GUEST EDITORIAL

Meta-analyses and systematic reviews on PET and PET/CT in oncology: the state of the art

Giorgio Treglia · Ramin Sadeghi

Evidence-based medicine (EBM) is the practice whereby the best available evidence is incorporated into daily clinical practice in a reproducible way. This reproducibility is achieved using standard methods to search for evidence and evaluate the validity of that which is found. Actually, this reproducibility of EBM practice is the main difference between this approach and traditional medical practice.

The process of evaluating the validity of evidence and assigning a level of evidence to each single study is called "critical appraisal" and it is an important step in EBM. For each level of evidence, the best studies are systematic reviews and meta-analyses, which are the hallmarks of EBM [1].

Narrative review articles are very important in medical literature; however, the validity of the data presented in narrative reviews cannot be ascertained. Sometimes narrative review articles can be very misleading since they are the personal opinion of individual authors, and are not prepared using any predefined established method. The aim of a systematic review, on the other hand, is to present a balanced and impartial summary of the existing evidence on a specific topic. This aim is achieved using a clear search strategy and clear inclusion criteria, which makes their results much more reliable [1].

Meta-analysis is a statistical technique for combining the findings of independent studies included in a systematic review and it is often used to assess the clinical effectiveness of healthcare interventions. All meta-analyses are actually systematic reviews with components of statistical pooling of data. Conversely, not all systematic reviews have a meta-analysis component.

A systematic review/meta-analysis is performed in four stages:

First (stage I), an explicit scientific question to be addressed in the systematic review should be designed. In other words: what is the systematic review about? This stage usually results in the formulation of a four-component question called PICO. PICO stands for "patients", "intervention", "comparison", and "outcome". An example of a PICO question is: "How does ^{18}F-FDG PET imaging (I) for diagnosis of inguinal lymph node involvement (O) perform in comparison to inguinal lymph node dissection (C) in patients with penile squamous cell carcinoma (P)?

In the second stage, a search strategy should be designed. First, information sources should be selected. These sources are usually electronic databases such as PubMed or SCOPUS. The information sources should cover the PICO question as completely as possible. Databases with the widest journal and study coverage (such as SCOPUS or Google Scholar) are the best ones to include in the search. These databases are then searched using a keyword search strategy, based on the use of selected keywords. This search strategy should be as clear and as sensitive as possible to ensure identification of all relevant studies. As an example, for the PICO question mentioned above, "PET AND peni*" would yield a complete list of relevant studies.

In the third stage, an appropriate set of data is extracted from the identified studies relevant to the PICO question. For this purpose, inclusion criteria are set: publication year, language (only English articles or no language limitation), publication status (journal articles, meeting abstracts, etc.), and study design (randomized controlled trials for

G. Treglia
Department of Nuclear Medicine and PET/CT Center, Oncology Institute of Southern Switzerland, Bellizona, Switzerland

R. Sadeghi (✉)
Faculty of Medicine, Nuclear Medicine Research Center, Mashhad University of Medical Sciences, Mashhad, Iran
e-mail: sadeghir@mums.ac.ir; raminsadeghi1355@yahoo.com

therapeutic studies or type of reference standard for diagnostic studies). The inclusion criteria should be sure to capture all the high-quality relevant studies. After selection of the relevant studies, data and variables such as demographic data of the patients included in the selected studies, characteristics of the diagnostic tests performed, summary measures (sensitivity and specificity for diagnostic studies), etc., should be extracted. These variables should be preselected on the basis of knowledge of the main outcome and the possible related factors. Selecting the variables post hoc can affect the results of a systematic review and is a practice that is being discouraged. Finally, the quality of the included studies should be verified according to validated study quality assessment checklists such as QUADAS or the Oxford Center for Evidence-Based Medicine checklists for diagnostic studies.

In the final (fourth) stage, the synthesis of the results is performed. For diagnostic studies, five measures of performance—sensitivity, specificity, positive and negative likelihood ratios, and diagnostic odds ratio—are pooled and reported as the final result of a meta-analysis. Measures of heterogeneity (Cochrane Q value and test, I^2 index), threshold effect assessment (SROC curve fitting and correlation between sensitivity and specificity of individual studies), and publication bias evaluation (funnel plots, trim and fill method, etc.) should be performed and reported to ensure unbiased synthesis of the available evidence. If possible, additional analyses such as sensitivity or subgroup analyses, and meta-regression should be performed to account for confounding variables. The effect of the pre-test probability (prevalence of a disease in a selected population in diagnostic studies) on the predictive values of a diagnostic test can also be explored through Bayesian analysis. In this method, different pre-test probabilities are used to calculate post-test probabilities according to the pooled sensitivity and specificity provided by the meta-analysis. Bayesian analysis shows how useful and robust the test's predictive values are for populations with different pre-test probabilities. A complete description of the statistical analyses performed to ascertain diagnostic accuracy in systematic reviews/meta-analyses is beyond the scope of the current paper and readers can refer to other resources for further information [1, 2].

