On the normativity of evidence

Lessons from philosophy of science and the "VALIDATE" project

Refolo, P.; Sacchini, D.; Bloemen, B.; Grin, J.; Gutierrez-Ibarluzea, I.; Hofmann, B.; Oortwijn, W.; Raimondi, C.; Sampietro-Colom, L.; Sandman, L.; van der Wilt, G.J.; Spagnolo, A.G.

DOI
10.26355/eurrev_202312_34560

Publication date
2023

Document Version
Final published version

Published in
European Review for Medical and Pharmacological Sciences

European Review for Medical and Pharmacological Sciences 2023; 27: 11202-11210

On the normativity of evidence – Lessons from philosophy of science and the "VALIDATE" project

P. REFOLO[1,2], D. SACCHINI[1,2], B. BLOEMEN[3], J. GRIN[4], I. GUTIERREZ-IBARLUZEA[5,6], B. HOFMANN[7], W. OORTWIJN[3], C. RAIMONDI[1,2], L. SAMPIETRO-COLOM[8], L. SANDMAN[9], G.J. VAN DER WILT[3], A.G. SPAGNOLO[1,2]

[1]Research Center for Clinical Bioethics and Medical Humanities, Università Cattolica del Sacro Cuore, Rome, Italy
[2]Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
[3]Department for Health Evidence, Radboud University Medical Centre, Nijmegen, The Netherlands
[4]Department of Political Science, University of Amsterdam, Amsterdam, The Netherlands
[5]Basque Foundation for Health Innovation and Research (BIOEF), Barakaldo, Spain
[6]Basque Office for HTA (Osteba), Barakaldo, Spain
[7]Department of Health Science, Norwegian University of Science and Technology, Gjøvik, Norway
[8]Assessment of Innovations and New Technologies Unit, Hospital Clinic, University of Barcelona, Barcelona, Spain
[9]National Centre for Priorities in Health, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

Abstract. – "Evidence" is a key term in medicine and health services research, including Health Technology Assessment (HTA). Randomized clinical trials (RCTs) have undoubtedly dominated the scene of generating evidence for a long period of time, becoming the hallmark of evidence-based medicine (EBM). However, due to a number of misunderstandings, the lay audience and some researchers have sometimes placed too much trust in RCTs compared to other methods of investigation. One of the principal misunderstandings is to consider RCTs findings as isolated and self-apparent pieces of information. In other words, what has been essentially lacking was the awareness of the value-context of the evidence and, in particular, the value- and theory-ladenness (normativity) of scientific knowledge.

This paper aims to emphasize the normativity that exists in the production of scientific knowledge, and in particular in the conduct of RCTs as well as in the performance of HTA. The work is based on some lessons learned from Philosophy of Science and the European project "VALIDATE" (VALues In Doing Assessments of healthcare TEchnologies"). VALIDATE was a three-year EU Erasmus+ strategic partnerships project (2018-2021), in which training in the field of HTA was further optimized by using insights from political science and ethics (in accordance with the recent definition of HTA). Our analysis may reveal useful insights for addressing some challenges that HTA is going to face in the future.

Introduction

"Evidence" is a key term in medicine and health services research, including Health Technology Assessment (HTA). This is particularly true since the emergence and subsequent rise of the so-called "Evidence-based medicine" (EBM) movement in the early 1990s. Basically, EBM affirms that health and clinical decisions should not be based solely on intuition, expert opinions, and pathophysiological reasoning since they alone are potentially biased and unreliable sources[1]. Instead, such decisions should be based on the integration of individual clinical expertise and patient preferences with the best external clinical evidence that is available[2].

What is or constitutes the "best clinical evidence" is a complex, debated, and challenging issue[3]. EBM recognizes different sources of clinical evidence when determining the effectiveness of a particular intervention and establishes a hierarchical system of trustworthiness. According to the scientific rigor or robustness of methodological designs, the classification of evidence reliability places Randomized Clinical Trials (RCTs) at the highest level, followed by controlled observational studies, and finally by non-controlled studies and expert opinions[3].

Corresponding Author: Dario Sacchini, MD; e-mail: dario.sacchini@unicatt.it

RCTs undoubtedly dominated the scene of generating evidence for a long period of time, becoming the hallmark of EBM[4]. However, due to a number of misunderstandings[5], the lay audience and some researchers at times have placed too much trust in RCTs compared to other methods of investigation. One of the principal misunderstandings – which has been widely clarified by literature and by the EBM movement itself over time and can be considered definitively surpassed – is to consider RCTs findings as isolated and self-apparent pieces of information. In our understanding, what has been essentially lacking was the awareness of the value-context of the evidence, and in particular the value- and theory-ladenness (normativity) of scientific knowledge.

