Evidence-Based Medicine in Rheumatology: How Does It Differ from Other Diseases?

Theodore Pincus and Hasan Yazici

The Absence of a "Gold Standard" in Rheumatic Diseases

The authors are now sufficiently senior to recall the early 1970s, at which time rheumatologists were considered elite members of the medical community in their zealous search for evidence in clinical care. Rheumatology fellows were using terms such as "sensitivity," "specificity," "true negatives" and "false positives" more than trainees in other fields. This emphasis may have resulted from an important difference in rheumatic diseases versus many other diseases – the absence of a single "gold standard" measure for diagnosis, prognosis, management and assessment of outcomes in each <u>individual</u> patient with a given diagnosis. Trainees in cardiology, endocrinology, nephrology and other fields had a lesser interest in complexities of clinical measures as they often had a definitive "gold standard," such as sustained elevated blood pressure in hypertension, sustained elevated glucose in diabetes mellitus, or a definitive biopsy in lymphoma, to guide clinical care.

The discovery in 1948 of rheumatoid factor [1] and the LE cell phenomenon [2] gave hope that a single gold standard biomarker would be available similarly for diagnosis, prognosis, management and assessment of outcomes in rheumatoid arthritis (RA) or systemic lupus erythematosus (SLE), the two most common inflammatory rheumatic diseases. However, despite extensive clinical research, that hope has not been met. Rheumatoid factor was described as present in 70 % of patients seen with RA in the initial report of Rose, Ragan *et al.* [1],

T. Pincus, MD (✉)
Division of Rheumatology, Rush University Medical Center, 1611 West Harrison Street, Chicago, IL 60612, USA

H. Yazici, MD (✉)
Division of Rheumatology, Department of Medicine, Cerrahpasa Medical Faculty, University of Istanbul, Istanbul, Turkey (retired)

virtually identical to 69 % in a recent meta-analysis [3]. Furthermore, rheumatoid factor is found in about 5–10 % of people in the general population [3], including patients with chronic infections and no apparent disease at all. Antibodies to citrullinated proteins (anti-CCP or ACPA) show increased specificity for RA, as they are seen in fewer than 5 % of individuals in the normal population; however, these antibodies are found in only 67 % of RA patients [3], quite comparable to rheumatoid factor.

Further biomarkers have been sought in RA based on the erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP). As with rheumatoid factor and ACPA, these measures are abnormal in the majority of patients. However, at this time, at least 40 % of patients do not show elevated values [4], although this proportion has declined from about 80 % in the early 80s to approximately 55 % in recent years [5], as RA patient clinical status has been improving [6]. The absence of a gold standard laboratory biomarker such as serum glucose, cholesterol, creatinine, hemoglobin or hemoglobin A1c, therefore, distinguishes rheumatic diseases from many chronic diseases for clinical trials, other clinical research and routine clinical care.

Pooled Indices as Quantitative Measures of Clinical Status in Rheumatic Diseases.

In the absence of "gold standard" laboratory tests or other quantitative biomarkers such as blood pressure or bone densitometry scores, pooled indices are required to assess quantitatively clinical status and responses to therapy in individual patients with a rheumatic diagnosis. The most successful pooled indices are seen in RA, based on a core data set of 7 measures: 3 recorded by a physician from a physical examination, *i.e.*, tender joint count, swollen joint count, and physician global estimate of status; 3 based on a patient self-report questionnaire – physical function, pain, and patient global estimate of status; and only 1 laboratory test, ESR or CRP [7]. Patients who may have many swollen joints and low pain levels, or a reciprocal pattern, are assessed according to an identical quantitative index. The core data set has been used for more than 2 decades and may be regarded as one of the major advances in rheumatology, prerequisite for the better status of patients at this time compared with previous decades [6].

