VIA MEDICA

Nuclear Medicine Review 2009
Vol. 12, No. 2, pp. 59–64

ISSN 1506–9680

Editorial

How to apply the evidence-based medicine concept to nuclear medicine diagnostic studies – a review

Ramin Sadeghi[1], Rasoul Zakavi[2],
Vahid Reza Dabbagh Kakhki[1]

[1]Nuclear Medicine Research Center, Imam Reza Hospital, Mashhad University of Medical Sciences, Mashhad, Iran
[2]Nuclear Medicine Research Center, Imam Reza Hospital, Mashhad University of Medical Sciences, Mashhad, Iran

[Received 29 VI 2009; Accepted 29 VIII 2009]

Abstract

Evidence-based medicine (EBM) is defined as using the best available evidence for managing patients in daily healthcare practice. Although this approach has been applied successfully in many medical fields, it has not been addressed fully in the radiological discipline in general and nuclear medicine in particular. In this review, the concept of EBM has been introduced briefly and four steps of EBM practice have been explained. Asking answerable questions and finding the best evidence that constitutes the first two parts of EBM practice are explained in brief. The next two steps (appraising the available evidence and applying the best evidence) are explained in more detail. Since the bulk of nuclear medicine studies are of a diagnostic nature and most of the daily practice of a nuclear medicine specialist is involved in diagnosis, we have focused on the diagnosis studies. Systematic reviews are also explained to some extent. Appraisals of other kinds of study, such as interventional or prognosis studies, are not included in this review.

Keywords: evidence based medicine, critical appraisal, nuclear medicine

Nuclear Med Rev 2009; 12, 2: 59–64

Correspondence to: Vahid Reza Dabbagh Kakhki, Assistant Professor Nuclear Medicine Research Center, Imam Reza Hospital, Mashhad University of Medical Sciences, Ebn Sina Street, Mashhad, Iran
Tel: (98) (511) 8599359, fax: (98) (511) 8593038
e-mail: dabbaghvr@mums.ac.ir

Introduction

Evidence based medicine (EBM) is a relatively new approach to provide health care by which the best evidence is found for a definite clinical question and is applied in the most efficient way. Although this kind of practice is becoming the standard in patient care, it can be very time consuming and the users of this approach have to be equipped with efficient ways to succeed [1, 2]. However, application of EBM in radiology and allied disciplines (such as nuclear medicine) has not been addressed fully to date [3–6].

In EBM, standard techniques are used for searching the bulk of literature, and appraising the relevant data (regarding the validity and effect size). This property makes the results of the EBM approach reproducible. Reproducibility is the main difference between traditional and evidence based approaches to medical practice [2, 7–8]. Usually EBM practice is divided to 4 steps [7–8], which are shown in Table 1.

In this review these 4 steps are explained with the main focus on the diagnostic studies which constitute the major part of nuclear medicine practice.

Step 1: asking an answerable question

This step is the cornerstone of EBM practice. Asking an answerable question is defined as taking clinical data and converting them into a format to be used in the next steps mentioned above [9].

To achieve this goal, our clinical question has to be divided into several distinct sections:

1 — the specific group the patient belongs to;
2 — the test we want to apply for the diagnosis of the disease;

Table 1. Four steps of evidence based medicine (EBM) practice

Step I: Asking an answerable question
Step II: Searching for the best evidence
Step III: Critical appraisal of the evidence
Step IV: Applying the evidence to an individual patient

Table 2. Answerable question for a clinical scenario

Patient or problem	Adult patient with a history of treated pulmonary tuberculosis
Intervention	^{99m}Tc-sestamibi scintigraphy
Comparison	Sputum culture
Outcome	Differentiation of active from inactive pulmonary tuberculosis
Question	In an adult patient with a history of treated pulmonary tuberculosis, how sensitive is ^{99m}Tc-sestamibi scintigraphy for differentiation of active from inactive pulmonary tuberculosis?

3 — the test with which we would like to compare the test in part 2 (usually the gold standard);
4 — the intended outcome.

"PICO" is the acronym for this 4-part question: "P" for patient, "I" for intervention, "C" for comparison and "O" for outcome. This format for questioning is extremely helpful for searching the available literature for the best evidence.