This lack of awareness of the normativity of evidence may have also affected the HTA field, and perhaps influenced its development. Indeed, HTA is frequently interpreted as an impartial retrieval, critical examination, and synthesis of evidence (preferably derived from RCT findings), leading to a subsequent deliberation regarding the value judgments that can be drawn from it.

This paper aims to emphasize the normativity that exists in the generation of scientific knowledge, and in particular in the conduct of RCTs as well as in the performance of HTA. The work is based on some lessons learned from Philosophy of Science and the European project "VALIDATE" (VALues In Doing Assessments of healthcare TEchnologies" available at: https://validatehta.eu/)[6]. VALIDATE was a three-year EU Erasmus+ strategic partnerships project (2018-2021), in which training in the field of HTA was further optimized by using insights from political science and ethics (in accordance with the recent HTA definition)[7]. The analysis may reveal useful insights for addressing some challenges that HTA is going to face in the future. To be clear, the aim of the present paper is not to criticize or question RCTs, to propose a new hierarchy of evidence, nor to say what is good and bad science.

On the Normativity of Science

Claude Bernard (1813-1878), the founder of the so-called "experimental medicine", was the first one to advance the claim that medicine should be highly scientific, meaning that it should be the science that applies the knowledge acquired in the laboratories at the patients' bedside[8]. Specifically, he promoted and analyzed the use of experiments in medicine and argued that the experimental method is carried out in three stages: observation, hypothesis, and experimentation[9]. Observation and experimentation are two extreme terms of "experimental reasoning". They provide the knowledge of "facts", but in between them there lies "experimental idea", also known as "idea a priori" or simply "hypothesis". The latter serves as the foundation of all scientific reasoning and it is the essential part of every discovery, but it is worthless if it is not followed and confirmed by experimental "verification" [10].

Bernard's perspective was highly congruent with Positivistic and neo-positivistic philosophy, two cultural movements that dominated the Philosophy of Science in the second half of the 19th century and early 20th century[6]. According to such cultural movements, science is a formal activity that generates and accumulates knowledge by directly observing and confronting (i.e., by doing experiments) the natural world. Science tries to discover how the world "really" is by observing it carefully. Hypotheses and theories are formulated and subjected to scrutiny based on observations, and only when theories make predictions that can be "verified" – which means that correspond to how the world really is (as known through human senses and the use of scientific instruments) – they can be deemed to be "true" [6].

The term "positivism" derives from the French word "positif" which means "imposed on the mind by experience". Popularized in the first half of the 19th century by the sociologist and philosopher Auguste Comte (1798-1857), its central claim was that sensory experience or observations are the only source of authentic knowledge[11].

In strict continuity with positivism, neo-positivism developed a program of radical "re-foundation" of knowledge exclusively based on logic and verification. Its aim was to construct scientific laws and theories to describe and express relationships among observable phenomena and to elaborate a "unified language" for science as a whole. Core notions of a neo-positivistic approach to science were:

1) a division exists between the tangible reality, referred to as "the given", and its interpretations and articulations made by humans;

2) the real world, "the given", can be grasped by empirical sciences;

3) the truth, ascertained by empirical sciences, can be articulated through protocol sentences (i.e., statements that depict direct experience or perception) and observation reports;

4) protocol sentences and observation reports can be understood, described and conceptualized by logical analysis, thereby bridging the division between the world and words, "the given" and its related concept, facts and theory[12].

The fundamental assumption on which the whole philosophical conception of neo-positivism was built is the well-known "verification principle", for which a statement is meaningful if it either states a logical tautology or is in principle capable of empirical verifiability. This implies that any statement that cannot be verified by an empiricist criterion is meaningless and that such verification is a complete and definitive establishment of truth. Metaphysics, ontology, as well as ethics fail such criteria, thus they are cognitively meaningless. In this way, neo-positivism celebrated a sort of divorce of science from ethics, of the empirical dimension from the normative one, or – by using a well-known philosophical lexicon – of facts from values.

However, how can we ascertain the truth of the verification principle? That is, how is it possible to ascertain through empirical investigation the truth of the proposition that expresses the principle of verification? Since this cannot be done, it must be admitted that such a proposition escapes verification. The neo-positivists aimed to eradicate metaphysics by employing a principle that, within the boundaries of the canons they set forth, paradoxically revealed itself to be metaphysical. Such problem was not resolved, and the entire program went away, being subsequently replaced by Karl Popper's falsificationism and by Thomas Kuhn's research paradigms.

