4-2012

Evidence-Based Medicine versus Personalized Medicine: Are They Enemies?

Jose de Leon
University of Kentucky, jdeleon@uky.edu

Evidence-Based Medicine versus Personalized Medicine: Are They Enemies?

Notes/Citation Information
Published in *Journal of Clinical Psychopharmacology*, v. 32, issue 2, p 153-164.

This is a non-final version of an article published in final form in *Journal of Clinical Psychopharmacology*. April 2012, Volume 32, Issue 2, p 153-164. doi: 10.1097/JCP.0b013e3182491383

Digital Object Identifier (DOI)
http://dx.doi.org/10.1097/JCP.0b013e3182491383

This is a non-final version of an article published in final form in *Journal of Clinical Psychopharmacology*. April 2012, Volume 32, Issue 2, p 153-164. doi: 10.1097/JCP.0b013e3182491383

Guest Editorial

Evidence-based medicine versus personalized medicine: are they enemies?

Jose de Leon, M.D.*****

*University of Kentucky Mental Health Research Center at Eastern State Hospital, Lexington, KY, and Psychiatry and Neurosciences Research Group (CTS-549), Institute of Neurosciences, University of Granada, Granada, Spain.

Corresponding author: Jose de Leon, M.D., UK Mental Health Research Center at Eastern State Hospital, 627 West Fourth St., Lexington, KY 40508. Phone (859) 246-7563. Fax (859) 246-7019.
e-mail: jdeleon@uky.edu

Acknowledgments: No commercial organizations had any role in the writing of this paper for publication. Dr. de Leon personally develops his presentations for lecturing, has never lectured using any pharmaceutical or pharmacogenetic company presentations, and has never been a consultant for pharmacogenetic or pharmaceutical companies. In the past, Dr. de Leon received researcher-initiated grants from Eli Lilly (one ended in 2003 and the other, as co-investigator, ended in 2007); from Roche Molecular Systems, Inc. (ended in 2007); and, in a collaboration with Genomas, Inc., from the NIH Small Business Innovation Research program (ended in 2010). He has been on the advisory boards of Bristol-Myers Squibb (2003/04) and AstraZeneca (2003). Roche Molecular Systems supported one of his educational presentations, which was published in a peer-reviewed journal (2005). His lectures were supported once by Sandoz (1997), twice by Lundbeck (1999 and 1999), twice by Pfizer (2001 and 2001), three times by Eli Lilly (2003, 2006, and 2006), twice by Janssen (2000 and 2006), once by Bristol-Myers Squibb (2006), and seven times by Roche Molecular Systems, Inc. (once in 2005 and six times in 2006). The author thanks Lorraine Maw, M.A., for editorial assistance. The author is grateful to the editor and reviewers who "forced" him to further review the literature, sharpen his thinking and asked him to provide solutions for the problem.

Running title: evidence-based versus personalized medicine

Key words: evidence-based medicine; individualized medicine; personalized medicine; personalized prescription; pharmacogenetics; pharmacogenomics; statistical heterogeneity.

Evidence-based medicine (EBM) and personalized medicine (PM) are fashionable concepts frequently endorsed by medical and pharmacological journals. Everyone appears to agree that both should be implemented in the real world of clinical practice as soon as possible. The author believes that it will not be easy to implement both simultaneously since their approaches collide. While the EBM approach focuses on using randomized clinical trials (RCTs) to establish the best treatment for the average patient and ignores the outliers, PM focuses on the outliers (Table 1).

To appreciate this, one must reckon with the statistical concept of heterogeneity, which seems widely ignored. Feinstein[1] (both a physician and a scientist with mathematical training) pointed out that statistical heterogeneity is largely overlooked by statistics textbooks. In the view of the author[2] it is not surprising that statistical heterogeneity is ignored, in that clinicians have no idea what statistical heterogeneity means; they just want their statistician to analyze their data and produce the right answer. Secondly, statisticians receiving data from clinicians do not want to raise the possibility that the clinician's data is heterogeneous since then it could not be analyzed with the usual statistical methods or at least could not be analyzed collectively. The author believes that when dealing with drug response, EBM assumes that the "reality" of the drug response is statistically homogeneous; thus, an average response represents individuals quite well. PM, on the other hand, assumes that reality is not homogeneous since the average response is not representative of a sample of individuals, and drug response is a statistically heterogeneous phenomenon (Table 1).

How is this misunderstanding possible? Is EBM coming from Mars and PM from Venus? Yes. These two concepts have completely different origins; EBM was developed in the context of medical education and PM in the context of clinical pharmacology. Moreover, the battle between EBM versus PM may just be another skirmish in the larger ongoing wars between those who think medicine is a science and those who think it is an art.[3,4] The conflict can also be exemplified by other battles such as those between defenders of: 1) empirical observation versus mechanistic disease models,[5] 2) probabilistic and

empirical thinking versus deterministic and explanatory thinking,[6] 3) evaluation of interventions versus discovery and explanation,[7] and 4) public health versus individual patient health.[8,9]

This article briefly describes: 1) the history of EBM, 2) the recent realization that statistical heterogeneity may be an obstacle in EBM's attempts to summarize medical practice, 3) the history of PM, 4) the author's opinion that the level of heterogeneity or the lack thereof is that which dictates whether or not PM can easily coexist with EBM, 5) the history of attempts to explore drug response heterogeneity in psychiatry, and 6) proposed next steps for moving beyond this stalemate between EBM and PM.

History of EBM

The developers of EBM, Sackett and coworkers,[10] proposed that EBM had its origins in the first attempts to use statistics in clinical practice in 19th century France. EBM has been compared in particular to the "Medical Observation" developed by Pierre Charles Alexandre Louis in Paris during the 1830s and 1840s, who crusaded against the use of bloodletting in pulmonary infections, using the observational method and numerical calculations.[5,11] Gerber and coworkers[12] argued that EBM is fundamentally different from the mere attempts to introduce statistics in medicine in 19th century France, since its fundamental innovation is that EBM relies on and enhances a more equal relationship between physicians rather than the historical tradition of relying on experts. This appears to agree with one of Sackett's lesser known articles[13] in which he stressed that, when compared with the traditional method of educating physicians, EBM "puts a much lower value on authority".

In spite of these disagreements about antecedents, everyone agrees that EBM was mainly developed at McMaster University in Canada in the 1980s and 1990s[14-16] by Guyatt and Sackett, among others. Sackett contributed to EBM's dissemination by moving to Oxford University.[14] EBM became a mainstream concept in medicine after the publication of two articles, one in *JAMA*[16] in 1992 and one in the *British Medical Journal*[10] in 1996. According to Vandenbrucke[17], however, the *British Medical Journal*, *Annals of Internal Medicine*, and *JAMA* have greatly championed EBM, whereas *The Lancet* and the *New England Journal of Medicine* have kept some distance.