The final question is usually expressed in a single sentence. The major parts of this sentence can be underlined for convenience in future searches. The following example would be helpful in this regard.

Assume that you are a Nuclear Medicine specialist, and the pneumonologist of your hospital wants to know if you can help in the differentiation between active and inactive tuberculosis in an adult patient with a history of treated pulmonary tuberculosis (6 months ago), since the results of the sputum culture take several weeks to be ready. The pneumonologist has heard about the application of ^{99m}Tc-sestamibi in this clinical situation in a scientific meeting and wants to know if ^{99m}Tc-sestamibi can help in this regard, or not. The answerable question for this scenario is shown in Table 2.

As mentioned above, without formulating a PICO question in every distinct clinical scenario, the other steps of EBM practice cannot be completed. It is recommended that adequate time be spent practicing this step [9].

Step 2: searching for the best evidence

In the growing world of internet and computer science, finding the best evidence is becoming increasingly difficult and time consuming. This is the main obstacle for finding the best evidence in the published medical literature. For example, Ely et al. reported that family physicians asked an average of 3.2 questions for every 10 patients seen in an ambulatory clinic, but only pursued answers to 36% of those questions [10]. In another study by Ely et al. inadequate time for search is introduced as one of the main reasons for not using the best evidence in daily practice [11]. If we want to use the best available evidence, we should be equipped with a good strategy for searching the literature.

Medical literature resources

Available medical resources for using in EBM practice are evolving very rapidly and not all resources are high in quality. It is recommended that the medical resources be categorized in a hierarchical manner [12]. Original studies are at the bottom of this hierarchy, synthetic literature (systematic reviews) are the next level, the synopses of studies and syntheses are the next, and finally evidence-based information systems (the most comprehensive resource) would be at the top of the hierarchy. This hierarchy is usually called the "4S" which stands for (from bottom to top): "studies", "summaries", "synopses", and "systems". When seeking the best available evidence, the highest-level resource available for the problem should be searched first [9, 12].

The evidence based "systems" include all relevant evidence regarding a clinical problem and are updated on a regular basis to provide the most up to date available evidence. These systems are usually under strict review process. Two of these "systems" are UpToDate (http://www.uptodate.com) and Clinical Evidence (http://www.clinicalevidence.com). They are not freely available online; however, those physicians who have University Athens system subscription can have access to them.

"Synopses" are the next level of medical resources. As the name implies, these are synopses of original studies and systematic reviews. Usually these "synopses" provide structured abstracts of studies that meet the preset criteria, in addition to an accompanying commentary by an expert on the study results. They are usually peer reviewed extensively as well. The "synopses" can be considered as a shortcut to find an answer to a clinical question. Two examples of this level of resources are Evidence Based Nursing and ACP Journal Club. They are not freely available on the net either.

Synthetic literature constitutes the next "S" which is "syntheses". Databases for systematic reviews are in this level. It should be noted that review articles, although very useful, are the opinions of individuals on a topic and can be very misleading since there is no standard method for preparing them. By way of contrast, systematic reviews and meta-analyses are objective reviews, which are more reliable than the review articles. The main characteristics of narrative review articles, systematic reviews, and meta-analyses are depicted in Table 3. The most comprehensive data-

Table 3. Characteristics of narrative review articles, systematic reviews, and meta-analyses

Narrative review articles	Systematic reviews	Meta-analyses
Methods to collect and interpret data are subjective	Methods to collect and interpret data are objective	Quantitative. Otherwise the same as systematic review
Not appraisable	Appraisable	
Replication impossible	Easily replicable	

base for systematic reviews is the Cochrane Library (http://www.cochrane.org). This database has a collection of systematic reviews prepared according to its own standards and it provides a list of other systematic reviews in the medical literature. It is not free but it can be accessed through an Athens system subscription.

Individual "studies" are in the bottom of the medical resources hierarchy. Searching for the best evidence in this level is time-consuming and found articles need to be critically appraised. Many databases are available for medical literature. Two of the most commonly used databases are SCOPUS (http://www.scopus.com) and Medline's Pubmed (http://www.ncbi.nlm.nih.gov/pubmed).