Scientific hypotheses can never be proven; they can only be falsified or rejected, and this is a crucial challenge identified by Karl Popper (1902-1994)[13]. The scientist enhances the generation of knowledge by reinforcing hypotheses and refuting opposing ones. If, following rigorous testing, the hypothesis is not refuted, it remains corroborated. Consequently, even though the scientific method is the most effective approach to comprehending the world, we must acknowledge that scientific knowledge is never complete. It will always be subject to revision and, ultimately, susceptible to fallibility.

Popper, along with other philosophers like Thomas Kuhn (1922-1996), also contended that our observations are theory-laden, meaning observations cannot exist independently of the theories within which they are made[14]. "Normal science" occurs when the scientific community collectively adopts a "paradigm", which is "the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community"[14]. A paradigm represents the shared context, and within that context, the scientists will operate on the issues at stake. Nonetheless, there are instances when an established paradigm experiences a crisis, leading to a "scientific revolution" that replaces an old paradigm with a new one. With the rise of a new paradigm, the scientific community has a new set of assumptions and a new set of problems to be solved; Kuhn refers to these revolutions as "changes of world view" and suggests that "after a revolution, scientists are responding to a different world"[14].

Subsequently, Pierre Duhem (1861-1916) and Willard V. Quine (1908-2000) noted that any given body of evidence may support multiple theories, possibly in conflict with each other. As scientific theories are deductively underdetermined by data, scientists must rely on extra-empirical criteria to determine the merits of a theory over its empirically adequate rivals. Consequently, "this "extra-empirical criteria" is subject to the whims, preferences, biases, and social agenda of the researching scientists, and not the rigor of evidence-based adjudication[15]. While the "theory-ladenness" objection challenged the stability of observations themselves, the "underdetermination" thesis (also called Duhem-Quine thesis) undermined the stability of evidential relations.

From the mid-twentieth century, the aforementioned fact/value dichotomy received further criticism, losing its appeal. In this context, the "philosophical hermeneutics" initiated by Edmund Husserl (1859-1938) and Martin Heidegger (1889-1976) and developed by Hans-Georg Gadamer (1900-2002) played a fundamental role. Building upon the previously mentioned perspectives, Gadamer asserted that understanding is never static and conclusive, but always steeped in language, and thus dynamic and fluctuating. "Prejudices" serve as the fundamental basis for all forms of knowledge, and therefore, understanding is always personal, human, and subjective.

In a similar manner, the German philosopher Karl-Otto Apel (1922-2017) pointed out that moral language analysis always requires criteria to distinguish moral language from any other form of language. This helps to support the thesis that the existence of merely descriptive propositions (such as scientific propositions) is an illusion and that it is impossible to separate the normative dimension from the descriptive one. Furthermore, Apel

(1972) argued that the fundamental deficiencies of positivism originate from a failure to consider "the fact that all cognition of objects presupposes understanding as a means of intersubjective communication". In other words, science becomes incomprehensible as a human endeavor, if one cannot understand the implicit and explicit conventions and notions, or more generally, the communication community or language game, which it presupposes. Even tacit conventions about the use of words, not to say explicit conventions about definitions, theoretical frameworks, or statements of facts in empirical science imply "an intersubjective consensus about situational meanings and aims of practical life"[16].

Finally, the American philosopher Hilary Putnam (1926-2016) recently traced the "collapse of the fact/value dichotomy". According to him, a distinction should be recognized between statements of fact and statements of value, particularly those concerning ethical considerations, and such distinction could be beneficial in specific contexts. Nevertheless, a strict dichotomy between fact and value is philosophically indefensible because, on the one side, normative (e.g., ethical and aesthetic) judgments necessarily have a factual basis, and on the other side, scientific judgments encompass normative elements. Consequently, science cannot be considered "value-free", since "science itself presupposes values which are in the same boat as ethical values with respect to objectivity"[17].

These and other reflections considerably reduced the significance of the fact/value dichotomy and exerted a decisive influence on the methodological assumptions on which science is presently conceived. This turns out to be clear in the science of medicine too, since it has some major value-laden core concepts such as health, disease, and dysfunction[18]. Moreover, the moral core of medicine, i.e., to reduce pain, suffering, and premature death, is reflected in its science(s).

According to the new viewpoint, the notion of a universal standpoint for attaining "neutral knowledge" is rejected, and all forms of understanding, including scientific understanding, must be regarded as contextual, prescriptive, and influenced by underlying theories.

On the Normativity of HTA

The VALIDATE project applied many of the above-described insights to the field of HTA to allow its practitioners to understand better how facts that are being identified, selected, examined and interpreted, are related to specific normative commitments. The project is grounded on (i) taking Apel's assertion as a point of departure and (ii) seeking to provide a way to "do" ethical evaluation in full acknowledgment of Putnam's corollary claims.

