

The 'Comedy of Errors': Putting the principles of 'false positive' and 'false negative' error into context

JÖRG RUOF

The 'Comedy of Errors': Putting the principles of 'false positive' and 'false negative' error into context

■ Abstract

The 'Comedy of Errors' is an early Shakespeare play telling the story of two sets of identical twins (Antipholus and his servant Dromio of Ephesus versus Antipholus and Dromio of Syracuse) that were separated right after birth. Equally, 'false positive' (e.g., falsely assuming that treatment 'X' is superior to 'Y') and 'false negative' (e.g., falsely assuming that 'X' is not superior to 'Y') errors are closely related 'twins' but conceptually different. While 'false positive' error is more granular and accessible, 'false negative' error is more difficult to control for. Planning for sample size and power in clinical trials therefore mainly relies on control of 'false positive' error. Keeping the balance across these two archetypes of error is the ultimate challenge of clinical medicine as well as clinical research. Within the wooden sculpture of the representative of humanity, Steiner has provided his most visible characterization of the two sources of error. Similar to the main characters in Shakespeare's comedy he discriminates a 'master' and a 'servant' version of both Lucifer and Ahriman.

■ Keywords

'False positive' error
'False negative' error
Randomized clinical trials
'Comedy of Errors'
Representative of humanity

Introduction

The 'Comedy of Errors' is an early Shakespeare play, written between 1592 and 1594 before he turned thirty (1). It is located in the Greek and Sicilian mystery centres of Ephesus and Syracuse. The starting point of the play are two pairs of identical baby twins. The wife of the rich Syracusan trader Egeon delivered twins, and Egeon purchased a second pair of twins as future servants for his sons that were delivered at the same time by a poor servant woman. On the sea voyage to Ephesus their boat is hit by a tempest. The storm rages and the father lashes himself with one son and his servant baby to one mast. The mother is lashed with the other couple to the other mast. The boat breaks apart. Both father and mother with their respective son and servant survive but are saved on different boats, and thereby lose track of each other. The main narrative of the play develops when the grown twin brothers, Antipholus of Ephesus and Antipholus of Syracuse together with their respective twin servants meet again in Ephesus. Chaos and comedy unfold as everybody (servants, masters, family, and friends) gets confused with the different, albeit similar looking, personalities.

The aim of this essay is to further explore the concept of 'false positive' and 'false negative' error, a fundamental concept in medicine.

A qualitative approach to the two types of error

If a physician is telling a 70-year-old male patient that he is pregnant, it has to be assumed that this statement is wrong. As a condition ('pregnancy') is diagnosed, the perception is that the statement is 'false positive'. Reversely, if the physician is telling a 28-year-old, obviously expecting female patient that she is not pregnant, he is again wrong. However, this time he is failing to diagnose an existing condition ('pregnancy') correctly. So, the nature of his failure is qualitatively very different from the first error. This time it's 'false negative'. While the first error (which in statistical terms is also called 'type 1' error or ' α ' error) is based on the assumption of an untrue reality, the second failure ('type 2' error or ' β ' error) captures the negation of an existing reality.

If we consider clinical situations for example, the conversation with a patient addressing the prognosis and

treatment of an oncological condition immediately demonstrates that a 'false positive' diagnosis may cause serious harm (e.g. performing surgery that is not required) to a patient and should be avoided by all means. On the other hand, however, not diagnosing an existing oncological condition could equally harm the patient, as effective treatments might be delayed. 'Do-No-Harm' (2), one of the fundamental paradigms in clinical medicine, therefore requires the sound balance of the two types of errors. Identifying the right balance for each individual patient and the right situation and suitable timing to disclose and discuss available clinical findings is not a mechanistic process. Instead, it requires experience of the physician and an empathic, subtle and qualitative interaction with the patient. It is an essential part or 'the art of medicine'.

