The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2021, Vol. 114(3) 132–139
DOI: 10.1177/0141076821992926

Conclusions and perspectives, part II: social, national, and long-term perspectives

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

This article concludes our series on probabilistic thinking and the evaluation of therapies, 1700–1900.

Who was interested in 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 transactions of scientific societies. They saw themselves, or were regarded by clinicians, as 'strangers at the bedside', and doctors probably took very little notice of them. Daniel Bernoulli was in fact a first medically qualified adopter of probabilistic thinking. 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 book of 1844 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.

More than a century before Gavarret James Jurin and others in Britain – for instance, his young associate, the Swiss Johann Kaspar Scheuchzer (b.1702) – initiated pre-mathematical quantified evaluation (Mode I in Table 1) 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, and were emulated by British arithmetic observationists, such as Lind, Gregory, Haygarth, Black, Millar (a particularly militant author), McGrigor, Blane and many others. They were practical clinicians who used Modes I and II in probabilistic thinking (see Table 1). I have characterised them as 'marginal men' in that they were typically of provincial origin, not 'Oxbridge' graduates, but Edinburgh-trained Scots, naval or army doctors, dispensary practitioners and dissenters. They were decried by those in the medical establishment as 'democrats and levellers' who challenged authoritarian traditionalism, unqualified opinions, prejudices, and who fought for transparency.[1,2]

In the 19th century, such features of 'marginal men' continued to apply, as exemplified by Alcock, Todd and Hodgkin, all of whom had been born in the heyday of arithmetic observationism and had not pursued academic careers (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). On leaving active military service, Blane, McGrigor and Balfour rose to high posts in naval and army medical administration, respectively. Blane and McGrigor ended up with a knighthood.

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. They, too, may therefore have been seen by clinicians as 'strangers at the bedside'.

In Paris, with his *méthode numérique,* Louis set out to propagate (unconscious) informal probabilistic thinking (Mode I in Table 1). He taught pathology at two Parisian hospitals, but he was a loner who occupied no important posts. By the end of the 1840s, after his son had died, Louis' efforts had waned. He was less influential with local French students than he was with foreign students from Switzerland, Germany, Britain and America (Lancet 1834–1835).[3,4]

Table 1. Three modes of probabilistic reasoning in clinical medicine.

		Some representatives	
Unconscious implicit numerical, but not mentioning probability	Mode I Informal Pre-mathematical	Typical representatives	Other representatives dealt with in the text
		Jurin et al. from 1720s Lind 1772 Black 1789 Blane 1819 Bisset Hawkins 1829 Louis et al. from 1835/1837	Faure 1747 McGrigor from 1815 Alcock 1823 Todd 1835 Cowan 1835 Balfour 1854
Conscious explicit numerical and mentioning probability	Mode II Informal Pre-mathematical	Gregory 1772 Pinel 1807 Guy 1839 Henle 1844 Schweig 1854	Lavoisier 1780s Condorcet 1785 Hodgkin 1834/1854 Griesinger 1848 Oesterlen 1852 Wunderlich 1851 Trousseau 1865 Rosenbach 1896
	Mode III Formal Mathematical Evoking/elaborating calculus of probabilities for clinical needs	*Theoretical* J Bernoulli 1713 D'Alembert 1760 Laplace 1814 D'Amador 1837 Guy1841 Guy 1860 Petersen 1877 Martius 1881 Ephraim 1893	*Practical* D Bernoulli 1760 Haygarth 1784 Poisson 1837 Gavarret 1840 Radicke 1858 Fick 1866 Jürgensen 1866 Jessen 1867 Hirschberg 1874 Liebermeister 1877

In the second half of the 19th century, after Gavarret's formal probabilistic evaluation had become mathematically sophisticated and had led to tests of statistical significance, the key authors were chiefly German. 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, and 'marginal' for that reason. This religious identity was a feature of 19th century German probabilists, just as Christian non-conformism had been a feature of 18th-century British arithmetic observationists. Whether anti-semitism was a reason for marginalisation is an open question that would require detailed individual biographical studies.[5,6]

Two paths were open for these ambitious men: either (i) 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 path, and became 'marginal men' in private practice. Lacking institutional authority, their work was ignored (Jessen, Oesterlen, Hirschberg, Ephraim, Rosenbach). For three decades, the small medical faculty of the University of Tübingen became centre of interest in probabilistic thinking (Griesinger, 1848; Oesterlen, 1852, Wunderlich, 1853, Liebermeister, 1861, 1877, Jürgensen as of 1873).