In the current practice of HTA (Figure 1), value judgments are, in fact, often considered external to HTA or separate from it and are addressed by experts (in ethics) after the assessment is finished. An HTA usually starts by conducting an empirical inquiry into a health technology's safety, clinical

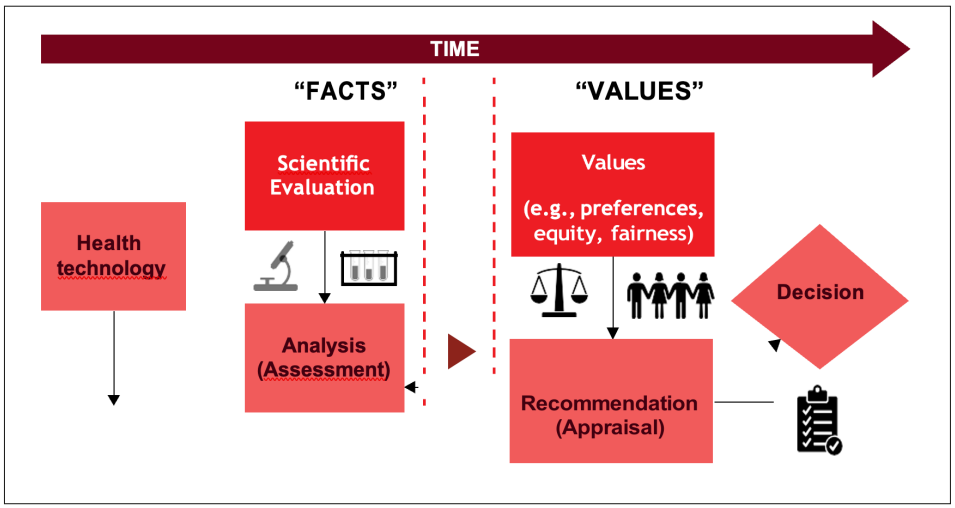

Figure 1. Conventional HTA decision process.

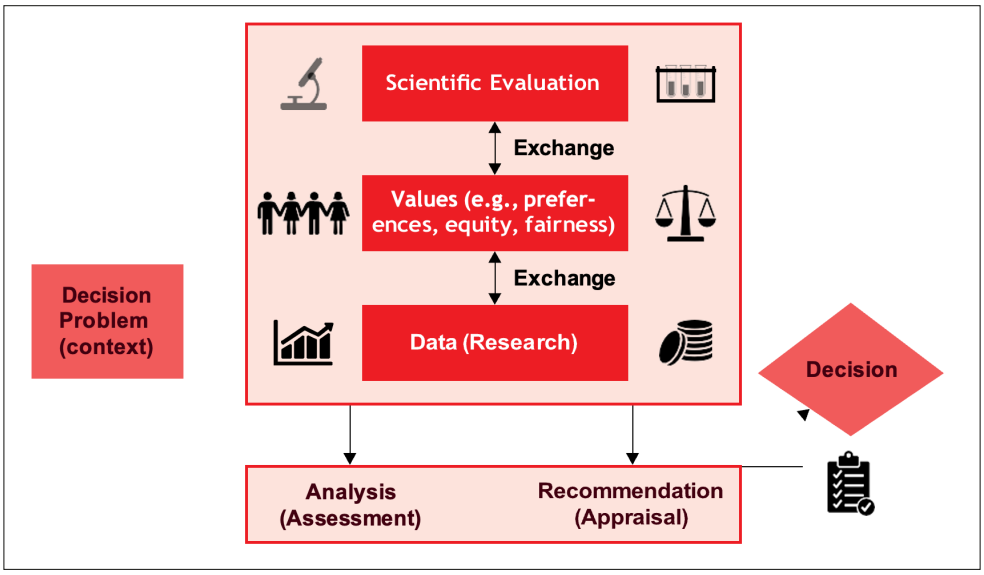

Figure 2. VALIDATE: integrative HTA process (Oortwijn and Sampietro-Colom 2022). VALIDATE proposes an integrative HTA approach, where the safety, clinical effectiveness, and cost-effectiveness of health technologies are thoroughly integrated with their wider ethical, legal, cultural, social, environmental, and organizational implications and stakeholders are involved in a more meaningful way during the entire HTA process.

effectiveness, and cost-effectiveness. This phase involves, for a large part, the analysis of RCTs. Subsequently, the question of whether any ethical issues may be anticipated, given some set of ethical commitments, such as beneficence, non-maleficence, autonomy, and justice, is asked[19]. In this way, HTA would consist of a sort of value-neutral retrieval, critical examination and synthesis of evidence, to be followed by a deliberation on the question of what value judgments can be derived from the available evidence.