The most prominent traditional index for RA has been the disease activity score (DAS) [8] and DAS28 [9], based on 4 measures: tender joint count, swollen joint count, ESR or CRP, and patient global estimate of status. The limitations of the DAS28 include a need for a laboratory test [ESR or CRP], which often is not available at the time of the visit, and is normal in up to 40 % of patients [4], and complex calculations, although easily accomplished at an excellent website. These limitations are overcome by the clinical disease activity index (CDAI) [10], which is simply a total of 4 measures: 28 tender joint count, 28 swollen joint count, and physician and patient global estimates 10 cm visual analog scales (VAS), total 0–76. An index of only the 3 patient self-report measures, known as routine assessment of patient index data (RAPID3), includes three 0–10 scales for physical function, pain and patient estimate of global status, total 0–30 [11]. Levels have been

established for high, moderate, low activity or severity of each index [12]; an index of only patient measures is not as specific to assess disease activity, since it might be sensitive to joint damage and chronic pain, but the other indices also are affected, though less so [13].

In analyses of clinical trials, essentially any 3 or 4 core data set measures will give very similar results, as was shown in analyses to establish remission criteria for RA [14]. Some rheumatologists support use of the simplest measure, RAPID3, as the patient does 95 % of the work and measurement involves the same single observer – the patient – at all visits. At the same time, other rheumatologists, particularly outside the USA, feel uncomfortable with only patient measures and include a CDAI or a DAS28.

Indices for Other Rheumatic Diseases

Indices exist for many other rheumatic diseases. In general, all include at least one measure from a physical examination and from patient self-report questionnaire, as well as a laboratory test. All are more complex than a "gold standard" measure. However, it also is possible that clinical decisions based on a gold standard measure may oversimplify what is needed for optimal patient care in chronic diseases. For example, functional status is as significant as ejection fraction to predict 3-year hospitalizations and deaths in congestive heart failure [15], CD4/CD8 ratios and other AIDS-specific measures to predict 3-year mortality in AIDS [16], and physiologic data and comorbidities to predict 1-year mortality in hospitalized elder patients [17]. Therefore, the importance of these measures may extend beyond rheumatology.

Some indices in rheumatology may be insensitive to clinical changes, which may account in part for some of the limitations in clinical trials. If an index includes, say, 10 measures, only 2 of which may change substantially and the others not at all, the index may indicate no change when an important clinical change has occurred in the 2 measures. Ironically, criteria for psychometric validation of indices based on statistical tools such as Cronbach's alpha and convergent validity generally may reduce sensitivity to change. Such sensitivity often is greatest with simple 10 cm visual analog scales (VAS). Nonetheless, it is essential to have an index for diseases in which certain clinical manifestations may vary widely and be prominent in some patients and absent in others, as noted for joint swelling and pain for RA.

Prominence of Patient History and Physical Examination in Clinical Decisions in Rheumatology

A survey was conducted in which 313 physicians, approximately half of whom were rheumatologists and half non-rheumatologists, estimated the relative importance of 5 elements of the clinical encounter – vital signs, patient history, physical examination, laboratory tests and ancillary studies (imaging, biopsy, endoscopy,

etc.) – in clinical decisions in 8 chronic diseases: congestive heart failure (CHF), diabetes mellitus, hypercholesterolemia, hypertension, lymphoma, pulmonary fibrosis, rheumatoid arthritis, and ulcerative colitis. The response options were 1–100 % in 5 equally divided intervals [18].

As expected, vital signs were most prominent in hypertension; laboratory tests were most prominent for diabetes and hyperlipidemia; and ancillary studies were most prominent for lymphoma, pulmonary fibrosis, ulcerative colitis, and congestive heart failure (Figure 1) [18]. RA was the only one of the 8 chronic conditions in which a patient history and physical examination accounted for more than 50 % of the information required for diagnosis and management (the total could be higher than 100 % due to "ties") [18]. These data provide evidence that the clinical encounter in rheumatology practice differs substantially from that in other subspecialties.

The results of this survey are reflected in the 7 items of the RA core data set, which includes 3 items from a patient questionnaire, 3 from a physical examination, and 1 laboratory test. A patient self-report questionnaire may be regarded as providing information for the patient history as quantitative data rather than narrative non-quantitative descriptions. A formal joint count may be regarded as providing information from the physical examination as quantitative data rather than narrative non-quantitative descriptions. The RA indices therefore reflect patient history and physical examination in contrast to gold standard biomarkers, which are most prominent in clinical decisions in many other chronic diseases.