In a recent review updating the progress of EBM, Guyatt and coworkers[18] gave credit to Cochrane, Sackett and Feinstein for laying its foundations. They credited the Scottish epidemiologist Archie Cochrane for insisting that the clinical disciplines should summarize evidence concerning their practices.[19] They acknowledged their colleague Sackett for developing the teaching innovations. They also credited Feinstein for defining the principles of quantitative clinical reasoning. It is interesting that Feinstein was recognized since he was critical of EBM until his death and helped to develop a patient-centered approach.[20]

According to Sackett and Rosenberg,[21] the rapid growth of randomized clinical trials, the slow pace of updating textbooks, the lack of time physicians have for keeping up with journals, and the lack of efficacy of continual medical education in improving clinical competence explains the need for the development of EBM. According to Woolf,[22] the main historical factor contributing to the introduction of EBM was the wild variation in medical practice for the same types of patients, including overuse (and sometime underuse) of services leading to increased costs, a major issue for Cochrane.[20] Sackett and coworkers[10] defined EBM as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research." Most would agree that the heart of EBM is the reliance on RCTs as the best alternative for guiding medical knowledge. [23] More recently, Reilly[24] very wisely has acknowledged that EBM is three different things: 1) a scientific hypothesis, 2) an ever-evolving body of evidence, and 3) an idealized way of practicing medicine.

In the opinion of this writer, EBM is a definitively a departure from the prior 2,500 years in the sense that traditionally the study of medicine was primarily based on mentorship with a more experienced physician, ideally somebody who was an "expert". The traditional approach in learning medicine was that of rotating with a mentor who taught the student physician the art of medicine; then the physician practiced by himself and acquired experience as he or she made his or her own mistakes (sometimes lethal

to the patients) and, in the case of a few, became a physician mentor. Thus, using the traditional way of educating physicians, the older the physician, the more wise and experienced he or she was supposed to be.[26] The EBM approach has inverted this process since older physicians tend to be less experienced with the updates provided by the EBM approach.[26]

EBM can be seen as the culmination of the introduction of the scientific method in medicine[27] since, during the last 500 years, greater scientific knowledge has progressively been introduced in the mentoring of physicians. In the 20th century, the development of the RCT approach and its progressive adoption by government drug agencies for marketing drugs has made the knowledge gained from RCT available and, more importantly, has led to the combination of available RCT in the so-called meta-analyses. The realization that RCT and meta-analyses should be the cornerstone for medical decisions and medical education has led to EBM. According to Feinstein,[27] this new "faith" is expanded mainly by epidemiologists, while physicians usually defer to epidemiologists or other experts, since they rarely master the "secretive" art of summarizing average drug responses using meta-analytic techniques. The believers in EBM rarely recognize a major practical problem: the RCT which are used to test drug efficacy and gain Food and Drug Administration (FDA) approval for marketing usually deal with short-term drug response in otherwise healthy and uncomplicated patients who are also willing to enter RCT. Physicians, however, often deal with chronically ill patients who usually take multiple medications and can be uncooperative in taking medications. Thus, physicians are frequently interested in drug response after many months or years of treatment in all types of patients.[28] More recently, an attempt has been made to increase the representativeness of RCT by conducting practical or pragmatic trials[29] that focus on effectiveness rather than efficacy, but there are very few or no pragmatic clinical trials for most medical problems. The author works in an acute hospital for patients with severe mental illness; it has 2,000 admissions/year and nearly all its patients would be excluded from pragmatic trials due to the risk of suicide attempts (half of the patients have previously attempted suicide) and/or the inability/unwillingness to sign a complex consent form due to severe cases of psychosis, mania or depression.[30] Thus, "current

EBM" does not include well-controlled studies for the types of very sick patients treated in his hospital. Knotterus and Dinant[31] wisely stated that "medicine based evidence is a prerequisite for EBM" and that "future research methods must find ways of accommodating clinical reality, not ignoring it". The sad truth is that the more difficult the patient, the less evidence is available to treat him or her. It is very unlikely that RCTs will ever be conducted for the types of patients seen in the author's hospital since these patients combine severe mental illnesses with extensive history of suicidality, high levels of co-morbid substance abuse, and high levels of medical co-morbidity.[32]

Statistical heterogeneity and EBM

Although in 1997 Feinstein and Horwitz[33] emphasized the problem of lack of homogeneity in RCTs, only recently have statisticians[34,35] started focusing on the issue that the average patient may not be a good representative of all patients. There is no general agreement on how to best deal with this complex problem. The best and most sophisticated way is through stratification,[34] separating patients according to a characteristic that has a crucial effect on drug response and then randomizing the groups separately. The number of RCTs utilizing this courageous approach[36] may increase in the future but it is likely to be a slow process if the RCTs are designed by drug companies; they may not be interested in taking risks since stratification decreases statistical power and therefore may require larger recruitment samples. When heterogeneity is not considered *a priori* and resolved by a design using stratification, it is only possible to deal with it *a posteriori*, using subgroup statistical analyses to test for heterogeneity in treatment effects;[35] which is a rather controversial subject among statisticians.

Regarding *a posteriori* classification, a statistician with long experience in working on industry RCTs, Senn[37-41] has provided a comprehensive critique on heterogeneity. He criticizes EBM's defenders by reminding them that RCTs only provides information on averages and that the calculation of an average number, such as the number needed to treat, assumes that all patients benefit equally;[38,41] and he incorporates the thinking from a PM expert in his solutions.[42,43] Senn reminds[37-41] us that a RCT has three main sources of error variability besides the differences between treatments (the average differences

between treatments over all randomizations). They include: 1) between-patient variability (the average differences between patients), 2) patient-by-treatment interaction (the extent to which difference between treatment differs from one patient to another); and 3) within-patient error (the variability shown from treatment period to treatment period when the same patient is given the same treatment).[38] Senn says that parallel RCTs cannot distinguish between the three types of error variability and cross-over RCTs (in which each patient only receives the treatment and control conditions) cannot distinguish between patient-by- treatment interaction and within-patient error.[38] He notes that differences in drug response may include genetic variability but genetic variability cannot exceed individual variability shown by patient-by- treatment interaction in RCTs; [38,41] In addition, the same data cannot be used to test for differences between groups in RCTs and then used to classify patients retrospectively.[37,41] Senn also points out that the assumption that pharmacogenetics is important in drug response is largely untested.[41,42]

History of PM

Physicians have traditionally practiced PM in their attempts to decide the best treatment for each of their patients.[44] However, they were not using the term "PM" and they were probably basing their traditional PM on their subjective preferences or their understanding of what was best for the patient, and not on scientific knowledge. The current use of the term "PM" can be explained by advances in clinical pharmacology.[45-48] Meyer[45] has presented the most comprehensive historical review of these pharmacological advances; here only a brief overview of how these advances have led to the development of PM is presented. At the beginning of the 20th century, around the time the word "genetic" was first used, Garrod[49] hypothesized that there was a "chemical individuality" and "those idiosyncrasies with regard to drug" response may be explained by evolution. In the 1950s[45-48] pharmacologists observed that some severe adverse drug reactions (ADRs) only occurred in a small number of patients. They included hemolytic anemia after the use of primoquine, apnea with succynilcholine, and peripheral neuropathy with isoniazid. In 1959, Vogel[50] coined the term "pharmacogenetics". In the 1960s-70s[45-48] two phenotypes, poor metabolizers and extensive metabolizers, were described for several drugs including

tricyclic antidepressants (TCAs), debrisoquine, spartein and mephenytoin. In the 1980s-90s,[45-48] 1) the various genes associated with the cytochrome P450 (CYP) isoenzymes were discovered, 2) the allele variants associated with poor metabolism were described, 3) the ultrarapid metabolizers were described in patients taking TCAs and were associated with a gene duplication or multiplication, and 4) other pharmacokinetic genes began to be identified.