Although there are many medical resources which provide the best available evidence, currently only a few of them address the issues of radiology and nuclear medicine. For the questions regarding interventions and treatment, usually the Cochrane Library should be searched for any systematic review or RCT. If this search does not yield any useful article, Pubmed (or other resources) is recommended. However, when the clinical question is of a diagnostic nature (as is the case for the majority parts of Nuclear Medicine practice), Pubmed is recommended, especially the Clinical Queries option within it [13–14].

A detailed explanation of searching medical databases such as Pubmed is beyond the scope of this review and can be found elsewhere in the literature [9, 14].

Full-text articles: how to get them

Many Nuclear Medicine journals are freely available online. The Iranian Journal of Nuclear Medicine, the Hellenic Journal of Nuclear Medicine, Nuclear Medicine Review, Nuklearmedizin, the Quarterly Journal of Nuclear Medicine and Molecular Imaging, the Annals of Nuclear Medicine (not the recent issues), the Journal of Nuclear Medicine (not the recent issues), and the Journal of Nuclear Medicine Technology (not the recent issues) are among such freely accessible journals. For those physicians who have a university Athens system subscription, most of the other journals are also freely available. This is also true for the Cochrane Library and its systematic reviews.

Step 3: critical appraisal

Not all published articles are of high quality regarding the level of evidence. Critical appraisal is the process of evaluating individual studies to determine the level of evidence, which is an easy and efficient way to do this task.

Levels of evidence

The Oxford Centre for Evidence Based Medicine provides a free online table of levels of evidence (Table 4) [15]. With the guide of these tables, we can assign a level of evidence to each article in order to select the highest quality studies. By this strategy, there would be no need to read all articles regarding a clinical question. Many poor quality articles can be omitted in this step.

For each study, several questions have to be answered in order to assign a level of evidence. Several free appraisal sheets are available online for this purpose [16, 17]. These sheets contain two main parts:

1 — questions regarding the validity which can be found in the materials and methods section.

2 — the effect size or strength of the study, which is presented in the results section. The first section is the main part to be used for assigning the level of evidence [18].

Question 1: What was the spectrum of patients who underwent the test in question?

Whole spectrum of target disease should be covered in the study considering the severity and temporality (mild and severe disease, acute and chronic disease). Randomization would ensure avoiding selection bias. The characteristics of the patients (such as gender, age, ethnicity, etc) should be considered to assure that

Table 4. Oxford centre for evidence-based medicine levels of evidence for diagnosis studies (reproduced with permission)

Level	Diagnosis
1a	SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres
1b	Validating** cohort study with good†† reference standards; or CDR† tested within one clinical centre; Independent blind comparison of an appropriate spectrum of consecutive patients, all of whom have undergone both the diagnostic test and the reference standard
1c	Absolute SpPins and SnNouts†††
2a	SR (with homogeneity*) of level > 2 diagnostic studies
2b	Exploratory** cohort study with good†††reference standards; CDR† after derivation, or validated only on split-sample§ or databases; Independent blind comparison but either in non-consecutive patients or confined to a narrow spectrum of study patients (or both), all of whom have undergone both the diagnostic test and the reference standard; or a clinical decision rule not validated
3a	SR (with homogeneity*) of 3b and better studies
3b	Non-consecutive study; or without consistently applied reference standards
4	Case-control study, poor or non-independent reference standard
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

By homogeneity we mean a systematic review that is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity need be worrisome, and not all worrisome heterogeneity need be statistically significant. As noted above, studies displaying worrisome heterogeneity should be tagged with a "–" at the end of their designated level. **Validating studies test the quality of a specific diagnostic test, based on prior evidence. An exploratory study collects information and trawls the data (e.g. using regression analysis) to find which factors are "significant"; † Clinical decision rule. (These are algorithms or scoring systems which lead to a prognostic estimation or a diagnostic category); †† Good reference standards are independent of the test, and are applied blindly or objectively, applied to all patients. Poor reference standards are haphazardly applied, but still independent of the test. Use of a non-independent reference standard (where the test is included in the reference, or where the 'testing' affects the reference) implies a level 4 study; ††† An "Absolute SpPin" is a diagnostic finding, the Specificity of which is so high that a Positive result rules-in the diagnosis. An "Absolute SnNout" is a diagnostic finding, the Sensitivity of which is so high that a Negative result rules-out the diagnosis; § Split-sample validation is achieved by collecting all the information in a single tranche, then artificially dividing this into "derivation" and "validation" samples

the study is applicable to the patients on whom the test is going to be used [16, 19]. The eligibility criteria should also be defined meticulously [19].