Limitations of Laboratory Findings

As noted above, when rheumatoid factor was discovered in 1948, it was initially thought that this autoantibody might be both causative and diagnostic, as with antinuclear antibodies in 1960 for SLE, HLA B27 in ankylosing spondylitis, and mutant gene associations in FMF and MEFV [19]. However, the information from the laboratory is relatively limited in rheumatic diseases, compared to lab tests in other subspecialties of internal medicine, such as hemoglobin A1c or serum glucose. Of course, laboratory markers are important in groups and as clues to pathogenesis and development of treatments. For example, the development of biological therapy for RA may be traced directly to identification of rheumatoid factor with subsequent recognition of cytokines.

Laboratory markers are not positive in 30–50 % of all patients with RA [20]. Furthermore, they are "abnormal" (false positive) in some individuals in the normal population who have other diseases or no disease whatsoever, unlike measures such as sustained hypertension or elevated glucose over time.

There is value in calculating the sensitivity and specificity and predictive value of different tests, for the probability of a certain disease being present in a patient. However, the individual patient who may not have any positive tests but has pathognomonic clinical features of a disease has a 100 % probability of having the disease, regardless of the test results. A test that is positive in only 70 % of patients has limited utility in daily practice, although most rheumatologists are not aware of this

problem. It is sobering to remember that information from the laboratory in rheumatology is not pathognomic as in other diseases.

Limitations of Imaging

Structural changes are prominent in many rheumatic diseases, which might suggest an expectation that imaging would be most informative in diagnosis and management. Magnetic resonance imaging (MRI) and ultrasound certainly have improved sensitivity compared to plain radiographs. However, these new imaging modalities have not improved specificity. It is also worth remembering that the severe outcomes of RA such as work disability and premature death are predicted at far higher levels of significance by physical function on a patient questionnaire and by comorbidities than by hand radiographs [21].

Ironically, one possible limitation of studies to analyze radiographs as prognostic of severe outcomes may be that radiographic data are derived from the hand, whereas work disability and death are far more prominently influenced by large joints, particularly knees, but also hips and shoulders. For example, the initial series reported on mortality in RA indicated that 6 joints, 2 shoulders, 2 hips and 2 knees could predict mortality as effectively as all joints [22].

Furthermore, radiographic findings and clinical symptoms are often highly dissociated. For example, joint tenderness and radiographic findings have no correlation whatsoever [23]. Many people who may have 4+ osteoarthritis of the knee report no pain [24].

Limitations of Histopathology

Rheumatic diseases may include biopsies in an effort to establish or feel more secure about a given diagnosis. However, many findings have little tissue specificity, such as the synovitis in RA which can be seen in many forms of inflammatory arthritis. While tissue specificity is seen in immune complexes in the kidneys or dermoepidermal junction in SLE, uric acid crystals in synovial fluid or tophi in gout, giant cells and in the vessel wall in giant cell arteritis and in Takayasu disease, lymphocyte infiltration in salivary gland biopsies in Sjogren's syndrome, and bacilli in the intestinal wall in Whipple's disease; it is difficult to further expand the scope of histopathology in rheumatic disease.

Table 1 Comparison of "biomedical model" and "biopsychosocial model" of disease

	Biomedical Model	Biopsychosocial Model
Cause	Each disease has a single "cause"	Disease etiology is multifactorial: external pathogens, toxins, and internal host milieu, genes, behavior, social support
Diagnosis	Identified primarily through laboratory tests, radiographs, scans; information from patients of value primarily to suggest appropriate tests	A patient medical history provides 50%–90% of the information needed to make many, perhaps most, diagnoses
Prognosis	Also established most accurately based on information from high-technology sources, rather than from a patient	Information provided by a patient often is the most valuable data to establish a prognosis
Treatment	Involves only actions of health professionals, eg, medications, surgery	Must involve patient, family, social structure
Role of health professionals and patients in general health and disease outcomes	Health and disease outcomes are determined primarily by decisions and actions of health professionals	Health and outcomes of chronic diseases are determined as much by actions of individual patient as by health professionals

Rheumatology Assessment as a Challenge to the Biomedical Model

The major paradigm for advances in medical care over the last two centuries is a "biomedical model," in which clinical observations are translated into quantitative high-technology data from laboratory tests and ancillary studies. Early examples include bacterial cultures and quantitative laboratory measures of organ function (e.g., liver, kidney function tests), which can be used to guide care as "gold standard" measures for diagnosis, management, prognosis and assessment of outcomes in individual patients.