In the late 1990s, advances in genetics allowed parallel genetic testing[51] (or testing for multiple genes at the same time) through the use of the so-called DNA microarrays. Thus, the new term "pharmacogenomics" was coined for the study of all genes that may influence drug response.[52] In 1997 the development of DNA microarray technology led *Science*[53] to define "personalized prescription" as "tailoring drugs to a patient's genetic makeup" and to predict that personalized prescription would "soon" reach clinical practice. The race for the human genome and the political decision in 2000 by President Clinton that the race was over[54] led to further hype in lay[55] and scientific journals[56] about the promises of genetics to personalize medicine. The generalized use of personalized prescription would begin in 2015, according to *Time* magazine,[55] or 2020, according to *JAMA*.[56] The first decade of the 21st century has led to greater acknowledgement of difficulties and complexities at three levels: 1) genetic science, 2) the use of genetic DNA microarrays in clinical practice, and 3) the development of other microarrays. The complexity of the increasing genetic knowledge led to the realization that the function of approximately one-third of human genes is unknown and that duplications and other variants, the so-called copy number variations (CNV) and epigenetics, may influence drug response.[57] The introduction of the first DNA microarray approved by the Food and Drug Administration (FDA) for pharmacogenetic testing in 2006[9,58] has led to very limited clinical use for reasons that are complex.[59] The further development of DNA microarray technology has allowed the testing of multiple other products besides DNA, including proteins, RNAs, and lipids, and to the development of new diagnostic branches such as "protenomics", "transcriptomics" and "metabonomics". All of these new techniques called "omics"[60] have brought

additional expectations about their use in PM and also the rapid realization of the complexities of the scientific and practical issues involved in taking these microarrays to clinical practice.[61,62]

In the time frame of 3-5 years these technologies, microarrays and genotyping methods continue to advance and make prior versions obsolete by becoming more powerful (providing more and more data – up to millions of pieces of data) and progressively cheaper. To make these advances more suitable for clinical practice, "better" and "cheaper" is important, but advances in making these tests easier to use are also crucial. In that sense, the introduction of nanotechnology in genotyping tests may be an important step.[63] The FDA approved a warfarin pharmacogenetic test using this technology in September 2007.[64] This technology can provide genotyping results in a few hours, providing busy clinicians with rapid answers without the need of waiting 1-2 days for the usual genotyping test before starting a new medication.

In the view of the author,[62] personalized prescription is the branch of PM focused on individualizing drug prescription. To predict drug response, one has to use genetic information in the context of environmental and personal factors;[65] furthermore, these other factors may be more important than genetics in determining drug response with some drugs. Thus, personalized prescription should be built around the idea that pharmacology is a mechanistic science. Chemistry (and physics) are foundational in understanding pharmacological response but pharmacologists who are familiar with pharmacological mechanisms at the cellular, tissue, and organic levels, rather than chemists, are the ones who can predict drug response.[66] Unfortunately, pharmacokinetic and pharmacodynamic mechanisms vary from drug to drug.[62] In a view of personalized prescription based on pharmacological mechanisms one can propose the following: [67,68] 1) a drug's pharmacokinetic and pharmacodynamic mechanisms are behind its efficacy and safety; 2) genetic, environmental and personal variables influence drug pharmacokinetic and pharmacodynamic mechanisms and through them its efficacy and safety; 3) personalizing drug selection is much more complex than personalizing drug dosing; 4) in the process of personalizing drug selection, eliminating a drug in the process is easier than choosing a drug; 5)

personalizing dosing is easier when the drug follows linear kinetics and has a narrow therapeutic window; and 6) personalizing dosing in wide-therapeutic-window drugs may not make practical sense since physicians may be very arbitrary in their selection of drug dosages.

Statistical heterogeneity or lack of homogeneity and PM

The crucial role of the level of statistical heterogeneity in PM and its implications related to the statistical power of the studies has not been addressed by the literature. The first question regards personalizing prescription for a specific drug response. Is it completely heterogeneous (each patient responds differently)? Is it somewhat heterogeneous? Is it homogeneous (most patients respond similarly to the drug)? Thus, knowing the statistical distribution of drug response is absolutely crucial in the design of PM studies. If drug response is absolutely heterogeneous (it lacks homogeneity) and cannot be extrapolated from individual to individual, it is impossible to study or predict unless we focus on the prior response of the same drug in the same patient. As pharmacokinetic and pharmacodynamic mechanisms are fairly similar across all humans, probably due to the effects of individual selection associated with evolution, drugs with non-homogeneous response do not appear to exist. Thus, patients appear to be relatively homogeneous in drug response, with differing levels of heterogeneity. However, heterogeneity varies from drug to drug due to the peculiarities of variations in pharmacokinetic and pharmacodynamic mechanisms. If one wants to defend EBM as being a good representation of drug response, one must assume that the mean represents the population well and that the standard statistical test can be properly used since the sample shows statistical homogeneity (Table 1). If one wants to defend PM as being a better representation of drug response than EBM, one must assume that a good number of subjects are outliers and should be excluded from the usual statistical analysis and that mean results do not represent them; therefore outliers should be studied separately (Table 1).

To illustrate, three outlier prevalences in the population are considered in this discussion as examples: <1% of the population; 1-10%, and 50%. If the outliers are so rare as to include <1% of the population, they are quite unlucky; it is very unlikely that well-designed studies using RCTs will ever be

conducted in them. This is the tragedy of EBM: the more difficult the patient is, the more likely that he/she will be ignored by the EBM approach.[65] An example may clarify this: some very unlucky subjects do not have CYP2D6 and CYP2C19. They number approximately 1 per 1000 or fewer in each race.[69] If you are one of these unusual subjects, it may be better for you not to need an antidepressant since almost all of them are mainly metabolized by one of these two enzymes (and frequently the other CYP is used as an auxiliary pathway). [70] If you are a pharmacologist interested in antidepressants and are trying to recruit 50 of these double poor metabolizers for a controlled clinical trial you have the herculean task of needing to screen 50,000 patients to identify 50 who meet the study criteria.

On the other hand, when "unusual" subjects include 50% of the population, such as with gender, there may be "too many outliers." Let's imagine that one scientist believes males can respond to a new drug but most females do not respond. He/she will have the difficult task of convincing a pharmaceutical company to develop this drug for only half of the market, in this case males, when the competing drugs are approved for both genders.

If the outliers (patients who show heterogeneity) account for 1-10% of the population, the possibility of studying them is dependent upon resources, and is influenced by the number of groups used in the randomization or for the tests of heterogeneity in treatment effects, the percentage of outliers, and the use or nonuse of some kind or enrichment strategy for increasing the number of outliers randomized into the study. In summary, it is a question of the financial cost and practicality of the RCTs.