A narrow spectrum of recruited patients would place a study in level 2 of evidence. For example, in a study by Ahmadihosseini et al., the application of ^{99m}Tc-MIBI in pulmonary tuberculosis diagnosis was evaluated [20]. A rapid review of the abstract of this article shows that a narrow spectrum of patients entered the study (only patients with active or inactive treated tuberculosis). A good spectrum of patients in this study would be all patients with respiratory symptoms that are indicative of tuberculosis.

Question 2: Was the reference standard the best available test?

The reference standard is the test used in a study to establish the diagnosis of the studied disease. Special care is needed to select the best available test for this purpose. Sputum culture for pulmonary tuberculosis and pulmonary angiography for pulmonary embolism are two examples in this regard. Sometimes, follow up instead of reference standard is used to find out if the patient has the disease or not. In this case, the period of follow up should be long enough considering the nature of the disease [16, 19]. In some clinical scenarios, there is no single best reference standard, and a combination of clinical findings and paraclinical tests should be used. For example, this strategy should be used for the diagnosis of recurrent lymphoma lesions [21]. A poor reference standard would place a study in level 4 of evidence.

For example, in a study by Fallahi et al. the value of ^{99m}Tc-MIBI for diagnosis of multiple myeloma lesions was evaluated. The reference standard in this study was a combination of clinical and paraclinical tests, which seems to be the best available reference standard [22].

An example of poor reference standard would be the detection rate of sentinel lymph node by lymphoscintigraphy instead of using axillary lymph node dissection (ALND) [23].

Question 3: Were the reference standard and the index test applied to all patients in the study regardless of the index test results?

The reference standard test should be used for all patients (regardless of index test results). The most common flaw in the diagnosis studies is failure to meet this criteria [18]. This would place a study in the level 3 of the evidence.

For example, in a study by Kostakoglu et al. [24] on ^{67}Ga diagnostic efficacy for diagnosis of active lymphoma in the residual masses after treatment of lymphoma, the authors obtained a sensitivity of 96% and a specificity of 80%. However, the gold standard of the study (biopsy) was not performed for all patients. Apparently, the positive result of ^{67}Ga was the main decision-maker for the application of biopsy in this study.

The period between reference standard and index test has to be considered since it should be short enough to assure that the target disease did not change between the reference standard and the index test [17]. For example, the time between a pulmonary embolism event and pulmonary angiography should not be too long, since the thrombolysis process can interfere in the diagnosis of pulmonary embolism.

Question 4: Was the comparison between the index test and reference standard blind and independent?

The interpreters of the index and reference standard tests should be blind to the results of each test in order to avoid expectation bias [19], which can place a study in level 4 of evidence.

Question 5: Were the index test and reference standard explained fully in the article?

Reliable duplication of the results of each study depends on this question. For example, the type of collimator, the type of gamma camera, etc. should be mentioned in the study.

The above-mentioned questions are very important to assign a level of evidence to each article. Many articles (with low validity) can be omitted at this point and not reviewed further. The remaining questions evaluate the results of the study.

Question 6: What were the results of the study?

For diagnostic tests, sensitivity, specificity, and positive and negative predictive values (PPV and NPV) should be considered and can usually be found in the results section of studies. Figure 1 shows a 2 × 2 chart for a dichotomous test. The definitions can be found in Table 5.

When the results of a test are of multilevel or continuous nature (such as gated SPECT indices), the cut-off value of the index test is the main determinant of the sensitivity and specificity. In these situations, there is a negative correlation between sensitivity and specificity. This is the basis of receiver operating characteristic (ROC) analysis. The ROC curve shows the sensitivity and specificity for different cut-off points of the test results. It is very important to con-

	Reference standard positive	Reference standard positive
Index test positive	True positive (a)	False positive (b)
Index test negative	False negative (c)	True negative (d)

Figure 1. 2 × 2 chart for a test with dichotomous results.