Over the last few decades there has been growing awareness that the traditional biomedical model, while spectacularly effective in acute disease and acute aspects of many chronic diseases, includes some significant limitations, particularly for chronic diseases. A classical statement was provided by George Engel, in a widely-read article in *Science* in 1978 [25]:

> "I contend that all medicine is in crisis and, further, that medicine's crisis derives from … adherence to a model of disease no longer adequate for its scientific tasks and social responsibilities…The biomedical model embraces both reductionism, the philosophic view that complex phenomena are ultimately derived from a single primary principle, and mind-body dualism, the doctrine that separates the mental from the somatic."

We may contrast the classical biomedical model with a biopsychosocial model of disease, which appears relevant to rheumatic diseases as complementary to a biomedical model (Table 1).

Some important differences between the biomedical and the biopsychosocial models are summarized in Table 1 [26].

The Biopsychosocial Model in Rheumatic and Other Diseases

Some essential elements to the biopsychosocial model (interestingly, without actually calling it as such) as it applies to rheumatology have been masterfully explained and discussed by H. Holman in his 1994 article "Thought Barriers to Understanding Rheumatic Diseases" [27]. Holman asserted that the main problem with the practice and the science of rheumatology is that "the prevailing conceptual base of our investigation is incommensurate with the rheumatic disease problems which we confront." He gives two main reasons:

1. While most rheumatologic diseases are chronic, the traditional medical teaching put the emphasis on acute pathology.
2. There is the prevailing notion of a "single cause" for a single disease. This, of course, has its roots in medicine's spectacular success in handling the infectious diseases. This reductionism is common in research, where we tend to overlook interactive biological pathways, common in many of our diseases.

As we note, Holman does not give a formal reference to the biopsychosocial model in this article. Perhaps he might have decided to underplay the psychological and social components of said model.

Nonetheless, the Vernon Riley experiment he relates in detail in this paper actually provides a brilliant example of the inadequacy of the biomedical model. The experiment concerned breast cancer in C3H mice [28]. This cancer, seen in C3H mice, is both genetic and environmental. The tumor appears around 1 year of age only among those mice that have been infected with a specific virus during suckling. All the cross-experiments of the biomedical dictum confirm these genetic and environmental components.

Riley introduced a third dimension, a psychosocial dimension if you will, to this model. He randomized C3H female offspring into 2 groups: one under usual experimental conditions of crowded cages and frequent blood samples, and the other in spacious cages and little if any bleeding. The outcome was that the latter group developed the expected tumor a median of almost 200 days later.

Limitations of EBM as Randomized Trials in Application to Clinical Care

The randomized controlled clinical trial may be considered a development in the tradition of the biomedical model. It is designed to mimic a laboratory experiment, in isolating a single variable that tests therapy while keeping all the other variables constant [29]. The clinical trial is most successful in acute infectious disease in

which the outcome may be known within a week or two. A trial becomes progressively more limited over time in chronic diseases, as discussed in detail in the chapter concerning limitations of randomized controlled clinical trials (see chapter "Limitations of traditional randomized controlled clinical trials in rheumatology"). Nonetheless, in these introductory comments, we recognize that even proponents of evidence-based medicine (EBM) as clinical-based trials recognize some limitations in application of results to clinical care.

Any evidence, in the last analysis, can be considered as a tool to convince either oneself or somebody else of the strength of a "truth" under consideration. More simply, evidence is what makes an object or a concept "evident" to us or to others. Verbose as it is, this definition – unlike its standard dictionary versions – has the advantage of emphasizing that the quality of evidence is quite dependent on *a)* who is to be convinced and *b)* the circumstances under which "the convincing" takes place. To convince your 3-year-old child that his plate is hot, you would never resort to the most direct evidence, the "gold standard" as the jargon goes, that he should touch it briefly and see for himself. However, for his mother, and if you are occasionally brave, this direct evidence might be used!