The two faces of error are not unique to medicine; criminal justice (3), for example, always has to balance the risk of convicting an innocent person ('false positive') versus letting the guilty go free ('false negative'). But in justice as in medicine the two sources of errors reach far beyond the individual situation and cover the various facets of 'public health': Regulatory bodies must weigh the risk and benefit of premature versus delayed provision of a marketing authorisation for a new medicine; payers continuously need to avoid additional costs for treatments without an additional benefit, but equally need to ensure funding for treatments that optimise care. In addition, patients are continuously balancing the risk of taking medicines without a clinical benefit or with large side effects versus missing treatments that would have been beneficial (4). Thus, the 'duality of errors', the two closely related albeit unequal 'faces of failure', is a universal principle with ramifications reaching into all aspects of clinical medicine and beyond.

A quantitative approach to the two types of error

Austin Bradford Hill, a founding father of evidence-based medicine (EBM) advanced medical science by his landmark epidemiological studies uncovering the interaction between smoking and lung cancer. Holding neither a degree in medicine nor in statistics he is perceived as the 'greatest medical statistician' of the twentieth century (5, p. 1521). His research was on the evidence of causation, and he identified nine different conceptual viewpoints (strength, consistency, specificity, etc.) that led to his hypothesis that 'smoking is a factor, and an important factor in the production of carcinoma of the lung' (5, p. 1522). The starting point of his research was the loss of any intuitive insights into causality: 'Here then are nine different viewpoints from all of which we should study association before we cry causation.' (6, p. 299) He absolutely honestly describes a situation where the 'DIA-GNOSIS', the gnostic insight, the intuitive judgement about what a particular illness is about, has been lost and he aims to approach this state of uncertainty by concluding on a causality only after a thorough assessment of all possible associations (7).

The Scottish physician Archie Cochrane, another founding father of EBM and a pioneer of randomised con-

Tab. 1: Cross table illustrating the principles of 'false positive' and 'false negative' error. Research Question: 'Is X superior to Y'? Null hypothesis: Treatment 'X' is NOT superior to 'Y'.

Research conclusion	Reality	
	'X' is not superior to 'Y'	'X' is superior to 'Y'
'X' is superior to 'Y'	'false positive' error	<i>no error</i>
'X' is not superior to 'Y'	<i>no error</i>	'false negative' error

trolled trials (RCTs) was taken prisoner of war (POW) in 1941 in Crete. Due to his medical background, he became a POW medical officer in various German POW camps, fighting Tuberculosis for example. Reflecting on the various available treatment options (e.g., bed rest, pneumothorax, pneumoperitoneum, thoracoplasty, etc.) he commented: 'I had considerable freedom of clinical choice of therapy: my trouble was that I did not know which to use and when. I would gladly have sacrificed my freedom for a little knowledge. I had never heard then of "randomised controlled trials", but I knew there was no real evidence that anything we had to offer had any effect on tuberculosis, and I was afraid that I shortened the lives of some of my friends by unnecessary intervention.' (8, p. 6) What he describes is a status of 'uncertainty', fundamental uncertainty in medical decision making. Referring to Shakespeare's 'Comedy of Errors', Cochrane depicts a situation where the boat has already been completely broken apart. Any intuitive insight in what could have been helpful for his patients has been lost and he is searching for external, factual evidence as a base for rational medical decision making. In his later career as a Professor of Tuberculosis and Chest Diseases at Cardiff University School of Medicine he aimed to overcome this uncertainty by developing clinical and epidemiological data: 'I have attempted over the years [...] the study of medical error and its control, and by encouraging RCTs, to reduce these inherent biases.' (8, p. 4)