The history of the attempts to explore treatment response heterogeneity in psychiatry

Psychiatry has a long history of trying to identify predictors of differential pharmacological response but these attempts began before the EBM approach was born, were not developed in the context of RCTs and were not tested in RCTs. These attempts came from a different tradition; the mechanistic tradition. The developers of these models hypothesized that psychiatric drugs work by modifying specific pharmacological mechanisms in the brain. Two of the most important attempts are briefly described: the

existence of a serotonergic versus noradrenergic type of depression, and the dexamethasone suppression test (DST).

The monoamine hypothesis of depression was first formulated in the 1960s[71] and some evidence that different antidepressants may have differential effects on serotonin and noradrenalin metabolites[72] led to efforts to classify depressive patients according to pharmacological mechanisms. In that sense, in 1975, Maas[73] simplistically hypothesized that there are two groups of depressive patients: A (with disorders of norepinephrine systems) and B (with disorders of serotonin systems). This approach was described early in the first attempts to develop pharmacological guidelines in psychiatry[74], before the birth of the EBM movement. These guidelines[74] were developed by experts using a comprehensive approach including the mechanistic and RCT approaches, trying to balance all biological and clinical data and, therefore, giving a lot of weight to data beyond those obtained from RCTs.

In psychiatry, the DST was the most important biomarker used as a potential index of heterogeneity of treatment response. The initial DST research studies in depression in the 1970s[75,77] led to enthusiasm with which mainstream psychiatry embraced a biological test of the heterogeneity of pharmacological response. However, in 1987, a guideline published by The American Psychiatric Association (APA) Task Force on Laboratory Tests[77] indicated the lack of definitive data of the DST's clinical usefulness in selecting treatment. In 1988 Nierenberg and Feinstein[78] used the history of the DST, a diagnostic test initially widely accepted and later rejected, as a cautionary example for avoiding similar problems in the future with diagnostic tests.

These two promising but, in the end, unsuccessful attempts to use biological markers in psychiatry are prime examples that the history of addressing heterogeneity in treatment response in psychiatry have been rather unproductive. Thus, all recent APA practice guidelines[79] incorporate EBM principles but also consider clinical consensus. These guidelines address treatment response heterogeneity but they correctly point out the limitations of the data.

Next steps for moving forward from the current stalemate between EBM and PM

This editorial has presented the unsatisfactory status of medical science due to the stalemate between EBM and PM and presented a caricatured view of their traditions in order to keep the interest of the reader.

To move forward, three issues appear important and are described in the next three sections: 1) advances in scientific methodology which would bring EBM and PM closer; 2) education of the defenders of EBM and PM about the weaknesses of their approaches and the virtues of the other side; and 3) the need for openness toward advances in science, in general, that may rescue medicine from this stalemate. Many specific solutions are presented in these three sections; unfortunately, many of these ideas are not likely to be implemented. Thus, the discussion of each specific idea is concluded with a comment on its limitations.

Advances in scientific methodology which would bring EBM and PM closer

Additional development is needed in the following areas: 1) scientific methodologies for PM, 2) knowledge of drug-response heterogeneity, 3) introduction of a PM approach to RCTs, and 4) use of a PM approach in the development of guidelines for EBM.

Development of scientific methodology for PM

Regarding the development of methodology for PM, it is important to remember that RCTs are usually funded by industry while most of the pharmacogenetic studies using genetic testing to expand PM have been conducted by research institutions.[80] Two excellent reviews describe some possible types of designs for pharmacogenetic studies including RCTs.[81,82] As long as no funding sources are identified for these studies, this discussion of design types for researching methodological progress in PM is mostly a theoretical exercise. Grant agencies have traditionally left drug study funding to industry. Pharmaceutical companies have no interest in conducting studies once a drug is on the market and especially when it is off patent. Companies developing PM tests have very small research budgets when compared with pharmaceutical companies.[61] In theory, pharmacogenetic studies may help to bring increased use of some

of the classic antipsychotic or antidepressant drugs by reducing ADRs[83] or may even rescue drugs withdrawn from the market.[84] In practice, the lack of funding and the enormous cost of any type of study that may produce competition with approved drugs on the market and supported by the pharmaceutical company marketing is a serious limitation.

Studying variability between individuals in drug response

The literature describes three methods for exploring the differences in drug response among individuals: 1) pharmacogenetic twin studies, 2) repeated drug studies, and 3) replicate cross-over trials after RCTs.

The typical means of studying genetic versus environmental influences in complex illness has been the use of twin studies to compare differences between dizygotic and monozygotic twins. In classic formulations monozygotic twins are said to share 100% of their genome while dizygotic twins are said to share on average only 50%; in both type of twins, the twin and his/her co-twin were believed to share similar levels of environmental influences.[85,86] Statistical models have been developed allowing causal associations to be made and estimations of heritability to be defined as the proportion of total variance that is accounted for by genetic components.[85] The assumption of shared environment has always been dubious but advances in genetics are beginning to question the genetic assumptions of the twin models since monozygotic twins may not share 100% of their genome and they may have differences in CNVs or epigenetic changes.[86] Thus, twin studies may be a powerful method with high potential for studying some aspects of drug response but there are practical limitations since it is not ethical to administer drugs prospectively in long-term studies to healthy twins under controlled conditions. The twin model can be used for short studies with little risk, particularly using single doses and focusing on pharmacokinetic effects. [87] As a matter of fact, recently some twin pharmacological studies have focused on the pharmacokinetics of some medications or substances of abuse.[85,88,89] Besides these controlled studies, "natural experiments" may occur when monozygotic twins take the same drug but show a discordant drug

response (e.g., ADRs). In summary, twin studies have some potential to explore the heritability of some aspects of drug response and advance scientific knowledge in PM but are limited by practical aspects.[85]

Kalow and coworkers[42,43] proposed investigating response variation by repeating drug administration to the same individuals and comparing the variability of response within and between individuals (intra-individual variation versus inter-individual variation). Inter-individual variation in response to the same drug may have environmental and/or genetic reasons. Intra-individual drug variation cannot be explained by genetic variations; it can only be explained by instability of measuring methods (test-retest reliability), chance, environment or changes in genetic function induced by environment (epigenetics).[43] For this type of study it is important to use reliable measures, exclude chance by using enough relatively large samples, and eliminate known environmental factors that influence drug response, including those that may change drug response over time (such as drug tolerance, which is probably mediated by epigenetic mechanisms).[57] The comparison of intra-individual variation versus inter-individual variation provides a measure of heritability. More recently, Kalow's colleagues[90] have: 1) clarified the types of drugs that can be studied using this method; 2) proposed that a comparison of inter- versus intra-subject variability in drug response under minimal environmental exposure may provide an upper-bound estimate of heritability of drug efficacy and safety; and 3) indicated that modest changes in population averages may underestimate the dramatic impact of a genetic variation at the tails of a population. According to Dorne,[91] the effects of polymorphic enzyme outliers may contribute to huge differences in metabolism up to a factor of 52 times in CYP2D6 and of 26 times on CYP2C19.