One example of some limitations of application of randomized controlled clinical trials to clinical care recognized by a leading proponent of EBM may be found in the introduction to the book *Philosophy of Evidence Based Medicine* by Dr. Jeremy Howick [30]. Here the author takes a direct quote from Dr. Chalmers, who saw many children with measles who also were malnourished and in general poor health, while working as a young doctor in a Palestinian refugee camp in the Gaza Strip. Unless there was clear evidence of superinfection he refrained from prescribing antibiotics, as he had been taught in medical school.

However, mortality among his patients was considerably higher than among those of his Palestinian colleague who routinely prescribed prophylactic antibiotics. Chalmers observed: "This clinical impression was very sobering. It made me wonder whether what I had been taught at medical school might have been lethally wrong, at least in the circumstances in which I was working, and precipitated a now incurable 'septicemia' about authoritarian therapeutic prescriptions and prescriptions unsupported by trustworthy empirical evidence" [30].

The catch line here is, of course, "…in the circumstances in which I was working…." The "evidence" about not starting prophylactic antibiotics in managing measles might have been true for the more fortunate locations where Chalmers' professors resided, but not for the Gaza strip. This business of "to whom" and "under which circumstances" – or the *external validity* (in the jargon) – is an important and often neglected aspect of EBM.

All evidence can either be direct or indirect. Some good examples of direct evidence as it concerns our discipline are: a wedge-shaped crushed vertebral body on a radiograph in osteoporosis; finding sodium mono urate crystals in the synovial fluid (making a diagnosis of gout); colchicine preventing attacks of familial Mediterranean fever (management); or anti-Ro antibodies sitting in the cardiac conduction system causing heart block in neonates (understanding disease mechanisms). Examples of such similar direct evidence from other disciplines are tell-tale EKG findings in a myocardial infarction, massive proteinuria in nephrotic syndrome, or antibodies to

acetylcholine receptors in myasthenia gravis. Much more common in our field is indirect evidence, as seen in the biopsychosocial model (Table 1). Here our young discipline differs substantially from other specialties, as noted above.

EBM as a Long-Standing Tradition

EBM surely did not surely begin abruptly in 1992 when it was publicly announced, nor did the Enlightenment and Industrialization bypass evidence in medicine. As early as the 17th century, Francis Bacon, the scientist and the philosopher, severely criticized Hippocrates for the anecdotal nature of his (as the name openly says) "aphorisms" [31]. A century later, the eminent English physician Francis Clifton made a strong plea for a meticulous tabulation of disease occurrence in English towns by sex, race, age and the type of illness [31]. The 19th century gave us much more objective tools for observation like the stethoscope and the microscope. Again in the 19th century, mankind began to benefit from scientifically powerful medicines for prevention: the smallpox vaccine, and antisera to treat diphtheria and tetanus. Medicine became much more scientific in the 20th century, not only with new and effective drugs and vaccines, but with spectacular advances in imaging and surgery. Moreover arithmetic, statistics, probability and randomness began to be taken seriously, discussed and also required by the physician practitioner and the medical scientist alike.

In 1948 the first properly randomized clinical trial [32] was conducted, and showed that the new drug streptomycin was superior in treating patients with tuberculosis compared to the available "standard of care," bed rest. Many other, similarly well-conducted trials in many other diseases followed. So what was wrong with prior EBM that made its founders declare the *new* EBM in 1992?

Possibly at least four, and somewhat related, reasons were behind the emergence of the new EBM as based on randomized controlled clinical trials.

The first was that the new EBM advocates wanted to give to the proponents of unconventional remedies one additional blow. This was surely timely, especially in the light of ever-rising medical costs in the setting of limited resources. Nothing more needs to be said here.

A second important issue behind the emergence of the new EBM was that, although the science of medicine had progressed substantially, this progression was often not translated into usual clinical care. In other words, the application was not commensurate with the level of science.

A third driving force may have been the relative inability of the biomedical model to address common, as well as less common, ills – as aptly exemplified both in Dr. Chalmers' story in the Gaza strip and in Riley's rats. It was that the traditional *science* of medicine fell short of explaining our ills and how to prevent or handle them to our satisfaction, and this is surely related to the second issue just discussed. However, when EBM is regarded as exclusively based on randomized controlled clinical trials, it does not give much headway to the biopsychosocial model. It mimics a laboratory experiment with a reductionist focus on a single variable,

attempting to control all other variables through randomization. Instead it does something else and this, perhaps, leads us to the fourth issue.

