

Of standard of reference and accuracy: the problem of truth in imaging

Stefano Fanti¹ · Elisabetta Lalumera²

Published online: 17 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Introduction

In this issue of *EJNMMI* we report a comprehensive review [1] of the most important clinical application of ¹¹C-choline PET, restaging of patients with biochemical recurrence of prostate cancer, in the form of a systematic review and meta-analysis of the literature. A major problem in conducting this review was the identification of a reference standard. It could be considered paradoxical that there is no agreement on the definition of a reference standard, but the difficulty was very clear from the beginning, and is one of the best known problems in the pooling of data, even if not frequently discussed. Out of 43 original articles that were judged potentially relevant and for which the full text was acquired for more detailed evaluation, no reference standard was reported in 5 studies, making it impossible to consider these studies further. Even considering the 29 articles finally selected, we had to use a composite reference standard (different for locoregional recurrence and distant metastases) because there was no consistency among the different studies.

In our analysis the reference standard for locoregional recurrence was the histopathological findings on TRUS-guided biopsy. For distant metastases any of the following were used: histopathological findings on lymphadenectomy or biopsy,

composite reference standard with clinical follow-up of at least 12 months, contrast-enhanced CT, MR imaging, bone scintigraphy, normalization or reduction of PSA values with salvage therapy, and clinical follow-up of at least 12 months for ¹¹C-choline PET-negative studies. The rationale for this approach was the impossibility of obtaining pathological confirmation for any finding, as well as the need for an unacceptably long follow-up in patients without imaging findings. We judged the composite reference standard as adequate and acceptable, thus allowing the calculation of sensitivity and specificity. The risk of bias in the interpretation of the reference standard was considered high when only contemporaneously acquired imaging without follow-up was used, and in such cases the term “detection rate” was used (i.e. when the standard of reference including other imaging methods were carried out at the time of investigation).

Truth

The standard of reference is essentially the closest thing to actual truth that we can achieve. Philosophers have speculated for centuries on the nature of truth. In brief, there is the metaphysical question ‘what is truth?’, and the epistemological question ‘can we know the truth, and how?’. Let us leave metaphysics for now. We discuss below three fictional philosophical characters who may exemplify three possible epistemological attitudes to the truth, and mirror the situations we faced:

1. Plato, or Platonism. Take all the facts of the world, and define god as whoever or whatever stores and scans the complete inventory of all the facts of the world. The truth is there to be seen from god’s point of view, independent of any possible or actual human judgment. It is possible

This Editorial Commentary refers to the article <http://dx.doi.org/10.1007/s00259-015-3202-7>.

✉ Stefano Fanti
stefano.fanti@aosp.bo.it

¹ Department of Nuclear Medicine, S.Orsola-Malpighi University Hospital, PAD 30, Via Massarenti 9, 40138 Bologna, Italy

² Department of Psychology and Philosophy, Bicocca University, Milan, Italy

that we know the truth from our limited perspective, but it is also possible that we do not know if we know it or not. This scenario is probably represented by obtaining a final pathological specimen of every finding, the super gold standard; unfortunately in practice in most cases it is not available.

2. Peirce, or Pragmatism. The truth is whatever appears at the end of scientific inquiry (convergence of methods, shared results). In this view, truth does not elude us, but it is hard to obtain. In Peirce's words "the opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real" [2]. This obviously sounds like a composite reference standard which is reasonably close to truth but at the same time reasonably feasible.
3. Descartes, or Methodological solipsism. Some truths are transparent to the expert's judgment, who trusts his own procedure. Paradigmatic examples are 'I am thinking', or 'I am here now', or 'I see a green patch of light'. To raise doubts about the true value of such reports would be not to understand the meaning of 'truth'. Notice that the report 'I see a green patch of light' is true also when you are hallucinating. This could be applied to studies where no efforts have been directed towards verifications that are not under investigators' control.

Example 1

Take the judgment 'Barack Obama is running in the White House park right now'.

For the Platonist, it is either true or false, and god knows which one. In the absence (or the temporary unavailability) of god's own privileged epistemological position, a drone over the White House or a GPS showing Obama's location wouldn't do, because none of such procedures is free from the possibility of error. Arguably, the Platonist is not much help if someone has to discover whether Obama is actually running or not – the ordinary domain of discourse and scientific judgments are poor candidates for platonic truths (for a platonic epistemology).