Senn,[41] a statistician, has taken Kalow's proposal[42,43] one step further by proposing the use of what he calls replicate cross-over trials. They are a particular type of cross-over design: those in which the same treatment is given to the same subject more than once. A replicate cross-over trial is the equivalent of doing '*n*-of-1 trials' in all the patients included in the RCTs. He also indicated that the data need to be analyzed using mixed effects models. This study design is only possible in drugs used for chronic illnesses, with reversible effects and no need for long treatments.[38] More importantly, it will make RCTs

much more expensive. Mant[92] opines that it is naïve to think that replicate cross-over trials are a feasible solution and that we should resign ourselves to using observational studies to answer questions about applicability to individual patients.

In summary, pharmacogenetic twin studies, repeated drug administration studies, and adding '*n*-of-1 trials' after RCTs can be used to establish an upper-bound estimate for individual variations (and possibly for heritability) of different aspects for some drug responses, but they have rarely been used and have some practical limitations. Thus, the current status of RCTs is that they can tell us which treatments are effective but not necessarily which patient should receive them.[93,94]

Personalizing RCTs

Predictions based on mechanistic knowledge of genetic or environmental influences on drug responses can be incorporated in the design of RCTs. An example would be to stratify patients by genetic markers according to our knowledge of pharmacogenetics. The CYP2D6 gene variations are the best understood; one extreme is the poor metabolizer who does not have any CYP2D6 due to lack of functional genes and the other extreme is the ultrarapid metabolizer who has too much CYP2D6 enzyme due to three or more active genes. Ideally, RCTs stratified by CYP2D6 genotype should be carried out for classical antidepressant and antipsychotic drugs that are off-patent but these RCTs will not be supported by the industry. Furthermore, the industry has eliminated CYP2D6 drugs from their pipeline;[9] no future CYP2D6 drugs will be studied using CYP2D6 knowledge to stratify RCTs.

Our knowledge of the influence of inhibitors and inducers can also be incorporated in RCTs. The industry, however, usually prefers to exclude patients taking inducers or inhibitors from their RCTs and ignore the issue completely. A good example of "ignorance" in psychiatry is the failure of risperidone as an adjunct treatment in mania compared with placebo when risperidone was added to carbamazepine in a RCT.[95] A case report,[96] case series[97,98] and in vitro studies[99,100] had demonstrated that carbamazepine is an inducer of risperidone metabolism, but the pharmaceutical company's ignorance of the literature led to a partial failure (in the carbamazepine arm) in this RCT.[95] To obtain a positive outcome in the

carbamazepine arm, the company should have paid attention to this prior mechanistic literature and randomized the carbamazepine patients to double the risperidone dose compared with the other manic patients taking non-inducers (lithium or valproate).

Personalizing EBM guidelines

Guidelines for EBM can be more PM-friendly by: 1) incorporating a more patient-centered approach, 2) incorporating a mechanistic pharmacological approach, and 3) rethinking the hierarchy of grading recommendations, particularly for drug safety.

The editors of the family medicine journals[101] have proposed that it is possible for guidelines to emphasize patient-centered outcomes rather than focus on illness-centered outcomes. Clinical pharmacologists and pharmacists have incorporated mechanistic knowledge in the development of pharmacogenetic guidelines[102-103] and EBM guidelines for diagnostic tests. [105] The approach is different than the approach of traditional EBM guidelines since the emphasis is not so much on the availability of RCTs to support these recommendations but also incorporates the available mechanistic information in the guidelines. Obviously, if one thinks that RCTs is the only way to establish reliable knowledge in medicine these types of guidelines incorporating mechanistic knowledge are worthless.

Vandenbroucke,[106-109] a physician with epidemiological training, has emphasized that the hierarchy of innovation in medicine is exactly the opposite of the hierarchy used by EBM in guidelines. RCTs are the less innovative and anecdotal findings are the most innovative. More importantly, he emphasized that EBM forgets that RCTs were developed to test drug efficacy, not drug safety. As a matter of fact, Ioannidis[110] proposes that ADRs are "neglected, restricted, distorted, and silenced" in RCTs. According to Vandenbroucke,[106-109] RCTs, due to their nature, are limited in detecting ADRs since they include: 1) a small number of patients (a few hundred), 2) a limited population (patients with comorbidities and poor health are excluded), and 3) a relatively short duration (usually a few months). Randomization is not helpful or important for unexpected and unpredictable ADRs. As a matter of fact, observational studies (case reports, case series and pharmacoepidemiological studies) have been crucial in

establishing new ADR knowledge.[109] Thus, strictly from the best scientific methodological point of view, EBM's hierarchy of putting RCTs at the top for drug efficacy may be correct, but for drug safety, RCTs may not be particularly helpful. This is demonstrated by the withdrawals from the market of many drugs that were proven to be safe in RCTs and were not safe when used in the general population[111]. It is also demonstrated by an aldosterone RCT that suggested that this drug was associated with decreased mortality in advanced congestive heart failure, leading to an EBM guideline and a subsequent increase in hospitalizations with no reductions in mortality when aldosterone was used in the general population of this type of patient.[112] These examples of major problems in extrapolating from the patients studied in RCTs to the general population reflects that RCTs may have major problems with external validity.

Educating EBM and PM defenders regarding their weaknesses and the virtues of the other side

It is fitting to be reminded of Reilly's[24] description that EBM is an idealized way of practicing medicine. Four physicians with epidemiological expertise are included in this editorial and the accompanying references: Feinstein,[1,20,27,33,78] Ioannidis,[60,110,114-116] Sackett,[10,13,16,21,] and Vandenbroucke.[5-7,17,106-109,117] If all physicians were like them, it would be very easy to implement EBM. Unfortunately, we (physicians) do not have the superb mathematical minds of these four experts. EBM assumes that physicians have the time and knowledge to perform complex computer searches and complete meta-analyses to resolve their clinical questions. The reality is that completing meta-analyses is not "easy" science. A statistician has recently complained that several of the published meta-analyses in top medical journals have significant flaws.[118]

The writer spends a substantial part of his time helping practicing psychiatrists and psychiatry residents with their practical pharmacological problems. This includes considerable effort in teaching them to: 1) use PubMed to search for reviews and meta-analyses, and 2) use pharmacokinetic and pharmacodynamic knowledge to resolve their questions. He has had some success in these tasks but he has had no success in teaching them basic statistical principles such that they could intelligently assess the typical article published in a psychiatric journal. The McMaster University students and physicians may

be particularly sophisticated in statistics and epidemiology but the average US practicing physician and resident may not be as capable. This is why Feinstein[27] insisted that the science of EBM was not developed by physician scientists but rather by scientists who were not physicians. A deceptively easy solution would be to increase the mathematical training of prospective physicians. In the US, pre-med curricula and medical school entrance exams emphasize biology, chemistry, and physics, with little emphasis on statistics or mathematical thinking. One can propose substituting the emphasis on physics with an emphasis on statistics. Is this likely to happen? No. Is this wise? That is uncertain. The psychology of science[119] is in its initial stages but there may be a dichotomy of personality traits among different individuals according to their career selections. Clinicians tend to be people-oriented and scientists tend to be oriented to things.[119] Mathematicians, engineers and physicists may have lesser people skills than social scientists. It would not be surprising, then, if increasing the mathematical requirements for medical students led to different populations of prospective students, with less interest in people and fewer social skills.