The new EBM with its authors, journals, books, governments and surely the drug control agencies like FDA considered that the medical field and profession needed many degrees and levels of central control – and, as in the first issue, the limited money/resource concern was the main reason. To put it another way, for the new EBM what was more important was the correct implementation of existing science rather than promoting science, and the monetary concern was also dominant. One effort toward control was placement of meta-analysis at the top of the hierarchy of evidence-based medicine (Figure 2) [28]. However, a meta-analysis is only as informative as its component clinical trials, and limitations of clinical trials such as patient selection, short-term time frame in chronic diseases, and reporting of data only in groups, may render a meta-analysis less accurate concerning clinical care than observational studies.

It must also be brought up here that this justified concern for money led to a central control that, ironically, kindled more money and resource problems – after the drug industry began to use this central control, in many instances, to their financial interests [33, 34]. It is as if medicine has learned little from the economists and business administrators that central control of business, in the long run, almost always takes a bigger chunk out of public money than private enterprise.

Over the last few years, some of the limitations of regarding evidence-based medicine only as clinical trials, invariably superior to other sources of clinical evidence, have gained increasing recognition. A more up-to-date view of "evidence-based medicine" is expressed by the Oxford Centre for Evidence-Based Medicine [35]: "While they are simple and easy to use, early hierarchies that placed randomized trials categorically above observational studies were criticized [28] for being simplistic [36]. In some cases, observational studies give us the 'best' evidence [28]. For example, there is a growing recognition that observational studies – even case-series [37] *and anecdotes* [38] can sometimes provide definitive evidence." Nevertheless, the principles of "evidence-based medicine" continue to evolve, hopefully leading to improved patient care and outcomes.

References

1. Rose HM, Ragan C, Pearce E, Lipman MO. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. Proc Soc Exp Biol Med 1948; 68:1–6.
2. Hargraves MM, Richmond H, Morton R. Presentation of two bone marrow elements: The "tart" cell and "L.E." cell. Proc Staff Meet Mayo Clin 1948; 23:25–8.
3. Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S, et al. Meta-analysis: Diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. Ann Intern Med 2007; 146(11):797–808.
4. Sokka T, Pincus T. Erythrocyte sedimentation rate, C-reactive protein, or rheumatoid factor are normal at presentation in 35 %-45% of patients with rheumatoid arthritis seen between 1980 and 2004: analyses from Finland and the United States. J Rheumatol 2009; 36(7):1387–90.

5. Abelson B, Sokka T, Pincus T. Declines in erythrocyte sedimentation rates in patients with rheumatoid arthritis over the second half of the 20th century. J Rheumatol 2009; 36(8):1596–9.

6. Pincus T, Sokka T, Kautiainen H. Patients seen for standard rheumatoid arthritis care have significantly better articular, radiographic, laboratory, and functional status in 2000 than in 1985. Arthritis Rheum 2005; 52:1009–19.

7. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993; 36:729–40.

8. van der Heijde DMFM, van't Hof M, van Riel PLCM, van de Putte LBA. Development of a disease activity score based on judgment in clinical practice by rheumatologists. J Rheumatol 1993; 20:579–81.

9. Prevoo MLL, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LBA, van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts: Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995; 38:44–8.

10. Aletaha D, Smolen J. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol 2005; 23:S100-S108.

11. Pincus T, Swearingen CJ, Bergman MJ, Colglazier CL, Kaell A, Kunath A, et al. RAPID3 on an MDHAQ is correlated significantly with activity levels of DAS28 and CDAI, but scored in 5 versus more than 90 seconds. Arthritis Care Res 2010; 62(2):181–9.

12. Pincus T, Hines P, Bergman MJ, Yazici Y, Rosenblatt LC, Maclean R. Proposed severity and response criteria for Routine Assessment of Patient Index Data (RAPID3): results for categories of disease activity and response criteria in abatacept clinical trials. J Rheumatol 2011; 38(12):2565–71.