Such an approach assumes that scientific facts (empirical inquiry of a technology's safety, clinical effectiveness, and cost-effectiveness) can be meaningfully separated from values (normative dimension). As such, the former would "work" as objective truth claims (they would tell something about some part of reality according to a positivistic approach to science), while the latter would be subjective and would come into play only after the facts have been collected.

As outlined above, to this day, a clear distinction between facts and values is philosophically difficult, if not impossible, to maintain. VALIDATE reaffirmed this notion and clarified that ethical commitments always guide empirical analysis, or at least they are implied in it. VALIDATE also clarifies that value frameworks are already operative at the stages when the questions to be addressed are proposed and the facts that answer those questions are collected (Figure 2). Ethical commitments are constitutive of HTA in the sense that, without acknowledging their existence, the very practice of HTA is not fully intelligible: practicing HTA means subscribing (usually tacitly) to these commitments. Therefore, value judgments permeate all levels of HTA[20,21].

The use of cochlear implants for prelingually deaf children is a well-known example in the field of HTA and it can clarify this fundamental point. A cochlear implant is a small and complex electronic device designed to offer a sense of sound to individuals who are profoundly deaf or experiencing severe hearing impairment. A lively debate has arisen among advocates of cochlear implants for prelingually deaf children and certain leaders in the deaf activist community. On the one hand, the latter consider the profoundly deaf as part of a distinct culture (minority culture) from the mainstream hearing society. Conversely, proponents of cochlear implants maintain that deafness is a disability, i.e., the failure to achieve an expected level of function. How to evaluate safety, effectiveness, and cost-effectiveness of cochlear implants? On the matter, data from RCTs have no real significance if they are not accompanied by a debate around the normative assumptions related to the health condition. It does not make sense and it is useless to collect facts without indicating which and whose problem they answer, and which

problem definitions are (may be unintentionally and implicitly) ignored.

Another example is bariatric surgery: when assessing the effectiveness of this intervention it may be highly important (to avoid a "Type III error": answering the wrong question) to examine the underlying assumption that reducing Body Mass Index is the most important endpoint. Obesity may be, in fact, framed as a problem that primarily results from poor lifestyle choices, or a problem that mainly results from an obesogenic environment. In the first case, education of the public is a logical solution. In the second case, regulation of food and marketing industries would be a more promising direction to take. Certainly, obesity may also be framed as a genetic problem. The point is that the choice of perspective is value-laden.

To be more precise, the aforementioned relationship between empirical dimension and normative dimension can be fruitfully explained – according to the VALIDATE project – in terms of plausibility, relevance, and amenability.

Plausibility refers to the fact that, during an HTA process, only those consequences of adopting a health technology are studied and are believed to be plausible. For instance, when someone is faced with the task of assessing a drug, there are some things that strike as potentially plausible (e.g., the patient will take the medication), while others as potentially implausible (e.g., the patient will throw the medication in the garbage at all times). In this sense, plausibility is a function of knowledge and understanding of the technology and the sort of impact(s) it may have. In short, plausibility refers to the question: what sort of effects are likely to occur, in view of our knowledge and understanding of the technology, and therefore which facts need to be considered?

Relevance is a function of norms and values. The commitment to norms such as the promotion of a certain conception of justice, or relieving suffering, or avoiding doing harm, or respecting autonomy determine the sort of data (outcomes) that become of interest. This means that not all data are relevant, or otherwise, that data of interest are influenced by normative commitments (from which perspective and why they have been selected). For example, if we need to know the long-term health benefit of a cancer drug to take a decision, data extrapolated from RCTs are useless, since they only show the health outcome over a limited time, and in a controlled setting.

Finally, amenability refers to the fact that it is possible to investigate those consequences of adopting a health technology through methodologically and epistemologically sound mode of empirical inquiry. For instance, if there is no way of determining how many patients will throw the medication in the garbage, no assessment of this will be possible. In this sense, amenability to scientific inquiry is a function of methodological and epistemological standards, and refers to the question: is this research feasible?