The starting point of any clinical research is the development of a sound research question, for example, if treatment 'X' is superior to treatment 'Y'. This research question is then usually transformed into a 'null hypothesis' (H0) that you want to reject i.e., 'X' is NOT superior to 'Y'. A 'false positive' error occurs when H0 is rejected and the alternative hypothesis is accepted although it is untrue. A 'false negative' error occurs if H0 is accepted, and the alternative hypothesis is rejected although it is true. It is important to recognise, that you don't test for your initial research question i.e., if 'X' is superior to 'Y'. Instead, you hypothesise the opposite (H0: 'X' is NOT superior to 'Y') and by rejecting H0 you aim to increase certainty that 'X' is superior to 'Y'. I.e., confidence in the superiority of a treatment is achieved by rejection of the hypothesis of non-superiority. Somewhat simplified, the cross table including both, the reality and the research conclusion is often used to visualise the principles of 'false positive' and 'false negative' error (see Tab. 1).

In their 1928 landmark article 'On the use and interpretation of certain test criteria for purposes of statisti-

cal interference', Neyman and Pearson (9) elaborate on how to best control for these two sources of error. The focus thereby is on inferential statistics to control for 'false positive' error. The authors acknowledge: 'The second source of error is more difficult to control [...]' (9, p. 177) In line with their considerations, until today, clinical trial planning aims to limit the 'false positive' error to achieve a high level of certainty when rejecting H_0 and thus accepting the alternative hypothesis (that 'X' is indeed superior to 'Y'). Thus, 'false positive' and 'false negative' errors point in opposite directions, but key characteristics are asymmetrical. While tight control for the 'false positive' error is essentially a core element in each clinical trial, sample size and power calculation, controlling for 'false negative' error, is not equally addressed. Another example is within the methods paper of the German Institute for Quality and Efficiency in Health Care (IQWiG), a contemporary core guidance document on medical statistics, where a detailed elaboration on how to control for 'false negative' error is lacking (10).

The quest for certainty and the importance of 'keeping the balance'

Commenting on the two types of error and on the key principles of medical statistics, the neurologist and founder of the German University of Witten/Herdecke, Gerhard Kienle, commented (11, p. 341): 'Im statistischen Sinne gibt es keine Sicherheit, sondern nur für den jeweiligen Fall festzulegende Irrtumswahrscheinlichkeiten.' ['In a statistical sense, there is no certainty, only the probability of error, which has to be determined in each case.'] In line with Bradford Hill's above-mentioned warning not to prematurely 'cry causation' and instead focus on association, Kienle points to the fact that statistics may only provide probability estimates and thereby do not satisfy the underlying quest for certainty in medical decision making.

Therefore, additional evidence is required to increase certainty. The concept of evidence-based medicine does reach beyond clinical data and includes the perspective of the patient and the clinician. As David Sackett has put it, EBM is 'a systematic approach to clinical problem solving by the integration of best research evidence with clinical expertise and patient values' (12, p. 70). Balancing the 'probability insights' of the findings from clinical research with the experience gathered as a clinician and the characteristic of each individual patient constitutes the EBM triad. The quest for certainty may not be satisfied by just knowing the data and the probability of error, it requires balancing the data with the reality of clinicians and patients.

Keeping the balance of those three pillars of evidence is a major challenge for modern medicine. While the quantity of available data is continuously increasing, and technologies to manage those data, including recent trends in artificial intelligence, are evolving, the two other pillars, inclusion of clinical experience and patient values, are far less dynamic and less developed. Furthermore, key

public health decisions such as the regulatory assessment and approval of new medicines and the subsequent health technology assessments, that precede national reimbursement decisions, occur at the beginning of the life cycle of new medical technologies. Consequently, those situations are characterised by a scarcity of clinical and patient experience, challenging the balance of the EBM triad (13). This further supports the contemporary trend in medical decision making that clinical data outweigh the importance of the experience of the treating physician and the values of the involved patient. Interestingly, the increasing number of single arm pivotal clinical trials for rare conditions decrease any certainty that may be derived from clinical data. It will be very interesting to observe how health care systems will approach the increasing uncertainty and probability of error regarding the benefits of respective innovative medicines (e.g. gene therapies and targeted oncology medicines for small subsets of patients) in the future (14).