The reality is that the application of pharmacological scientific advances is a complex process. In the US, where there are few physicians with formal clinical pharmacological training (versus European countries that have clinical pharmacology residencies), the mechanistic science of pharmacology is mainly developed by pharmacologist scientists and some pharmacists with basic skills; physician scientists make minimal contributions. In the area of RCTs some US physician scientists may be the first authors of articles but most of the complex statistical designs and analyses are done by biostaticians or epidemiologists working for the industry. The meta-analyses of pharmacological studies are usually authored by biostaticians or epidemiologists. On the other hand, in the US most of the prescribing is done by practicing physicians (or physician extenders) who have very limited pharmacological and statistical training. If these practicing clinicians have access to a pharmacist and are wise, they consult the pharmacist, who could provide a more comprehensive and mechanistic pharmacological view in treating problematic patients. In summary, the author believes that using EBM and conducting meta-analyses for

making decisions about drug prescription is an "ideal" way of practicing medicine, but is beyond the abilities of current US physicians (and physician extenders).

Many of the scientists practicing EBM and writing meta-analyses appear not to understand that theories and mechanistic interpretations are a fundamental aspect of medical practice. Any pharmacologist knows that selective serotonin reuptake inhibitors (SSRIs) frequently exhibit no linear relationship between dosage and plasma concentration, wide ranges between therapeutic and toxic doses, and some are powerful self-inhibitors of their metabolism.[65] This means that SSRIs are not good candidates for CYP genotyping.[65] The TCAs are much better candidates for CYP genotyping. However, when the epidemiologists from the Agency for Healthcare Research and Quality[120,121] decided to examine the clinical utility of CYP genotyping in psychiatry, they selected SSRIs instead of TCAs. Not surprisingly, the outcome was not good for PM defenders.[120,121] Pharmacokinetic knowledge in the CYP area would have saved the effort and expense of completing a comprehensive review and a meta-analysis, but researchers with an EBM background[120,121] did not see the need to review the literature on the pharmacokinetic science of SSRIs.

Medical science includes epidemiological principles and basic mechanistic science; medical practice is the difficult art of taking this knowledge to clinical practice with individual patients. Due to the complexity of the task, physicians need to rely on and learn from epidemiologists and basic scientists, including pharmacologists. Thus, it is obvious that medicine needs both EBM and PM, and collaboration between defenders of EBM and PM is needed to advance medicine. Defenders of each view must acknowledge the limitations of their approach, their lack of expertise in the other approach, and that collaboration is needed. Tegarden[122] used a clinician's way of thinking by describing a risk-benefit analysis to justify the use of warfarin pharmacogenetics in spite of the current lack of available RCTs.

Advances in science in general that may rescue medicine from this stalemate

Three major new advances in science may be relevant for those who think that "outside" help could resolve this stalemate in medicine: 1) the progressive understanding of how science advances, along

with its limitations; 2) a new way of scientific thinking called "complexity"; and 3) the initial steps in using the scientific approach for defining an expert.

An interesting advance that is finally being accepted by the experts is that scientific thinking: 1) has been very successful but has its weaknesses, 2) has no completely clear demarcations from "pseudoscientific" thinking, and 3) is not so different from the thinking used in other sophisticated human activities (such as law or the arts).[117,123-125] In summary, science is a complex trial-and-error historical process led by experts, the scientists.[117,123-125] Thus, the influence of subjective and personal issues cannot be avoided, to the point that major advances in science can be explained by a complex mix of 1) "charge" (discoveries solve problems that are quite obvious but in which the way to solve the problem is not so clear), 2) "challenge" (discoveries are explained by a new acknowledgment of the limitations of scientific thinking, and 3) "chance" (serendipitous findings made by "prepared minds").[126] Ioannidis,[114-116] a physician with mathematical training, has conceptualized for medical science the limitations of scientific thinking and the difficulties in advancing it. In a series of seminal articles, he has reminded us that the majority of published medical findings are false, including many of those obtained through RCTs.[114-115] He also stated that "important" articles and "important" journals are not free from the problem of spreading false findings; moreover, they are even more prone to it due to the competition to publish "significant" findings.[115] The "boring job" of replication,[128] the very cornerstone of establishing a scientific finding as truth, is discouraged. Finally, Ioannidis pointed out that significant biases, both personal and financial, contribute to the dissemination of false findings.[116] In this context, it seems important to remind the defenders of the PM and EBM approaches that their fields may not be so different in that both are full of false findings; both are biased, but their varying traditions result in different biases. Working together and using the strengths of the other side to balance their weaknesses appears wise since, according to Naylor[116], "clinical medicine seems to consist of a few things we know, a few things we think we know (but probably don't), and lots of things we don't know at all."

A new current in scientific thinking is called "complexity"[130-131], reflecting the complexity of scientific exploration and the need for complex computer models. This theory and its accompanying computer models are being introduced in medicine, the so-called "network" medicine.[132] Network analysis may have potential to explore the myriad of genes that might contribute to PM. Two applications are briefly discussed here: 1) gene-gene interactions, and 2) ethnicity. When one is sure that a single gene has important and definitive effects on response to a specific drug, the task of probing it is "relatively" simple; one only needs to establish how all genetic and epigenetic changes modify the activity of the gene and determine the clinical relevance of these changes. Therefore, one assumes that this gene has "large effect sizes" on the drug's safety or efficacy so that other genes can be ignored. However, the "elephant in the living room" of PM is how to deal with the possibility that myriads of gene-gene interactions are likely to be important. The problem of dealing with gene-gene interactions may be mind-boggling if one moves from 2 to 3 genes to 20 or 50 genes interacting. Network analyses[133], by developing gene pathways, may be a reasonable way of dealing with unknown complex gene-gene interactions.

From the genetic point of view the concepts of race or ethnicity do not make sense, but there are obvious genetic differences related to ancestry, due to the evolution of human populations.[134] Unfortunately, what we know about the functionality of genes used in PM (e.g., CYP2D6 and CYP2C19) cannot be predicted easily by ancestry markers. Network analysis[135] has been used recently to control for ancestry but it is not known whether this will be a successful method or not to deal with this complex issue. Badcott[136] has argued that putting too much faith in the idea of multiple genes facilitating PM may be misleading. The aleatory process of evolution ensures that, with the possible exception of monozygotic twins, every human being is genetically unique and that predictive power may be enhanced by knowledge and technology, but is ultimately constrained by insuperable probabilistic considerations. It is illusory to relate population-level probabilistic causation with individual-level causation.