In conclusion, VALIDATE clearly affirmed that HTA always involves a close entanglement of facts and values, even though the latter may remain implicit or tacit. HTA is not a matter of collecting facts about a certain health technology, but rather "a matter of collecting facts that are considered plausibly associated with the use of the technology, relevant, and amenable to accepted methods of scientific inquiry"[22]. There can be no universal standpoint from which health technologies can be "evaluated", and all information must be considered as contextual, prescriptive, and value-laden. Therefore, understanding the different values and views becomes fundamental when determining which types of data should be used or which type of considerations should be done to address policy relevant questions. How to do this work is a different task. VALIDATE proposes to make use of the "reconstructing interpretive frames" method. This allows to (i) understand how different "pre-judgements" at the core of these frames generate different problem definitions, as well as to a focus on different facts and causal relations; and (ii) critically scrutinize different frames and hence appreciate how different assessments relate to each other[6].

Discussion

The accounts of scientific knowledge as "situated knowledges" offered by Philosophy of Science and the VALIDATE project reveal that "evidence is not so evident", in the sense that no kind of scientific evidence can be deemed to be self-apparent or neutral. Rather, it acquires relevance depending on the agreed questions to be addressed and the values at stake, and, therefore, it should be always placed in framework of meaning.

Similar considerations can help to better clarify the role of RCTs and, more generally, the whole "fabric" of evidence.

Firstly, RCTs are not theory-free, and RCTs findings are not automatically simply generaliz-

able, or automatically usable outside of the context in which they are obtained (poor external validity and generalizability). They do not reveal the "truth" about the world, but the truth in the trial sample only and within the conditions of the trial, that is, of an ideal or at least controlled world.

Secondly, the type of scientific evidence and the methodologies needed to analyze that evidence will depend on the research question being asked. The specific question at stake and the kind of background assumptions that can be acceptably employed will determine the most suitable methods to be used and in which combinations. Under this perspective, RCTs have no special status, and "RCT or no RCT" is not the same as "evidence or no evidence", or worse "truth or non-truth". The EBM itself recognizes that the best core of evidence for different questions cannot be reached at all times through RCTs. RCTs are not the gold standard in all situations; rather, they are the best experimental design available in some circumstances and with reference to a certain type of questions. In other words, the challenge is to find manners to identify and avoid unwarranted cognitive, affective, and normative biases in assessing evidence in order to be able to support well-supported decisions on health technologies. RCTs are part of the solution of this common enterprise or could be part of the problem if they are not correctly used or interpreted.

Thirdly, affirming that RCTs are the best experimental design in some circumstances does not exclude that further and also better designs could be developed for the same circumstances. Prior to the COVID-19 pandemic, the practice of clinical research had seen minimal revisions, having remained relatively unchanged for at least three decades. Somebody defined this phase as "stagnation of RCTs"[23]. Due to a number of factors, such as the changing nature of health technologies (heterogenous mix of chemicals, biologicals, digital, cell and gene therapies), a different approach to medicine (more patient-centered), advances offered by technology (like machine learning, AI, or digital platforms), and well-known limitations (most of which have been exacerbated by the pandemic), in recent times substantial progress has been made in the design, conduct and implementation of "master" protocols, which is leading to many changes in the practice. Umbrella study, basket study, platform study and master observational trial (MOT) are examples of this important transformation[23]. Furthermore, there are distinguished voices and scientific approaches that demonstrate that other alternatives could produce robust evidence to support decisions under certain conditions and can demonstrate even causality[24-26].

Fourthly, there is no dichotomy between RCTs findings and other types of evidence. In recent times, a number of issues have been raised about the increasing incorporation of non-randomized studies[27-29] and Real-World Data (RWD)[30-32].

In the light of the considerations stated above, it becomes clear that the future is not about RCTs *vs*. RWE but RCTs *and* RWE, with a more proper understanding of how they inform each other[33]. If one recognizes the role of the whole "fabric" of evidence and the normativity which accompanies any kind of empirical research, then that incorporation will not be seen as source of concern but, rather, as a resource. The key lies in acknowledging that RWD can serve various purposes: for instance, RWD can be utilized to help to complement information when extrapolating the long-term survival curve beyond the trial period for economic evaluation. Additionally, RWD can help to provide information about the comparators, such as the choice of relevant comparators reflecting clinical practice and treatment effects. Moreover, RWD can be employed to increase information on the generalization of evidence which is hardly captured in RCTs[32]. All of this is based on the varying values involved and the research questions that need to be addressed.

The challenge in the future should focus on how to integrate the different pieces of information. For example, National Institute for Health and Care Excellence (NICE) is now working hard to enhance the methods for embracing all available evidence to appraise innovative health technologies[34]. In February 2020, it announced a statement of intent that a broader range of data would be utilized to address evidence gaps, including electronic health record data and RWD from registries and clinical audits. To support this, NICE developed a Real-World Evidence (RWE) framework, which was published in June 2022, accompanied by a launch webinar[35]. The lifecycle approach to HTA[36-39] is another interesting example of how similar integration can begin to be realized.

