Conclusions and perspectives for the 20th century, part 1: 200 years of discussion

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, together with the next one to be published in the forthcoming March issue of the JRSM, concludes our series on probabilistic thinking and the evaluation of therapies, 1700–1900.

Modes of probabilistic thinking

Arguably the first outflow of probabilistic thinking in medicine was to evaluate the effects of smallpox inoculation in 18th-century England: numerical comparisons of death rates of inoculated and un inoculated groups were made by mathematically inclined clinicians such as James Jurin, Secretary of the Royal Society. From then on, probability became a problem of numbers, of quantification. 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 observationists’ 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 by my research.

Another mode of thinking spoke of probability explicitly, but in practice still used informal, pre-mathematical quantification, that is, 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, concerning inoculation of smallpox in particular, that reasoning became mathematically probabilistic, and therefore conscious (Mode III in Table 1).

This perspective emerged in a violent debate between a French mathematician who had contradicted a Swiss colleague in the Paris Académie des Sciences.

This probabilistic 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. Gavarret’s work 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 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 and Tröhler). Probabilistic thinking became common practice from the second half of the 19th century onwards. I have selected the authors considered in the present essay because of 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. From the 1840s onwards, members of both these groups referred to Gavarret (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,4 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 of treatment effects (Mode II in Table 1), but not for physiological phenomena.

Practitioners of 18th-century British arithmetic observationism and its later French sibling the méthode numérique 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,5 p. 396). They can be seen as a reaction against the arbitrary exercise of authoritarian personal powers characteristic of the Ancien Régime. A more democratic society would call for trustworthy action 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 for ‘fevers’? In other words, what were the risks of medical innovation? This led also to questioning the

---

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

<table>
<thead>
<tr>
<th>Features</th>
<th>Modes</th>
<th>Typical representatives</th>
<th>Other representatives dealt with in the text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconscious implicit Numerical, but not mentioning probability</td>
<td>Mode I</td>
<td>Jurin et al. from 1720s, Lind 1772, Black 1789, Blane 1819, Bisset Hawkins 1829, Louis et al. from 1835/1837</td>
<td>Faure 1747, McGrigor from 1815, Alcock 1823, Todd 1835, Cowan 1835, Balfour 1854</td>
</tr>
<tr>
<td>Conscious explicit Numerical and mentioning probability</td>
<td>Mode II</td>
<td>Gregory 1772, Pinel 1807, Guy 1839, Henle 1844, Schweig 1854</td>
<td>Lavoisier 1780s, Condorcet 1785, Hodgkin 1834/1854, Griesinger 1848, Oesterlen 1852, Wunderlich 1851, Trouseau 1865, Rosenbach 1896</td>
</tr>
</tbody>
</table>
value of long-established standard therapies themselves, like bloodletting, when checked against the natural course of disease, as reported, for instance, by Hamilton (1816).2,6,7

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 urged 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 were those not mere speculations! The truth lay in observed facts, and for many of them, assembled in groups described statistically. Fostering this philosophy motivated some clinicians. In turn it meant that 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 uncertainty about whether results calculated from groups might be applied to individual patients. I refer to this as the ‘group-versus-single patient/case problem’. It had been identified and addressed in the 1760s in the Paris inoculation debate; subsequently, in the Paris disputes in the 1830s and later in the 19th century by Henle (1846), Griesinger (1848) and Trousseau (1862). The conundrum persists and seems likely to remain a bone of contention.

And then, questions also arose concerning the moral status of inference: 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 century2 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, p. 66, 183).

In mid-1830s Paris (and in the following decade by Henle), the necessity of ‘good faith’ when working with statistics was stressed (Murphy,8 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-propter-hoc-fallacy’: it had already been identified in the 17th century9 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). This fallacy, which might be believing that association means causation, seems eternal to me.

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

Another controversial issue concerned the essential interpretation of data. Since it could imply value judgments, inferences and generalisations 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’.10

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 reversals’ and it was repeatedly mentioned (Richerand, 1825; Lancet, 1834/1835; 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 between treatments. Nowadays, the term ‘medical reversal’11,12 has sometimes been used when a medical intervention introduced enthusiastically but without adequate evidence, has been abandoned when better evidence revealed not only that it did not help but that it 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 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.\textsuperscript{2,7,13} 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,\textsuperscript{14} p. 1)

The term ‘risk’, derived from the French ‘risque’, 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 namely, numerical data to reduce uncertainty (possibly even to certainty...).

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 now know with hindsight – that the human brain does not recognise probability. 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. These prerequisites included 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, while deeming clinicians to be mathematically incapable (Martius). Researchers did not even apply those 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 hand and reason 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 conception of disease; Virchow’s cellular pathology; anaesthesia and its sudden consequence, modern surgery; and diagnostic innovations (such as the stethoscope, the ophthalmoscope, laboratory methods and radiology). Students could not be bothered with complicated epistemistic 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? This question will be addressed in the next and final article in this series.

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 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.

Provenance: Invited contribution from the James Lind Library.

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.

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

References