13. Pincus T, Yazici Y, Bergman MJ. RAPID3, an index to assess and monitor patients with rheumatoid arthritis, without formal joint counts: similar results to DAS28 and CDAI in clinical trials and clinical care. Rheum Dis Clin North Am 2009; 35(4):773–8.

14. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum 2011; 63(3):573–86.

15. Konstam V, Salem D, Pouleur H, Kostis J, Gorkin L, Shumaker S, et al. Baseline quality of life as a predictor of mortality and hospitalization in 5,025 patients with congestive heart failure. Am J Cardiol 1996; 78(8):890–5.

16. Justice AC, Aiken LH, Smith HL, Turner BJ. The role of functional status in predicating inpatient mortality with AIDS: A comparison with current predictors. J Clin Epidemiol 1996; 49:193–201.

17. Covinsky KE, Justice AC, Rosenthal GE, Palmer RM, Landefeld S. Measuring prognosis and case mix in hospitalized elders. The importance of functional status. J Gen Intern Med 1997; 12:203–8.

18. Castrejón I, McCollum L, Durusu Tanriover M, Pincus T. Importance of patient history and physical examination in rheumatoid arthritis compared to other chronic diseases: Results of a physician survey. Arthritis Care Res 2012; 64(8):1250–5.

19. Esen F, Celik A, Yazici H. The use of diseased control groups in genetic association studies. Clin Exp Rheumatol 2009; 27(2 Suppl 53):S4-S5.

20. Pincus T, Yazici Y, Sokka T. Complexities in assessment of rheumatoid arthritis: absence of a single gold standard measure. Rheum Dis Clin North Am 2009; 35(4):687–97.

21. Pincus T, Bergman MJ, Maclean R, Yazici Y. Complex measures and indices for clinical research compared with simple patient questionnaires to assess function, pain, and global estimates as rheumatology "vital signs" for usual clinical care. Rheum Dis Clin North Am 2009; 35(4):779–86.

22. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. Ann Intern Med 1994; 120:26–34.
23. Fuchs HA, Callahan LF, Kaye JJ, Brooks RH, Nance EP, Pincus T. Radiographic and joint count findings of the hand in rheumatoid arthritis: Related and unrelated findings. Arthritis Rheum 1988; 31:44–51.
24. Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol 2000; 27:1513–7.
25. Engel GL. The biopsychosocial model and the education of health professionals. Ann N Y Acad Sci 1978; 310:169–81.
26. McCollum L, Pincus T. A biopsychosocial model to complement a biomedical model: patient questionnaire data and socioeconomic status usually are more significant than laboratory tests and imaging studies in prognosis of rheumatoid arthritis. Rheum Dis Clin North Am 2009; 35(4):699–712.
27. Holman HR. Thought barriers to understanding rheumatic diseases. Arthritis Rheum 1994; 37:1565–72.2011.
28. Riley V. Mouse mammary tumors: alteration of incidence as apparent function of stress. Science 1975;189(4201):465–7.
29. Feinstein AR. An additional basic science for clinical medicine: II. The limitations of randomized trials. Ann Intern Med 1983; 99:544–50.
30. Howick J. The Philosophy of Evidence-Based Medicine. Oxford: Wiley-Blackwell;
31. Trohler U. An early 18th-century proposal for improving medicine by tabulating and analysing practice. J R Soc Med 2010; 103(9):379–80.
32. MRC Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. Br Med J 1948; 2(4582):769–82.
33. Hickey S, Roberts H. Tarnished Gold: The Sickness of Evidence Based Medicine. CreateSpace; 2011.
34. Stamatakis E, Weiler R, Ioannidis JP. Undue industry influences that distort healthcare research, strategy, expenditure and practice: a review. Eur J Clin Invest 2013; 43(5):469–75.
35. OCEBM Levels of Evidence Working Group. The Oxford 2011 Levels of Evidence. http://www.cebm.net/index.aspx?o=5653. 2011. Oxford U.K., Oxford Centre for Evidence-Based Medicine
36. Smith GCS, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 2003; 327:1459–61.
37. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. BMJ 2007; 334(7589):349–51.
38. Aronson JK, Hauben M. Anecdotes that provide definitive evidence. BMJ 2006; 333(7581):1267–9.