For the Pragmatist, 'Barack Obama is running...' is true if it appears to be true at the end of the inquiry, and it is false otherwise. Therefore the pragmatist would check his own sources, make a few phone calls, go to Washington himself, or send someone there; all such different means are fallible, and most of them rely on other people's knowledge, procedures and capacities. For the Pragmatist the search for the truth value is a communitarian, shared and fallible enterprise.

The Cartesian we consider here is endowed with a special power: like all Cartesians, he can issue a priori necessarily true judgments such as 'I am here now', 'I am thinking', 'now my

hands are cold'. In addition to normal simple introspective reports, on which he is completely authoritative like any of us would be, he could also produce reports on Obama's actions and thoughts, with an impressive score of true outputs. In the past, he has reported truly 'Obama is sneezing in Michigan', 'Obama is sleeping on his right side in Chicago', 'Obama is eating salted almonds in Salt Lake City'. Thus people tend to ask him about Obama's surroundings; unfortunately, however, he cannot explain how he produces such reports. When asked, he replies 'trust me, they are likely to be true'.

Example 2

Take the choline PET/CT imaging diagnosis 'Increased uptake (hot spot) located in a bone attributable to prostate cancer metastasis'.

For a Platonist nuclear medicine physician this could be acceptable (and true) only if there is pathological evidence of it. Given that in real clinical medicine it is impossible to biopsy every hot spot, the truth is almost unknowable. This argument has been frequently raised by clinicians reluctant to incorporate PET/CT into guidelines and clinical practice: EAU guidelines state that the accuracy of choline PET is "limited by the lack of a reliable histological gold standard" [3].

However for a Pragmatist nuclear medicine physician, the diagnosis is acceptable if confirmed by other unequivocal imaging approaches (such as a positive MRI scan) or even better if confirmed on comparative follow-up imaging (ceCT or MRI or bone scintigraphy at 4 – 8 months). In this case truth is reasonably achievable, and is based on comparison with all available data.

For a Cartesian nuclear medicine physician, the diagnosis is always acceptable, as the Cartesian considers his own finding as sufficient. The argument is that PET/CT is more sensitive than other methods, and false-positive findings are unlikely to occur. It is evident that such a perspective is scientifically quite weak and methodologically debatable. Nonetheless, a number of studies in which such criteria were used have been published in major journals.

Another problem derives from the lack of a universally recognized standard of truth, namely the overestimation of sensitivity by the Platonist and the Cartesian.

The Platonist approach should be as close as possible to the truth, which is simply not feasible in practice. Indeed several trials have required a pathological specimen of every positive finding, in our scenario mainly for local relapse. The resulting accuracies were frequently very high, and more than expected (in our paper, for any relapse it was 89 %). This is related to bias in patient selection. In fact, if a trial considers an invasive procedure such as biopsy mandatory for every PET/CT-positive case, it is obvious that referring clinicians will only

enrol patients with a very high likelihood of disease, as nobody likes to perform an aggressive procedure if it is likely to be inconclusive. Therefore, it is not surprising that the accuracy values we found using a composite reference standard (the Pragmatist's approach) were lower than in studies in which biopsy was performed systematically (in our paper 62 %); we feel that these values are more representative of real clinical practice. On the other hand, it is trivial to note that accuracy values found in studies using a Cartesian approach are always very high, given that what is reported is considered true. However, it is unlikely that these numbers turn out to be reproducible in daily practice.

In conclusion, the identification of a reference standard is a major problem in diagnostic imaging. We should be very careful when preparing new research protocols, as the choice of reference standard is crucial. We should be very careful when writing papers and using the terms “sensitivity” and “accuracy” as sometimes they do not reflect what we are measuring.

We should be consistent as a nuclear medicine community when identifying the truth if we are to be considered reliable by referring clinicians.

References

1. Fanti S, Minozzi S, Castellucci P, Balduzzi S, Herrmann K, Krause BJ, et al. PET/CT with 11C-choline for evaluation of prostate cancer patients with biochemical recurrence: meta-analysis and critical review of available data. *Eur J Nucl Med Mol Imaging*. doi:10.1007/s00259-015-3202-7.
2. Peirce CS. How to make our ideas clear. *Popular Science Monthly*. 1878;12:286–302
3. Mottet N, Bellmunt J, Briers E, van den Bergh RC, Bolla M, van Casteren NJ, et al. Guidelines on prostate cancer. *European Association of Urology*; 2015. p. 82. <http://uroweb.org/wp-content/uploads/EAU-Guidelines-Prostate-Cancer-2015-v2.pdf>