The marginal social positions of most clinicians publishing on epistemological questions, and therewith thinking and fostering probabilistic thinking, was fairly typical of the 18th-century British and 19th-century French, British and German authors I have studied. Opponents of this 'unorthodox' way of thinking, in contrast, were well-established members

of the academic community who wished to maintain the status quo and their personal prestige.

Were there national differences?

The question now arises whether there were 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 these questions.

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 way of thinking developed in Paris, first over four generations in a *master-to-pupil-chain* of mathematicians. Probability remained mathematical even when Gavarret finally tried to apply it in clinical practice formally in 1840 – albeit in a mode that turned out to be practically unusable. This 'state-of-the-art' in France endured for the rest of the 19th century.

The French ignored the pragmatic mode of British arithmetical observationism, as well as the later German thinking. What is remarkable is that open-mindedness to Louis' work existed among young foreign students. German and English translations of Louis' relevant publications appeared promptly. As shown in 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.[1] By the mid-1830s, when some realised that their type of informal probabilistic thinking was also being practised in Paris, it was favourably but also critically reviewed. Thomas Guy then combined it with Gavarret's formal French achievements and thereby made it applicable in practice, although he realised that it was not ideal. This pragmatism can be seen as a genuine British tradition.[7] Guy's 1860 state of the art seems to have remained the British standard for the rest of the 19th century.

True, there was some interest in epistemology as manifested in 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 these remain the only translated German works on probabilistic issues I have been able to find. Two lectures by Jürgensen and Liebermeister on cold water used to treat fever were also translated and published by the Sydenham Society in 1877 (Liebermeister, Jürgensen, 1877b). Yet, although they included some of Liebermeister's and Jürgensen's statistics, the study design and the methods underlying them, let alone Liebermeister's work on formal probabilistic issues and his 'four-table-test (described in the December 2020 issue of the *JRSM*), were not mentioned, let alone translated. Trousseau's diatribe against the numerical method continued to be hidden in the English edition of his long-winded *Leçons cliniques* (Trousseau, 1868; Tröhler, 2020). Clearly, the (editor's) interest was in clinical issues.

When German authors became involved in probabilistic thinking the 1840s, they could – and actually did – draw from French and British sources. Local contemporaneous networks emerged in Tübingen, Kiel and Breslau through personal collegiality. They tackled some problems in applying formal probabilistic techniques to clinical practice. Over time, solutions were proposed, then deemed too complicated, prompting proposed new solutions and so on, in repeating cycles.

I do not yet 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 these two centuries, and in all three of the principal countries studied, there was an increasing awareness among clinicians of a need to publish dependable information about medical achievements. The simple fact that evolving probabilistic modes of

Table 2. Translations of works treated in the text.

Original language Author/Date	Translation into: English	French	German	Other languages
Latin				
J Bernoulli 1713	–	–	–	
French				
Cabanis 1798	+		–	
Pinel Nosographie 1798	+		+	
Traité 1800	+		+	
Laplace 1812	+		+	Dutch
Gavarret 1840	–		+	
Louis Anatomo-pathologie de la phtisie 1825 Effets de la saignée 1835	+ +		+	
Trousseau 1865	+		+	
English				
Lind 1772		–	+	
Gregory 1772		–	+	Italian, Spanish
Haygarth 1784		–	+	
Black 1789		–	+	
Blane 1819		–	+	
German				
Oesterlen 1852	+	–		
Radicke 1861	+	–		
Jürgensen 1876	+	–		
Liebermeister 1876	+	–		

thinking and clinical action had encroached on minds over 200 years made a difference, particularly in the long run. The meaning of probability 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 over the two centuries, suggest 10 enduring features.

