DIAGNOSTIC IMAGING IN ONCOLOGY

Clinical trials for evidence-based radiology: to randomize or not to randomize? This is the question

Anna Colarieti[1] · Francesco Sardanelli[1,2]

Editorial

The goal of every physician is to optimize patient outcome through accurate decision-making. Over the last thirty years, evidence-based medicine (EBM) has been the cornerstone of this decisional process, providing an approach structured on the results of clinical research, i.e. the best "external evidence", as well as personal clinical expertise and patient's values and preferences [1, 2]. As stated by Dave L. Sacket in 1996 [3], "evidence-based medicine is the conscientious, explicit, and judicious use of current best evidence in making decision about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external evidence from systematic research". Technically speaking, "evidence-based medicine is the use of mathematical estimates of the risk of benefit and harm, derived from high-quality research on population samples, to inform clinical decision making in the diagnosis, investigation or management of individual patients", as underlined by Trisha Greenhalgh and Anna Donald in 2000 [4].

The "evidence" is given for a variety of levels, from experts' opinion (the lowest one, so-called eminence-based medicine) to randomized clinical trials (RCTs) and their meta-analyses (the highest level). The consequence of this ranking was that RCTs are widely regarded as the best research design providing a structure for data collection, mitigating bias, and generating dependable and applicable results. RCTs are typically conducted in controlled environments, involving specific patient populations, well-defined inclusion and exclusion criteria, and standardized

interventions [5]. However, RCTs have inherent limitations such as, not assured equalization between intervention and control group, not always precise intervention effect estimate, possible not fully explored effect of covariates, and applicability on strictly selected populations [6]. The last limitation may imply, for instance, difficulties in the case of comorbidities, more and more possible due to population aging. Important to note, RCTs should be conceived as "*part of a cumulative program [...] to discover not 'what works', but 'why things work"* [5]. At any rate, notwithstanding all these limitations, RCTs, when available, provide the evidence for choosing a treatment, including drugs, surgery, and any other treatment type such as interventional radiology or radiation therapy.

We must note that the deep logic of RCTs is based on the difficulty to try two (or more) different treatment/interventions in the same patient at the same time, which would practically erase any variability and confounding factors. In fact, ophthalmologist and dermatologists can administer two different treatments at the same time to the same patient, each of them to one eye or arm/leg. Outside these cases, different treatments/interventions cannot be administered to the same patient at the same time without mutual interference. For this reason, *RCTs were thought and are practiced as a surrogate of an intraindividual comparison*, which is the theoretical true best way to provide clinical evidence.

Thus, we firstly underline that *for diagnostic imaging intraindividual comparison is mostly possible*. In fact, the Oxford Center for EBM (http://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009) affirms that the highest level of evidence for studies on diagnostic performance is obtained through well-designed prospective cohort studies and comparative research. As a matter of fact, nobody asks for RCTs showing the advantages of MRI versus CT for diagnosing brain tumors, of CT pulmonary angiography versus radiography for diagnosing pulmonary

✉ Anna Colarieti
anna.colarieti@gmail.com

1 Unit of Radiology, IRCCS Policlinico San Donato, Via Morandi 30, 20097 San Donato Milanese, Milan, Italy

2 Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy

embolism, or of ultrasound versus clinical examination in the diagnosis of acute cholecystitis.

A special case is that of breast MRI for local and contralateral staging of breast cancer, where surgeons and oncologists sometimes ask for evidence from RCTs considering the underlying issue of risk of overtreatment, i.e. of more extensive surgery than that needed [7]. Here, there is the inherent difficulty of the relationship between a diagnostic procedure and patient's outcomes: numerous confounding factors come into play, including a large spectrum of surgery, systemic therapy, radiation therapy options, including the variability among surgeons. This complexity makes it challenging to attribute a clinical outcome specifically to a diagnostic test for local and contralateral staging. Insights on preoperative breast MRI are coming from prospective observational data from a very large study [8–10].

Conversely, prostate MRI has gained acceptance for the opposite reason, being proven to be a valuable tool in reducing unnecessary biopsies or surgical overtreatment, guiding targeted biopsy, and opening the possibility of MRI surveillance instead of surgery [11–13]. Here, the practical impossibility of evaluating MRI-targeted biopsy versus non-MRI-targeted biopsy in the same patient resulted into a RCT, which demonstrated the superiority of MRI-targeted biopsy [14].

Whenever possible, intraindividual prospective comparative studies focusing on sensitivity and specificity are the best way for evaluating the diagnostic performance. Thus, these studies, not RCTs, are the best evidence on which the choice of a diagnostic test must be based. Notably, this reasoning holds specifically for medical imaging applied in the "clinical" context, i.e. for patients who present symptoms or signs (or suspicious findings at a screening test). First level population-based screening tests are a different issue, even if performed with the same imaging modalities used in the clinical context. In fact, in the screening context, we need to eliminate lead time bias (when an early disease detection through screening appears to improve survival rates but only lengthens the time between diagnosis and clinical presentation) and length bias (screen-detected cancers show a lower growth rate than interval cancers). Thus, in case of screening, rigorous RCTs are required, with a sufficiently large sample size to deal with the low incidence, also in high-risk population. RCTs can provide evidence of the effectiveness of screening strategies by comparing outcomes in the screened population against a control group. One recent relevant example is the proposal of breast MRI for screening women with extremely dense breasts [15–17].