Only time will time will tell if these "complexity" and "network" models are a passing fad or a contribution to individual patient care through the incorporation of complex information from multiple

sources, the final goal of PM. Certainly, it seems hard to convince current practicing physicians that in the future specific treatment decisions will be made by ignoring their "subjective" experience and introducing hundreds or thousands of pieces of data from genetic or other biomarkers into a computer model that will "magically" provide the answer. The word "magically" is used here in the sense that physicians will have little clue as to how the software works or was developed. Another new statistical methodology that may be relevant but is even harder to discuss than network analysis, due to its novelty, is the so-called dynamic treatment regimens.[137] These regimens allow individualization of treatment (type, dosage, timing) of a single patient by using a special type of randomized design (called sequential multiple assignment randomized trial). Another regimen, a regression-based analysis called Q-learning, may allow exploring similar data in non-randomized naturalistic long-term treatment designs. [137]

An even more "futuristic" scientific development may be science advancement to the point that one can study complex human concepts such as "what is an expert."[138] The scientific study of how individuals become experts should be important for both: 1) the defenders of EBM, who believe that training with an "expert" was a key tradition in the "art" of medicine, and that EBM is "revolutionary" in its attempt to make all physicians more expert and more scientific; and 2) the enemies of EBM, who believe that EBM is a part of social process in which the "old" authoritarian structures are hidden behind statistical techniques that still need to be interpreted by "experts." [139,140]

Physicians do not talk about who is an expert but they demonstrate it with their behavior. In a simplistic way, in clinical medicine, "experts" are probably those physicians who are consulted on difficult cases by other physicians from the same specialty. From the point of view of medical trainees, "experts" are those who are more frequently chosen as "mentors" by the young physicians who want to learn medicine. In the scientific area, due to large amounts of data available from publications, a more "statistical" and "numerical" approach toward determining who is an expert is possible by testing how influential the researcher's articles are. An example of these attempts may be the index "h" developed by Hirsch[141] to compare, using one number, the most frequently quoted authors in a scientific area. As

English is definitively established as the *lingua franca* for disseminating research and journals from countries other than Anglo-Saxon countries have fewer barriers toward being read by researchers, this type of index may help to establish the identities of the "current" world experts in a specific area of publication. Unfortunately, there is no guarantee that the current medical experts would be the ones considered "experts" when the researchers of the next century are looking back. Moreover, history tells us that great discoverers are usually ignored by their contemporaries. The scientific approach has not been particularly successful in studying and explaining some of the more complex concepts of human life, such as expertise and creative spirit, but researchers from different areas including educational sciences are trying to define what an "expert" is.[138] Even with these advances in understanding what an expert is and how to train them, one must hold out the possibility that inherent limitations exist in what science can offer for better training experts among physicians. According to Polanyi,[125,142] a thinker who began as a physician, then became an internationally known chemist and finally a philosopher of science, the process of learning medicine with a mentor requires some implicit learning, which is by nature difficult to translate into words and is only learned by example.[142]

Conclusion

Prior articles have described the tension between EBM and PM regarding economic issues[143] and practical implementation in the clinical environment[144-148]; one has even proposed that EBM has "been effectively sidelined and marginalized" by PM.[149]

This editorial focuses on historical and statistical issues that cannot be ignored in understanding this tension. The more we believe that PM is needed for one drug, the more this means that drug response is not homogeneous and that an unmodified EBM approach will not be helpful. If we decide that EBM is the way to go and drug response is homogeneous and well-represented by the mean, we are simply ignoring the patients who need personalized prescription. How bad is this? It depends: are we ignoring <1%, 10% or 50% of the sample? The pharmaceutical company marketing a drug may feel safe in ignoring "a few" outliers. The outlier patients and their families may not be so happy about being ignored.

The dilemma entails balancing the public health approach and the individualized approach.[8,9] It is easy to see why the EBM approach and the personalized prescription approach can be enemies in times of limited economic resources and the growth of health expenses due to technological developments.

The history of using biological markers in psychiatry to address heterogeneity in treatment response has been rather unproductive. To move beyond the stalemate between EBM and PM, this editorial proposes three types of steps: 1) advancing the scientific methodology to bring EBM and PM closer; 2) educating the defenders of EBM and PM about the weaknesses of their approaches and the virtues of the other side; and 3) openness to advances in science, in general, that may rescue medicine from this stalemate.

An armistice between EBM and PM is possible but it will require changing how the various interested parties (government agencies, health organizations, health providers and patients) think about the complex pharmacological mechanisms of drug responses. It will also require the use of new methodological designs for building a PM approach for studying new drugs through the use of stratification following mechanistic hypotheses in RCTs and replicate cross-over trails after RCTs. This may, however, leave already-marketed drugs with no information on personalization unless funding for pharmacogenetic twin studies and repeated drug administration designs is found.

"Personalized EBM" [150] is a new term being introduced in the literature. In the view of the author, the development of such an approach can only be based on prescribers gaining a better understanding of pharmacological mechanisms,[62] as well as a general acknowledgment by all parties (government agencies, health organizations, health providers and patients) of the limitations of EBM and PM. They are both complementary and antagonistic in their approaches, such that collaboration between the experts in both fields is needed in the advancement of pharmacological science and its applications in the treatment of individual patients.

References

1. Feinstein AR. Clinical biostatistics I. A new name and some other changes of the guard. *Clin Pharmacol Ther.* 1970;11:135-148.
2. de Leon J, Spina E, Diaz FJ. Pharmacokinetic drug interaction studies must consider pharmacological heterogeneity, use of repeated dosing, and translation into a message understandable to practicing clinicians. *J Clin Psychopharmacol.* 2009;29:201-205.
3. Tröhler U. Commentary: 'Medical art' versus 'medical science': J Civiale's statistical research on conditions caused by calculi at the Paris Academy of Sciences in 1835. *Int J Epidemiol*. 2001;30:1252-1253.
4. Grahame-Smith D. Evidence based medicine: Socratic dissent. *BMJ.* 1995;310:1126-1127.
5. Vandenbroucke JP. Evidence-based medicine and "médecine d'observation". *J Clin Epidemiol.* 1996;49:1335-1338.
6. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med.* 2008;5:e67.
7. Vandenbroucke JP. Clinical investigation in the 20th century: the ascendancy of numerical reasoning. *Lancet.* 1998;352(Suppl 2):SII12-SII16.
8. Buetow S, Getz L, Adams P. Individualized population care: linking personal care to population care in general practice. *J Eval Clin Pract.* 2008;14:761-766.
9. de Leon J, Susce MT, Murray-Carmichael E. The AmpliChip CYP450 genotyping test: integrating a new clinical tool. *Mol Diagn Ther*. 2006;10:135-151.
10. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312:71-72.
11. Rangachari PK. Evidence-based medicine: old French wine with a new Canadian label? *J R Soc Med.* 1997;90:280-284.

12. Gerber A, Lungen M, Lauterbach KW. Evidence-based medicine is rooted in Protestant exegesis. *Med Hypotheses.* 2005;64:1034-1038.

13. Sackett DL. Applying overviews and meta-analyses at the bedside. *J Clin Epidemiol.* 1995;48:61-66.

14. Cohen L. McMaster's pioneer in evidence-based medicine now spreading his message in England. *CMAJ.* 1996;154:388-390.

15. Gray GE. Evidence-based medicine: an introduction for psychiatrists. *J Psychiatr Pract.* 2002;8:5-13.

16. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA.* 1992;268:2420-2425.