First, it reveals that most French, British and German clinicians throughout the two centuries were not aware of the underlying probabilistic nature of their thinking and action when counting and analysing their cases using Mode I quantification (according to Table 1).

Second, it is equally clear that only a 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 material on this topic.

Fourth, I have noted a consistently forward-looking, future-oriented attitude among these authors, explicitly expressed, for example, by Todd (1831); Hodgkin (1834/1854); Griesinger (1848); Wunderlich (1851); Schweig (1854); Guy (1860); Jürgensen 1866; Jessen (1867); Petersen (1877) and Rosenbach (1891).

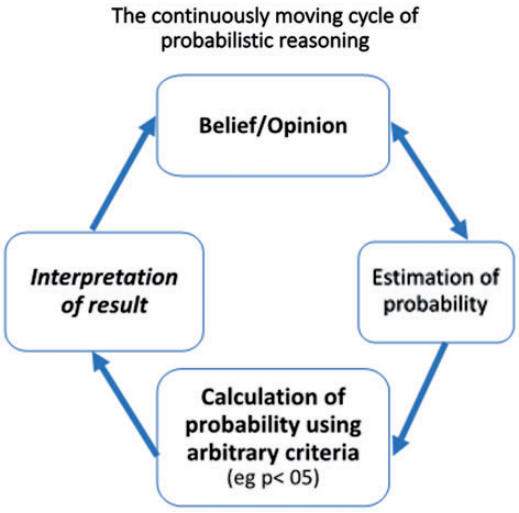
Fifth, most of the clinicians were (still) in marginal social positions when publishing on mathematically orientated probabilistic evaluation. For various reasons, all those who later achieved a recognised academic position dropped the interest manifested during their younger years. Some may have resigned 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, mathematically orientated authors marginalised themselves and the calculus of probabilities intellectually over the 200 years covered in this study.

Eighth, over the two centuries, 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.[8] 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.[8] 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, is medicine a 'Science' or a (healing) 'Art' – both requiring definitions? These are enduring questions with no universally agreed answers.

Ninth, the long-term perspective makes clear that the arguments advanced for or against probabilistic thinking persisted over the two centuries, and they were used in whatever way supported each of the differing analyses and perspectives of the disputatious parties. In fact, interpreting the results of calculations is an essential step in the cycle of probabilistic reasoning.



Tenth, the few supporters of probabilistic thinking – who were really fighting an uphill battle – were all engaged in clinical research. This suggests that only very rarely before 1900 was clinical practice guided by conscious probabilistic reasoning pursued through informal quantification, let alone by using formal methods (according to Table 1).

The above 10 persistent features may help to explain why clinical medicine never really adopted even informal probabilistic thinking before the end of the 19th century, let alone the rigorous conditions of statistical testing 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 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 of these – 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, 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,[9] p.14)

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

Outlook into the 20th 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 commentators 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, anti-sepsis and asepsis. New therapeutic possibilities were 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 historical controls (Tröhler,[10] pp.97–120; Tröhler[2]). 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 declare publicly 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 a 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 standard conditions; using uniform diagnostic methods; documented adherence to treatment; clinicial trials extending over sufficient time; and the concepts of placebo and blinding. In fact, the latter had been in the minds of some people since the 18th century (Jütte, 2013).[11–14] Some saw these as possibilities for improving 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, 1834/1854; 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 its time came. With hindsight, the 18th and 19th centuries were the long dawn of a 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). Yet none of them ever mentioned their 19th-century forerunners. Were they really unaware of them?

During the second half of the 20th century, evaluation science would become a more widely recognized discipline, particularly among young idealists from a variety of backgrounds, for example, by those who initiated the Cochrane Collaboration. But were they, once again, easily ignored 'marginal men'?