In conclusion, the type of evidence required depends on the context. For treatment evaluation, RCTs remain the best way. In case of diagnostic imaging (where symptoms or sign are present), intraindividual comparison studies focusing on sensitivity and specificity provide valuable evidence for diagnostic performance. For screening programs, RCTs are essential to eliminate biases and evaluate the societal impact. However, in the current era of big data and artificial intelligence, the position of RCTs as the top level of evidence could be reshaped by the increasing role of the so-called real-world evidence [18], which incorporates data from routine clinical practice and provides additional insights into treatment and diagnostic outcome. The comparison between RCTs and real-world evidence is an interesting perspective to be pursued.

Declarations

References

1. Sardanelli F, Hunink MG, Gilbert FJ, Di Leo G, Krestin GP (2010) Evidence-based radiology: Why and How? Eur Radiol 20:1–15. https://doi.org/10.1007/s00330-009-1574-4

2. Giovagnoni A, Ottaviani L, Mensa A, Durastanti M, Floriani I, Cascinu S (2005) Evidence based medicine (EBM) and evidence based radiology (EBR) in the follow-up of the patients after surgery for lung and colon-rectal carcinoma. Radiol Med 109:345–357

3. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS (1996) Evidence based medicine: what it is and what it isn't. BMJ 312:71–72. https://doi.org/10.1136/bmj.312.7023.71

4. Greenhalgh T, Donald A (2000) Evidence based health care workbook: understanding research for individual and group learning. BMJ Books, London

5. Hariton E, Locascio JJ (2018) Randomised controlled trials - the gold standard for effectiveness research: study design: randomised controlled trials. BJOG 125:1716. https://doi.org/10.1111/1471-0528.15199

6. Deaton A, Cartwright N (2018) Understanding and misunderstanding randomized controlled trials. Soc Sci Med 210:2–21. https://doi.org/10.1016/j.socscimed.2017.12.005

7. Houssami N, Ciatto S, Macaskill P, Lord SJ, Warren RM, Dixon JM, Irwig L (2008) Accuracy and surgical impact of magnetic resonance imaging in breast cancer staging: systematic review and meta-analysis in detection of multifocal and multicentric cancer. J Clin Oncol 26:3248–3258. https://doi.org/10.1200/JCO.2007.15.2108

8. Sardanelli F, Trimboli RM, Houssami N, Gilbert FJ, Helbich TH et al (2020) Solving the preoperative breast MRI conundrum: design and protocol of the MIPA study. Eur Radiol 30:5427–5436. https://doi.org/10.1007/s00330-020-06824-7

9. Sardanelli F, Trimboli RM, Houssami N, Gilbert FJ, Helbich TH et al (2022) Magnetic resonance imaging before breast cancer surgery: results of an observational multicenter international prospective analysis (MIPA). Eur Radiol 32:1611–1623. https://doi.org/10.1007/s00330-021-08240-x

10. Cozzi A, Di Leo G, Houssami N, Gilbert FJ, Helbich TH et al (2023) Screening and diagnostic breast MRI: how do they impact surgical treatment? Insights from the MIPA study. Eur Radiol. https://doi.org/10.1007/s00330-023-09600-5

11. Schoots IG, Roobol MJ, Nieboer D, Bangma CH, Steyerberg EW, Hunink MG (2015) Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. Eur Urol 68(3):438–450. https://doi.org/10.1016/j.eururo.2014.11.037

12. Brown LC, Ahmed HU, Faria R, El-ShaterBosaily A, Gabe R et al (2018) Multiparametric MRI to improve detection of prostate cancer compared with transrectal ultrasound-guided prostate biopsy alone: the PROMIS study. Health Technol Assess 22:1–176. https://doi.org/10.3310/hta22390

13. Del Monte M, Leonardo C, Salvo V, Grompone MD, Pecoraro M, Stanzione A, Campa R, Vullo F, Sciarra A, Catalano C, Panebianco V (2018) MRI/US fusion-guided biopsy: performing exclusively targeted biopsies for the early detection of prostate cancer. Radiol Med 123(3):227–234. https://doi.org/10.1007/s11547-017-0825-8

14. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Collaborators PSG et al (2018) MRI-targeted or standard biopsy for prostate-cancer diagnosis. N Engl J Med 378:1767–1777. https://doi.org/10.1056/NEJMoa1801993

15. Bakker MF, de Lange SV, Pijnappel RM, Mann RM et al (2019) Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med 381:2091–2102. https://doi.org/10.1056/NEJMoa1903986

16. Veenhuizen SGA, de Lange SV, Bakker MF, Pijnappel RM, Mann RM et al (2021) Supplemental breast MRI for women with extremely dense breasts: results of the second screening round of the DENSE trial. Radiology 299:278–286. https://doi.org/10.1148/radiol.2021203633

17. Mann RM, Athanasiou A, Baltzer PAT, Camps-Herrero J, Clauser P et al (2022) Breast cancer screening in women with extremely dense breasts recommendations of the European society of breast imaging (EUSOBI). Eur Radiol 32:4036–4045. https://doi.org/10.1007/s00330-022-08617-6

18. Sheldrick RC (2023) Randomized trials vs real-world evidence: How can both inform decision-making? JAMA 329:1352–1353. https://doi.org/10.1001/jama.2023.4855