Conclusions

Normativity is a key element of any kind of scientific enterprise. Scientific evidence-production is value-laden. This is particularly true in the

production of medical evidence, considering that medicine is a moral undertaking. Thus, disclosing the normative dimensions which accompanies any kind of empirical research can help to recognize the role of evidence. RCTs have no special status and no form of evidence is self-apparent. The research questions themselves define the best research design possible.

This observation can reinforce the idea that HTA is not a mechanical activity. It could be considered an art rather than a geometry, where the mosaicists (HTA researchers) paint their picture (evaluation) of a health technology by putting together the different pieces (information) with effects of light and shade[40-41].

Conflict of Interest
The authors declare that they have no conflict of interests.

Funding
This article received no grant from any funding agency in the public, commercial or not-for-profit sectors.

Informed Consent
Not applicable.

Authors' Contribution
All authors contributed equally to this work. All authors read and approved the final version of the manuscript.

ORCID ID
P. Refolo: 0000-0003-1055-160X
D. Sacchini: 0000-0002-1581-3018
B. Bloemen: 0000-0001-8856-2216
J. Grin: 0000-0001-7933-2574
I. Gutierrez-Ibarluzea: 0000-0002-4080-7963
B. Hofmann: 0000-0001-6709-4265
W. Oortwijn: 0000-0003-4499-8602
L. Sampietro-Colom: 0000-0001-7182-0231
G.J. van der Wilt: 0000-0002-5856-762X
A.G. Spagnolo: 0000-0002-5762-2164

References

1) Buts J, Baker M, Luz S, Engebretsen E. Epistemologies of evidence-based medicine: a plea for corpus-based conceptual research in the medical humanities. Med Health Care Philos 2021; 24: 621-632.
2) Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312: 71-72.
3) Worrall J. Evidence in Medicine and Evidence-Based Medicine. Philosophy Compass 2007; 2: 981-1022.
4) Spieth PM, Kubasch AS, Penzlin AI, Illigens BM, Barlinn K, Siepmann T. Randomized controlled trials - a matter of design. Neuropsychiatr Dis Treat 2016; 12: 1341-1349.
5) Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. Soc Sci Med 2018; 210: 2-21.
6) Oortwijn W, Sampietro-Colom L. The VALIDATE handbook: an approach on the integration of values in doing assessments of health technologies. Radboud University Press; 2022. Available at: https://www.jstor.org/stable/j.ctv2phprpn.
7) O'Rourke B, Oortwijn W, Schuller T. Announcing the New Definition of Health Technology Assessment. Value Health 2020; 23: 824-825.
8) Bernard C. Introduction to the Study of Experimental Medicine. Dover Pubns: Mineola; 1927.
9) Normandin S. Claude Bernard and an introduction to the study of experimental medicine: "physical vitalism," dialectic, and epistemology. J Hist Med Allied Sci 2007; 62: 495-528.
10) Bernard, Claude. Encyclopedia.com. Oxford University Press; 2018. Available at: https://www.encyclopedia.com/people/medicine/medicine-biographies/claude-bernard (Retrieved November 28, 2023).
11) Feigl H. Positivism. Britannica 2023. Available at: https://www.britannica.com/topic/positivism/The-critical-positivism-of-Mach-and-Avenarius (Retrieved November 28, 2023).
12) Burkard M. Conflicting Philosophies and International Trade Law. Switzerland: Palgrave Macmillan; 2018.
13) Popper K. The Logic of Scientific Discovery. New York, NY: Hutchinson & Co; 1959.
14) Kuhn TS. The Structure of Scientific Revolutions. 2nd ed. Chicago, IL: The University of Chicago Press; 1970.
15) Goldenberg MJ. On evidence and evidence-based medicine: lessons from the philosophy of science. Soc Sci Med. 2006; 62: 2621-2632.
16) Apel KO. The a priori of communication and the foundation of the humanities. Man and World 1972; 5: 3-37.
17) Putnam H. The Collapse of the Fact/Value Dichotomy and Other Essays. Cambridge, MA: Harvard University Press; 2002.
18) Hofmann B. Acknowledging and addressing the many ethical aspects of disease. Patient Educ Couns 2022; 105: 1201-1208.
19) Beauchamp TL, Childress JF. Principles of Biomedical Ethics, 5th edn. Oxford: Oxford University Press; 2001.

20) Hofmann B, Cleemput I, Bond K, Krones T, Droste S, Sacchini D, Oortwijn W. Revealing and acknowledging value judgments in health technology assessment. Int J Technol Assess Health Care 2014; 30: 579-586.