A checklist and a statement (see also Preferred Reporting Items for Systematic Reviews and Meta-Analyses - PRISMA [2]) have been developed specifically to ensure complete reporting of systematic reviews/meta-analyses, and many journals require authors of systematic reviews/meta-analyses to adhere to the PRISMA statement. The PRISMA statement consists of a 27-item checklist and a four-phase flow diagram and it can also be used for quality assessment of systematic reviews/meta-analyses [2].

All researchers are familiar with the concept of randomized controlled trials (RCTs). The RCT is the best

study design for therapeutic studies and systematic review of RCTs can provide the best evidence for a given treatment. The well-known Cochrane Database of Systematic Reviews (CDSR) is mostly dedicated to therapeutic studies and systematic reviews of RCTs.

On the other hand, the best design for a diagnostic study is a cross-sectional study with consistent reference standard application and blinding (best if prospective and multicenter). Not all diagnostic studies have such a high-quality design and the single studies may have included too few or too highly selected patients to ensure validity or generalization. This is the reason why we need systematic reviews and meta-analyses of diagnostic studies. High-quality meta-analyses can provide a precise estimate of a test's accuracy and can also allow the researchers to perform subgroup analyses to define accuracy in different subgroups of patients [1]. It should be emphasized that a good systematic review would include the best available evidence, consisting of prospective blinded studies with consistent application of the best available reference standard. However, these high-quality studies are rare, even in the "hot" diagnostic modalities such as PET and PET/CT. The readers of diagnostic systematic reviews should be aware of this fact and always look at the level of evidence of the individual studies included in a systematic review.

The recent medical literature shows a considerable growth in published systematic reviews and meta-analyses of diagnostic studies. This is also true in the field of nuclear medicine and specifically of ^{18}F-FDG PET and PET/CT in oncology. We conducted a comprehensive computer literature search of the PubMed/MEDLINE database to find published meta-analyses dealing with PET or PET/CT in oncology. We used a search algorithm that was based on a combination of the terms: (a) "meta-analysis" OR "metaanalysis" AND (b) "PET" OR "positron emission tomography". No beginning date limit was used; the search was updated to 31st December 2012. No language restriction was used.

The comprehensive computer literature search of the PubMed/MEDLINE database identified 438 articles. After reviewing the titles and abstracts, 302 articles were excluded because they were not within the field of interest. There, thus, remained 136 meta-analyses on PET or PET/CT in oncology published so far. Two-thirds of these meta-analyses on PET or PET/CT in oncology were published in the last 3 years and 40 % in the last year alone. This finding underlines the increasing interest in articles of this type in the field of nuclear medicine.

With regard to the PET radiopharmaceuticals used, most (93 %) of the meta-analyses on PET or PET/CT in oncology included studies that used ^{18}F-FDG, whereas only 7 % included studies using other PET radiopharmaceuticals.

Most meta-analyses on ^{18}F-FDG PET or PET/CT in oncology were diagnostic accuracy meta-analyses, having

sensitivity, specificity and accuracy as measures of outcome.

Different types of tumor were evaluated, mainly lung, head and neck, lymphomas, breast, colorectal, esophageal and uterine. But other types were evaluated too.

Several published meta-analyses pooled data from statistically heterogeneous studies. This heterogeneity is likely to stem from diversity in inclusion criteria, patient characteristics and methodological aspects between the included studies. In meta-analyses on the diagnostic performance of PET and PET/CT in oncology, findings from studies using PET and those using PET/CT are often pooled together. This could represent a possible bias on the final outcome, because the diagnostic accuracy of PET/CT is usually superior to that of PET alone.

In conclusion, there is an increasing bulk of synthetic literature (systematic reviews/meta-analyses) on the performance of PET or PET/CT in oncology. This body of literature mostly comprises ^{18}F-FDG PET studies in various malignancies. In the future, we definitely need to see more synthetic literature on the performance of PET and PET/CT in oncology, with particular attention to non-^{18}F-FDG PET tracers.

Conflict of interest   None.

References

1. Sadeghi R, Zakavi R, Kakhki VR (2009) How to apply the evidence-based medicine concept to nuclear medicine diagnostic studies—a review. Nucl Med Rev Cent East Eur 12:59–64
2. www.prisma-statement.org/statement.htm