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 mathematisation and/or by adoption of Bayes' theorem?

Propagators of a quantifying probabilistic approach did not foresee the emergence of difficulties in quantifying medicine (Sheynin, 1978, p. 285)[8]: when one intricacy was recognised 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?[15,16] How should one deal confidently with uncertainty when the replication of trials addressing the same question yield apparently incompatible results?[17] Will the relevance of mathematically sophisticated, probabilistic evaluation decrease in view of 'personalized healthcare', with which new genetics dazzles us?[18]

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 realised 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'.

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.

Declarations

Competing Interests: None declared.

Funding: None declared.

Ethics approval: Not applicable.

Guarantor: UT

Contributorship: Sole authorship.

Provenance: Invited contribution from the James Lind Library.

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 and Marcel Zwahlen, 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.

Supplementary Material: 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.

References

1. Tröhler U. *"To Increase the Evidence of Medicine". The 18th Century British Origins of a Critical Approach*. Edinburgh: Royal College of Physicians, 2000.
2. Tröhler U. 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, 2014. See https://www.jameslindlibrary.org/articles/statistics-and-the-british-controversy-about-the-effects-of-joseph-listers-system-of-antisepsis-for-surgery-1867-1890/ (last checked 25 January 2021).
3. Müllener ER. Pierre-Charles-Alexandre Louis' Genfer Schüler und die «Méthode numérique». *Gesnerus* 1967; 24: 46–76.
4. Hannaway C. Louis, Pierre-Charles -Alexandre. In: Bynum WF and Bynum H (eds) *Dictionary of Medical Biography*. Vol 3, Westport, London: Greenwood, 2007, pp.814–815.
5. Hammerstein N. *Antisemitismus und die deutschen Universitäten*. Frankfurt/M: Campus, 1995.
6. Weber T. Anti-semitism and philo-semitism among the British and German elite in Oxford and Heidelberg before the first World War. *Engl Hist Rev* 2003; 118: 86–119.
7. Pickstone JV. *Ways of Knowing: A New History of Science, Technology and Medicine*. Manchester: Manchester University Press, 2000.
8. Matthews RAJ. 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, 2020. See https://www.jameslindlibrary.org/articles/the-origins-of-the-treatment-of-uncertainty-in-clinical-medicine-part-1-ancient-roots-familiar-disputes/ (last checked 25 January 2021).
9. Vandenbroucke JP. Clinical investigation in the 20th century: the ascendancy of numerical reasoning. *Lancet* 1998; 352: 12–16.
10. Tröhler U. *Der Nobelpreisträger Theodor Kocher 1841–1917*. Basel-Boston-Stuttgart: Birkhäuser, 1984.
11. Booth CC. John Haygarth FRS (1740–1827). JLL Bulletin: Commentaries on the history of treatment evaluation, 2002. See https://www.jameslindlibrary.org/articles/john-haygarth-frs-1740-1827/ (last checked 25 January 2021).
12. Donaldson IML. 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, 2016. See https://www.jameslindlibrary.org/articles/antoine-de-lavoisiers-role-in-designing-a-single-blind-trial-to-assess-whether-animal-magnetism-exists/ (last checked 25 January 2021).
13. Stolberg M. Inventing the randomized double-blind trial: The Nürnberg salt test of 1835. JLL Bulletin: Commentaries on the history of treatment evaluation, 2006. See https://www.jameslindlibrary.org/articles/inventing-the-randomized-double-blind-trial-the-nurnberg-salt-test-of-1835/ (last checked 25 January 2021).
14. Kaptchuk TJ. Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998; 72: 389–433.
15. Amrhein V, et al. Retire statistical significance. *Nature* 2019; 567: 305–307.
16. Matthews RAJ. 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, 2020. See 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/ (last checked 25 January 2021).
17. Prasad V and Cifu A. *Ending Medical Reversal*. Baltimore, MD: Johns Hopkins University Press, 2019.
18. Senn S. Individual response to treatment: is it a valid assumption? *BMJ* 2004; 329: 966–968.