21) Hofmann B, Bond K, Sandman L. Evaluating facts and facting evaluations: On the fact-value relationship in HTA. J Eval Clin Pract 2018; 24:957-965.

22) Van der Wilt GJ, Reuzel R, Banta HD. The ethics of assessing health technologies. Theor Med Bioeth 2000; 21: 103-115.

23) Subbiah V. The next generation of evidence-based medicine. Nat Med 2023; 29: 49-58.

24) Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol 2016; 183: 758-764.

25) Hernán MA. Methods of Public Health Research - Strengthening Causal Inference from Observational Data. N Engl J Med 2021; 385: 1345-1348.

26) Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. JAMA 2022; 328: 2446-2447.

27) Kent S, Salcher-Konrad M, Boccia S, Bouvy JC, Waure C, Espin J, Facey K, Nguyen M, Rejon-Parrilla JC, Jonsson P. The use of nonrandomized evidence to estimate treatment effects in health technology assessment. J Comp Eff Res 2021; 10: 1035-1043.

28) Hatswell AJ, Freemantle N, Baio G. Economic Evaluations of Pharmaceuticals Granted a Marketing Authorisation Without the Results of Randomised Trials: A Systematic Review and Taxonomy. Pharmacoeconomics 2017; 35: 163-176.

29) Crispi F, Naci H, Barkauskaite E, Osipenko L, Mossialos E. Assessment of Devices, Diagnostics and Digital Technologies: A Review of NICE Medical Technologies Guidance. Appl Health Econ Health Policy 2019; 17: 189-211.

30) Pongiglione B, Torbica A, Blommestein H, de Groot S, Ciani O, Walker S, Dams F, Blankart R, Mollenkamp M, Kovács S, Tarricone R, Drummond M. Do existing real-world data sources generate suitable evidence for the HTA of medical devices in Europe? Mapping and critical appraisal. Int J Technol Assess Health Care 2021; 37: e62.

31) Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU-actions for stakeholders. Int J Technol Assess Health Care 2020; 1-10. Epub ahead of print.

32) Kang J. Real-World Data in Health Technology Assessment: Do We Know It Well Enough? In: Bremer A, Strand R (eds.). Precision Oncology and Cancer Biomarkers, Human Perspectives in Health Sciences and Technology 5. Available at: https://doi.org/10.1007/978-3-030-92612-0_12 (retrieved May 5, 2023).

33) Eichler HG, Pignatti F, Schwarzer-Daum B, Hidalgo-Simon A, Eichler I, Arlett P, Humphreys A, Vamvakas S, Brun N, Rasi G. Randomized Controlled Trials Versus Real World Evidence: Neither Magic Nor Myth. Clin Pharmacol Ther 2021; 109: 1212-1218.

34) National Institute for Health and Care Excellence (NICE). 2020. Reviewing our methods for health technology evaluation: Consultation, 2020. Available at: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation (retrieved May 5, 2023).

35) National Institute for Health and Care Excellence. NICE real-world evidence framework, 2022. Available at: https://www.nice.org.uk/corporate/ecd9. Accessed March 9, 2023 (retrieved May 5, 2023).

36) Gutiérrez-Ibarluzea I, Chiumente M, Dauben HP. The Life Cycle of Health Technologies. Challenges and Ways Forward. Front Pharmacol 2017; 8: 14.

37) Kirwin E, Round J, Bond K, McCabe C. A Conceptual Framework for Life-Cycle Health Technology Assessment. Value Health 2022; 25: 1116-1123.

38) Regier DA, Pollard S, McPhail M, Bubela T, Hanna TP, Ho C, Lim HJ, Chan K, Peacock SJ, Weymann D. A perspective on life-cycle health technology assessment and real-world evidence for precision oncology in Canada. NPJ Precis Oncol 2022; 6: 76.

39) Trowman R, Migliore A, Ollendorf DA. Health technology assessment 2025 and beyond: lifecycle approaches to promote engagement and efficiency in health technology assessment. Int J Technol Assess Health Care 2023; 39: e15.

40) Hofmann B. Health Technology Assessment - science or art? GMS Health Technol Assess 2013; 9: Doc08.

41) Refolo P, Sacchini D, Brereton L, Gerhardus A, Hofmann B, Lysdahl KB, Mozygemba K, Oortwijn W, Tummers M, van der Wilt GJ, Wahlster P, Spagnolo AG. Why is it so difficult to integrate ethics in Health Technology Assessment (HTA)? The epistemological viewpoint. Eur Rev Med Pharmacol Sci 2016; 20: 4202-4208.