Double-entry bookkeeping – A guidance to anthroposophical medical doctors

In 1923 Steiner advised young physicians and medical students to apply double-entry bookkeeping to fully reflect exoteric facts and insights as well as esoteric and anthroposophical considerations (15). In an attempt to apply that method, one might contrast the outlined concept of duality of 'false positive' and 'false negative' error with key elements in anthroposophic medicine.

Polarity and the quest to keep the balance across two opposite poles is a fundamental concept in anthroposophic medicine. The contrast of inflammatory versus oncological conditions, of the nervous system versus the digestive and movement system, or the polarity of the small- versus big-headed children in curative pedagogy: polarity and the dynamic that is created by overcoming it occurs all throughout anthroposophic medicine.

Within the wooden sculpture of the representative of humanity between the two adversary powers Lucifer and Ahriman, Steiner has provided his most visible and detailed legacy of the ongoing quest for keeping the balance. The sculpture represents the archetypal battle of humanity between the opposite forces and the ultimate task to 'keep the balance'. Referring to the wooden sculpture, Steiner explains (16, p. 152): '[...] dass der Mensch, wie wir ihn vor uns haben, eigentlich nur vorzustellen ist dadurch, dass wir alles an ihm als einen Gleichgewichtszustand vorstellen' ['that the human being as we have him before us, can actually only be imagined by imagining everything about him as a state of equilibrium']. The task ahead for anthroposophic medicine is always to seek this balance.

Clinical medicine has to match the principal conceptual insights into types of diseases, physiological or pathophysiological understandings, or the generic concepts of treatment with each individual bedside situation. While the latter represents the individualised version i.e., the biographical occurrence and relevance of a certain condition, the former represents the under-

lying general concepts. Experience, of a rheumatologist for example, develops after he has witnessed a variety of patients suffering from connective tissue conditions, thereby developing a 'conscious understanding and feeling' regarding this condition. By approaching the individual patients that are expressing a certain condition, the physician gradually gathers experience, which allows him to get an ever better understanding of the underlying condition.

To leverage the intuitive language of Shakespeare's brilliant and intuitive 'Comedy of Errors' one might consider a hierarchical relationship of these two levels, the 'Begriff' (i.e. the generic underlying concept) of a disease versus the individualised, visible expression of the disease: 'master' and 'servant'. The polarity of the 'master' level frames the conceptual dynamics of Shakespeare's play while the polarity of the 'servant' level is more tangible, more accessible and often very important to advance the storyline of the play.

Equally, Steiner's wooden sculpture displays the two facets of each type of error, a 'master' and a more individualised 'servant' version for both Lucifer and Ahriman. In a lecture on December 13, 1919 (17) Steiner described the 'double twins' (Fig. 1): 'Master' Lucifer on the top left, 'master' Ahriman on the bottom and the individualised version on the same level as the representative on his right side: 'Sie werden dann an dieser Gruppe zwei Figuren sehen: Hier den stürzenden Luzifer, hier den hinaufstrebenden Luzifer. Hier unten, gewissermaßen verbunden mit Luzifer, eine ahrimanische Gestalt, und hier eine zweite ahrimanische Gestalt. [...] Der Ahriman-Gestalt: alldem, was den Menschen führt zur Petrifizierung, zur Sklerose; der Luzifer-Gestalt: Repräsentanz alles dessen, was den Menschen fiebrig über das Maß derjenigen Gesundheit hinausführt, das er ertragen kann.' ['You will then see two figures in this group: Here the falling Lucifer, here the ascending Lucifer. Here below, in a certain sense connected with Lucifer, an Ahrimanic figure, and here a second Ahrimanic figure. [...] the Ahriman figure: all that leads man to petrification, to sclerosis; the Lucifer figure: a representation of all that leads man feverishly beyond that measure of health which he can bear.] (17, p. 183)

Leveraging the statistical terminology, the 'false positive' error, the Ahrimanic error, appears more granular, is easier to characterise, more amenable to quantification and – in clinical trial planning – to control for. The master version of 'false negative' error, the Luciferic deviation, is dominating the top left scenery of the sculpture. It is far less granular, and in statistics – without surprise – less amenable to control.