Table 5. Definitions of the characteristics of the test shown in Figure 1

Characteristic of the test	Definition	Formula
Sensitivity	How good is this test at detecting people with the disease?	a/(a + c)
Specificity	How good is this test at correctly excluding people without the disease?	d/(b + d)
Positive predictive value	What is the probability that a person with a positive test has the disease?	a/(a + b)
Negative predictive value	What is the probability that a person with a negative test does not have the disease?	d/(c + d)
Accuracy	What proportion of all tests yielded the correct result?	(a + d)/ /(a + b + c + d)

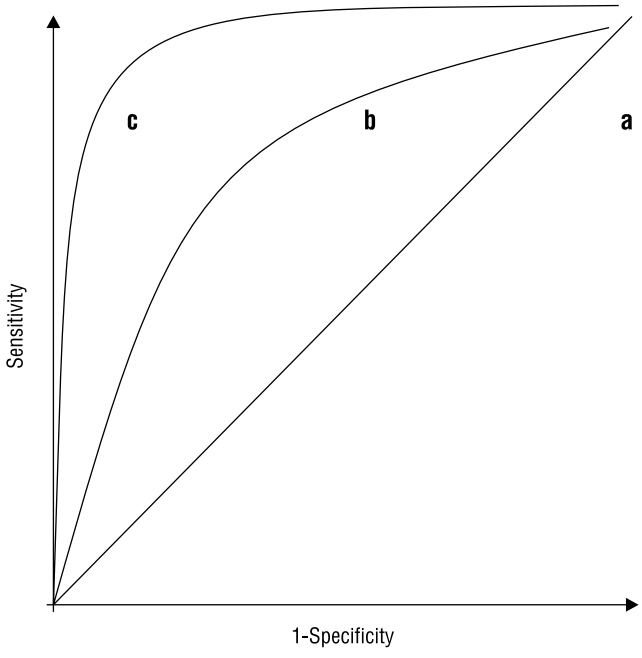

Figure 2. Receiver operating characteristic (ROC) curves of three different tests. Larger area under the curve for each test (location of the curve to the left and top) indicates better performance. For example the test "c" performs best, and test "a" performs worst.

sider the area under the curve in ROC analysis. If this area equals 0.5, the test is of no use. When this area approaches unity, the performance of the test increases (Figure 2).

Sensitivity and specificity are inherent properties of a test and are constant no matter what the prevalence of the disease in the tested population. In contrast, predictive values [both negative (NPV) and positive (PPV)] are dependent on the prevalence of the disease in the population. This is the cornerstone of the Bayes' Theorem, which is discussed further at the end of this article [5–6, 25].

Question 7: Were the confidence intervals mentioned for the test results?

In addition to p-values, sensitivity, and specificity, other important values should be mentioned for sufficient interpretation. The P-value is only a probability that an outcome has occurred by chance. It is definitely not a substitute for effect size. To appreciate the effect size of a study, the confidence intervals are very useful. A wide confidence interval range indicates a small sample size [26]. For example, the sensitivity of a test can be 80% in two different studies; however, the confidence intervals for this sensitivity can be 70–90% and 40–97%, respectively. The first study provides better evidence in this regard. Online calculators are freely available for confidence interval calculation [27].

A full explanation of this issue is beyond the scope of this article, but an excellent book written by Cohen addresses this issue in detail [28].

Step 4: applying the evidence to a particular patient

As mentioned before, the predictive values of a test are dependent on the prevalence of a disease in a society. This prevalence is usually called pre-test probability. The post-test probability is the refined probability of a disease when both pre-test probability and test results are considered together. For this purpose, likelihood ratios (LRs) are used, which are defined in Table 6. For a given pre-test probability (prevalence), pre-test odds can also be calculated.

