References

1. Wiesing U. *Kunst oder Wissenschaft. Konzeptionen der Medizin in der deutschen Romantik*. Stuttgart-Bad Cannstadt: Fromann-Holzboog, 1995.
2. Bailey R and Howick J. Did John Stuart Mill influence the design of controlled clinical trials? *JLL Bulletin: Commentaries on the history of treatment evaluation*. See www.jameslindlibrary.org/articles/john-stuart-mill-influence-design-controlled-clinical-trials/ (last checked 7 October 2020).
3. Rothschuh KE. Friedrich Oesterlen (1812–1877) und die Methodologie der Medizin. *Sudhoffs Archiv* 1968; 52: 97–129.
4. Coleman W. Experimental physiology and statistical inference: the therapeutic trial in nineteenth century Germany. In: Krüger L, Gigerenzer G and Morgan MS (eds) *The Probabilistic Revolution*. Vol. 2, Cambridge, MA/London: MIT Press, 1990, pp.201–226.
5. Matthews RAJ. *Chancing It. The Laws of Chance and How They Can Work for You*. London: Profile Books, 2017.
6. Ineichen R. Der Viererfelddtest "von Carl Liebermeister (Bemerkudengen zur Entwicklung der medizinischen Statistik im 19.Jahrhundert). *Hist Math* 1994; 21: 28–38.
7. Seneta E, Seif F, Liebermeister H and Dietz K. Carl Liebermeister (1833-1901): a pioneer of the investigation and treatment of fever and the developer of a statistical test. *J Med Biograph* 2004; 12: 215–221.