An important aside from a medical perspective: In the sculpture 'master' and 'servant' Lucifer display different dynamics: located on the left versus right side of the representative; falling versus rising gesture. Equally, in statistics you may never generalise from the overall clinical trial findings to the individualised bedside situation. A medicine that has been shown in the pivotal

Fig. 1: Drawing derived from (17, p. 183)



clinical trial to significantly extend overall survival by, for example, 6 months, might still cause more harm than benefit for a distinct patient situation, and vice versa. The 'art of medicine' needs to balance not just the fundamental concepts of 'ahrimanic' and 'luciferic' deviation, it also needs to be aware of individual dynamics that may counteract the overall trend observed in a clinical trial or in an epidemiological analysis. The direction the 'servant' takes might not necessarily be predestined by the 'master'. The words that Christ revealed to one of the 'malefactors' next to him on Golgotha: 'Verily I say unto thee, To day shalt thou be with me in paradise' (18) exemplifies this opportunity to reverse the overall 'master' trend in an individual 'servant' situation.

Both, the representative and Shakespeare's comedy also indicate a relation of the sources of error to the three dimensions of time: past/present/future. The representative is taking a step forward with his right leg suggesting a move. His 'presence' unfolds by balancing Lucifer (the 'extended' past) and Ahriman (the 'immature' future). Remaining bound to the past is related to enhanced Luciferic exposure while advancing too fast comes with increased Ahrimanic dynamics. The initial words of Dromio of Ephesus, when entering the stage for the first time, are related to the time dynamics between the two adversaries 'past' and 'future'. When asked by 'the other master' Antipholus of Syracuse: 'What now? How chance thou art return'd so soon?' he responds: 'Return'd so soon? Rather approach'd too late.' (1, p. 20) As a consequence of the developing error and confusing actions the servant is repeatedly scolded and beaten by the 'other master' for actions he did not do.

A risk in current medicine is that clinical judgement and patient perspective may be disregarded and that our technically dominated perspective on innovation and progress does narrow the insights into the real

origins of error and disease. Thus, unfolding ‘presence’ and obtaining a balanced, in-depth, and patient-centric understanding of diseases, as well as approaching and developing effective therapeutic interventions is a key task ahead for anthroposophic medicine.

To close this short essay, it is important to refer to the healing force of ‘warm humour’. Shakespeare designed his play as a comedy. The eternal tragedy and truth that unfolds early in the play reflecting the polarity of Antipholus of Ephesus and Antipholus of Syracuse and their respective servants Dromio of Ephesus and Dromio of Syracuse becomes ‘digestible’ through comedy and humour. Equally, Steiner positioned ‘humour’ in his wooden sculpture on top of all other figures in the right upper corner, reminding everybody that both seriousness and humour are required when facing the two adversary twin couples. Right in that context, the 2018 publication in the British Medical Journal covering the outcomes of an RCT researching parachute use to prevent death and major trauma when jumping from aircraft (19) does provide a great and humorous approach to ‘error’ in our data-centric medical environment.

Conflicts of interest

No conflict of interest has been raised.