Table 6. The definitions of likelihood ratios (LRs) for the test shown in Table 1

Characteristic of the test	Definition	Formula
Likelihood ratio of a positive test	What is the likelihood of a positive test result being found for a person with the disease compared to a person without it?	sensitivity/ /(l-specificity)
Likelihood ratio a negative test	What is the likelihood of a negative test being found for a person without the disease compared to a person with it?	(l-sensitivity)/ /specificityof

$$PRE\text{-}TEST\ ODDS = PRE\text{-}TEST\ PROBABILITY / (1 - PRE\text{-}TEST\ PROBABILITY)$$

These ratios (odds and likelihood ratios) can be combined together by multiplication [18]. No matter how many test are used to refine a probability, this method can be applied as follows:

$$POST\text{-}TEST\ ODDS = PRE\text{-}TEST\ ODDS \times LR\ of\ test\ 1 \times \times LR\ of\ test\ 2 \times LR\ of\ test\ 3...$$

Finally, the post-test probability of the disease can be calculated from the calculated odds:

$$POST\text{-}TEST\ PROBABILITY = POST\text{-}TEST\ ODDS / (POST\text{-}TEST\ ODDS + 1)$$

The concept of pre and post-test probability is shown in Figure 2 in another way.

Usually a threshold is set for the probability of each disease for treatment, above which treatment is justified. For example, this threshold for malignant bone tumours is very high (the diagnosis of malignant bone tumours should be almost definite (near 100%) to justify treatment). This is also the case for pulmonary embolism: the probability of the presence of pulmonary embolism should be more than 80% (high probability) to begin treatment.

The main application of a para-clinical test is refining the pre-test probability for a particular patient. If the post-test probability becomes higher than the treatment threshold, the treatment of the disease is justified and vice versa (Figure 3).

Other important issues to be addressed

For implementing the best found evidence in every day practice, cultural and religious issues should be considered. Some tests and procedures may not be culturally acceptable (no matter how good the performance is), and before requesting a test this important fact should be born in mind.

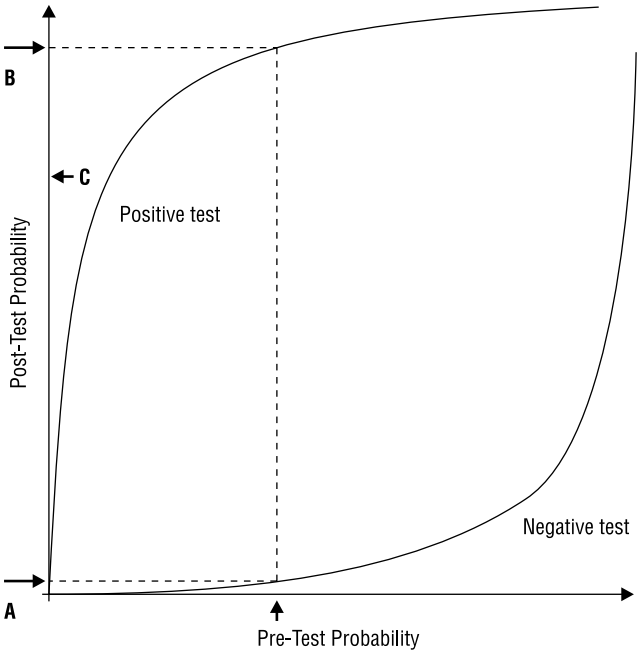

Figure 3. Graph showing the correlation between pre-test probability and both negative and positive results of a particular test. For a patient with a pre-defined pre-test probability (vertical arrow), the post-test probability after getting positive and negative results (A and B, respectively) would be those shown by horizontal arrows. The treatment threshold is shown by a white arrow (C). Since B > C, a positive result of this particular test would justify treatment. In contrast, A < C so a negative result of this test would not justify treatment.

Economical issues are another important aspect to be considered. With limited budget for health care, many expensive procedures and tests should not be ordered. This is beyond the scope of this review, but you can find a more detailed explanation elsewhere in the literature [29].

References

1. Wood BP. What's the evidence? Radiology 1999; 213: 635–637.
2. Evidence-Based Radiology Working Group. Evidence-based radiology: a new approach to the practice of radiology. Radiology 2001; 220: 566—575.
3. van Beek EJ, Malone DE. Evidence-based practice in radiology education: why and how should we teach it? Radiology 2007; 243: 633–640.
4. Del Mar C, Glasziou P, Mayer D. Teaching evidence based medicine [editorial]. BMJ 2004; 329: 989–990.
5. Alvarez Ruiz S, Cortés Hernández J, Rodeno Ortiz De Zárate E, Alonso Colmenares JI, Alcorta Armentia P. Evidence based medicine. Generalizations on the application to nuclear medicine. Part I. Rev Esp Med Nucl 2001; 20: 313–328.
6. Alvarez Ruiz S, Canut Blasco A, Rodeno Ortiz de Zárate E et al. Evidence based medicine. Application to nuclear medicine. Diagnostic slope. Part II. Rev Esp Med Nucl 2001; 20: 393–412.
7. 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.
8. Introduction. In: Sackett DL, Strauss SE, Richardson WS, Rosenberg W, Haynes RB. Evidence based medicine: how to practice and teach EBM. 2nd ed. Churchill Livingstone, Edinburgh 2000; 1–12.
9. Staunton M. Evidence-based radiology: steps 1 and 2 — asking answerable questions and searching for evidence. Radiology 2007; 242: 23–31.
10. Ely JW, Osheroff JA, Ebell MH et al. Analysis of questions asked by family doctors regarding patient care. BMJ 1999; 319: 358–361.
11. Ely JW, Osheroff J, Ebel M et al. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. BMJ 2002; 2: 265–268.
12. Haynes RB. Of studies, summaries, synopses, and systems: the "4S" evolution of services for finding current best evidence. Evid Based Nurs 2005; 8: 4–6.
13. The Cochrane Library Help. http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/cochranedemo.ppt. Accessed April 5, 2009.
14. Greenhalgh T. How to read a paper. The Medline database. BMJ 1997; 315: 180–183.
15. Levels of evidence. Oxford Centre for Evidence-Based Medicine Web site. http://www.cebm.net/index.aspx?o=1025.Accessed April 12, 2009.
16. Diagnostic Critical Appraisal Sheet. Oxford Centre for Evidence-Based Medicine Web site. http://www.cebm.net/index.aspx?o=1096. Accessed April 13, 2009.
17. Scottish Intercollegiate Guidelines Network (SIGN) Web site. http://www.sign.ac.uk/methodology/checklists.html. Accessed April 12, 2009.
18. Dodd JD. Evidence-based practice in radiology: Steps 3 and 4 Appraise and apply diagnostic radiology literature. Radiology 2007; 242: 342–354.
19. Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. BMJ 1997; 315: 540–543.
20. Ahmadihosseini H, Sadeghi R, Zakavi R, Dabbagh Kakhki VR, Haghighi Kakhki AR. Application of Technetium-99m-sestamibi in differentiation of active from inactive pulmonary tuberculosis using SPECT method. Nucl Med Commun 2008; 29: 690–694.
21. Hoda S. Role of nuclear medicine in detection and management of Hodgkin's disease and non-Hodgkin's lymphoma. Iran J Nucl Med 2002; 16-17: 17–25.
22. Fallahi B, Saghari M, Fard A et al. The value of 99mTc-MIBI whole body scintigraphy in active and in remission multiple myeloma. Hell J Nucl Med 2005; 8: 165–168.
23. Heuts E. Excision biopsy of breast lesions changes the pattern of lymphatic drainage. Br J Surg 2007; 94: 1573.
24. Kostakoglu L, Yeh SDJ, Portlock C et al. Validation of Gallium-68-citrate single-photon emission computed tomography in biopsy-confirmed residual Hodgkin's disease in the mediastinum. J Nucl Med 1992; 33: 345–350.
25. Diagnosis and screening. In: Sackett DL, Strauss SE, Richardson WS, Rosenberg W, Haynes RB. Evidence based medicine: how to practice and teach EBM. 2nd ed. Churchill Livingstone, Edinburgh 2000; 1–12.
26. Greenhalgh T. How to read a paper. Statistics for the non-statistician II: "Significant" relations and their pitfalls. BMJ 1997; 315: 422–425.
27. CAT maker software. Oxford Centre for Evidence-Based Medicine Web site. http://www.cebm.net/index.aspx?o=1216.Accessed April 13, 2009.
28. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Lawrence Erlbaum Associates, 1988.
29. Greenhalgh T. How to read a paper. Papers that tell you what things cost (economic analyses). BMJ 1997; 315: 596–599.