Prof. Dr. med. Jörg Ruof
European Institute of Health Outcomes
Hauensteinstr. 132
4059 Basel, Schweiz
joerg.ruof@r-connect.org

References

- 1 Shakespeare W, Pache W (editor and translator). The Comedy of Errors. Die Komödie der Irrungen. Stuttgart: Philipp Reclam jun.; 1998.
- 2 Marsh H. Do No Harm. Stories of Life, Death, and Brain Surgery. London: Weidenfeld and Nicolson; 2014.
- 3 Givati Y. Preferences for Criminal Justice Error Types: Theory and Evidence. Pepperdine Caruso School of Law 2019. Available at <https://law.utexas.edu/wp-content/uploads/sites/25/Givati-Preferences-for-Criminal-Justice-Error-Types-4.pdf> (02.05.2024).
- 4 Rutten-van Mölken M, Ijzerman M. Editorial. Finding the Right Balance. European Access Academy. EAA Convention Proceedings 2024;5:6–9. Available at <https://irp.cdn-website.com/e52b6f19/files/uploaded/Proceedings%20EAA%2004%202024.pdf> (02.05.2024).
- 5 Doll R. Sir Austin Bradford Hill and the progress of medical science. *BMJ* 1992;305(6868):1521–1526. DOI: <https://doi.org/10.1136/bmj.305.6868.1521>.
- 6 Hill AB. The environment and disease: Association or causation? *Proc R Soc Med* 1965;58(5):295–300. DOI: <https://doi.org/10.1177/003591576505800>.
- 7 Ruof J. Aristotelische Ursache-Wirkungsprinzipien in der Medizin – Versuch eines Beitrags zur „konzeptuellen Integration“. *Der Merkurstab* 2016;69(2):110–116. DOI: <https://doi.org/10.14271/DMS-20617-DE>.
- 8 Cochrane A. Effectiveness and efficiency: Random reflections on health services. The Nuffield Provincial Hospital Trust 1971. Available at https://web.archive.org/web/20160527234617/https://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/Effectiveness_and_Efficiency.pdf (02.05.2024).
- 9 Neyman J, Pearson ES. On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. *Biometrika* 1928;20a(1/2):175–240. DOI: <https://doi.org/10.2307/2331945>.
- 10 Institute for Quality and Efficiency in Health Care (IQWiG). General Methods. Version 7.0. 19.09.2023. Available at https://www.iqwig.de/methoden/general-methods_version-7-0.pdf (02.05.2024).
- 11 Kienle G. Arzneimittelsicherheit und Gesellschaft. Eine kritische Untersuchung. Stuttgart, New York: Schattauer Verlag; 1974.
- 12 Sackett DL, Rosenberg WMC, Gray JAM, et al. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312(7023):71–72. DOI: <https://doi.org/10.1136/bmj.312.7023.71>.
- 13 Schlegl E, Ducournau P, Ruof J. Different weights of the evidence-based medicine triad in regulatory, health technology assessment, and clinical decision making. *Pharm Med* 2017;31(4):213–216. DOI: <https://doi.org/10.1007/s40290-017-0197-3>.
- 14 Julian E, Gianfrate F, Solamoraes O, et al. How can a joint European health technology assessment provide an ‘additional benefit’ over the current standard of national assessments? Insights generated from a multi-stakeholder survey in hematology/oncology. *Health Econ Rev* 2022;12:1–12. DOI: <https://doi.org/10.1186/s13561-022-00379-7>.
- 15 Neisecke J. Rudolf Steiners „doppelte Buchführung“ zum Thema Entzündung und Geschwulst: Notizen vom Dezember 1920 (NB 43). *Der Merkurstab* 2020;73(6):415–421. DOI: <https://doi.org/10.14271/DMS-21287-DE>.
- 16 Steiner R. Die vierte Dimension. Mathematik und Wirklichkeit. GA 324a. Dornach: Rudolf Steiner Verlag; 1995.
- 17 Steiner R. Die Sendung Michaels. Die Offenbarung der eigentlichen Geheimnisse des Menschenwesens. GA 194. 4. Aufl. Dornach: Rudolf Steiner Verlag; 1994.
- 18 The New Testament with Psalms. Authorized King James Version. Nashville, TN: Holman Publishers; 2006.
- 19 Yeh RW, Valsdottir LR, Yeh MW, et al. Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. *BMJ* 2018;363:k5094. DOI: <https://doi.org/10.1136/bmj.k5094>.