17. Vandenbroucke JP. 175th anniversary lecture. Medical journals and the shaping of medical knowledge. *Lancet.* 1998;352:2001-2006.

18. Guyatt G, Cook D, Haynes B. Evidence based medicine has come a long way. *BMJ.* 2004;329:990-991.

19. Shah HM, Chung KC. Archie Cochrane and his vision for evidence-based medicine. *Plast Reconstr Surg.* 2009;124:982-988.

20. Jensen UJ. The struggle for clinical authority: shifting ontologies and the politics of evidence. *BioSocieties.* 2007;2:101-114.

21. Sackett DL, Rosenberg WM. The need for evidence-based medicine. *J R Soc Med.* 1995;88:620-624.

22. Woolf S. Evidence-based medicine: a historical and international overview. *Proc R Coll Physicians Edinb.* 2001;31(Suppl 9):39-41.

23. Sehon SR, Stanley DE. A philosophical analysis of the evidence-based medicine debate. *BMC Health Serv Res.* 2003;3:14.

24. Reilly BM. The essence of EBM. *BMJ*. 2004;329:991-992.

25. Choudry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. *Ann Intern Med.* 2005;142:260-273.

26. Claridge JA, Fabian TC. History and development of evidence-based medicine. *World J Surg.* 2005;29:547-553.

27. Feinstein AR. What can clinicians contribute to mutual challenges in medical statistics. *J Clin Epidemiol.* 1997;103:529-535.

28. Hope T. Evidence based medicine and ethics. *J Med Ethics.*1995;21:259-260.

29. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA.* 2003;290:1624-1632.

30. de Leon J, Mallory P, Maw L, et al. Lack of replication of the association of low serum cholesterol and suicide attempts in another country opens more questions. *Ann Clin Psychiatry.* 2011;23:163-170.

31. Knottnerus A, Dinant GJ. Medicine based evidence, a prerequisite for evidence based medicine. *BMJ.* 1997;315:1109-1110.

32. Susce MT, Villanueva N, Diaz FJ, et al. Obesity and associated complications in patients with severe mental illnesses: a cross sectional survey. *J Clin Psychiatry.* 2005;66:167-173.

33. Feinstein AR, Horwitz RI. Problems in the "evidence" of "evidence-based medicine". *Am J Med.* 1997;103:529-535.

34. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA.* 2007;298:1209-1212.

35. Kent DM, Rothwell PM, Ioannidis JP, et al. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials.* 2010;11:85.

36. Johnson BA, Ait-Daoud N, Seneviratne C, et al. Pharmacogenetic approach at the serotonin transporter gene as a method of reducing the severity of alcohol drinking. *Am J Psychiatry.* 2011;168:265-275.

37. Senn SJ. Applying results of randomised trials to patients. N of 1 trials are needed. *BMJ*. 1998;317:537-538.

38. Senn SJ. Individual therapy: new dawn or false dawn. *Drug Infor J.* 2001;35:1479–1494.

39. Senn SJ. Author's reply to Walter and Guyatt. *Drug Infor J.* 2003; 37:7–10.

40. Senn S. Individual response to treatment: is it a valid assumption? *BMJ.* 2004;329:966-968.

41. Senn S, Rolfe K, Julious SA. Investigating variability in patient response to treatment—a case study from a replicate cross-over study. *Stat Methods Med Res.* 2010 Aug 5 [Epub ahead of print].

42. Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. *Pharmacogenetics.*1998;8:283–289.

43. Kalow W, Endrenyi L, Tang BK. Repeat administration of drugs as a means to assess the genetic component in pharmacological variability. *Pharmacology*. 1999;58:281–284.

44. Ruano G. Quo vadis personalized medicine. *Personalized Med.* 2004;1:1-7.

45. Meyer UA. Pharmacogenetics - five decades of therapeutic lessons from genetic diversity. *Nat Rev Genet*. 2004;5:669-676.

46. Caldwell J. Drug metabolism and pharmacogenetics: the British contribution fields of international significance. *Br J Pharmacol.* 2006;147(Suppl):S89-S99.

47. Kalow W. Perspectives in pharmacogenetics. *Arch Pathol Lab Med*. 2001;125:77-80.

48. Motulsky AG. From pharmacogenetics and ecogenetics to pharmacogenomics. *Med Secoli*. 2002;14:683-705.

49. Garrod AE. *Inborn errors of metabolism.* London: Oxford University Press, 1909.

50. Vogel F. Moderne probleme der Humangenetik. *Ergeb Inn Med Kinderheild.* 1959;12:52-125.

51. Fodor SP. Massively parallel genomics. *Science.* 1997;277:393-395.

52. Pirmohamed M. Pharmacogenetics and pharmacogenomics. *J Clin Pharmacol.* 2001;52:345-347.

53. Science. New research horizons. *Science.* 1997;278:2039.

54. Golden F, Lemonick MD. The race is over. *Time.* 2000;July3:18-23.

55. Lertola J. Deciphering the code and what might come from it. *Time.* 1999;Nov 8:68-69.

56. Collins FS, McKusick VA. Implications of the human genome project for medical science. *JAMA*. 2001;285:540-544.

57. Nebert DW, Zhang G, Vesell ES. From human genetics and genomics to pharmacogenetics and pharmacogenomics: past lessons, future directions. *Drug Metab Rev*. 2008;40:187-224.

58. de Leon J. The AmpliChip CYP450 Test: personalized medicine has arrived in psychiatry. *Expert Rev Mol Diagn*. 2006;6:277-286.

59. de Leon J, Arranz MJ, Ruaño G. Products for pharmacogenetic testing in psychiatry: A review of features and clinical realities. *Clin Lab Med*. 2008;28:599-617.

60. Ioannidis JP. Expectations, validity, and reality in omics. *J Clin Epidemiol*. 2010;63:945-949.

61. de Leon J. Pharmacogenomics: the promise of personalized medicine for CNS disorders. *Neuropsychopharmacol*. 2009;34:159-172.

62. de Leon J. The future (or lack of future) of personalized prescription in psychiatry. *Pharmacol Res*. 2009;59:81-89.

63. Lee HK, Lewis LD, Tsongalis GJ, et al. Validation of a CYP2D6 genotyping panel on the NanoChip Molecular BiologyWorkstation. *Clin Chem*. 2007;53:823-828.

64. Nanosphere, Inc. Nanosphere announces first FDA cleared genetic test for warfarin sensitivity and nanotechnology-based molecular diagnostics platform. September 18, 2007. Available at: http://ir.nanosphere.us/phoenix.zhtml?c=214748&p=irol-newsArticle&ID=1075936&highlight Accessed on August 8, 2011.

65. de Leon J. The potential of genotyping (letter). *Science*. 2008;321:769.

66. Polanyi M. Life's irreducible structure. *Science*. 1968;160:1308-1312.

67. de Leon J. Future of personalized prescription in psychiatry. In: Schwab M, Kaschka WP, Spina E eds. *Pharmacogenomics in Psychiatry. Advances in Biological Psychiatry 25*. Basel, Switzerland: Karger, 2010:118–134.

68. de Leon J. Paying attention to pharmacokinetic and pharmacodynamic mechanisms to progress in the area of anticholinergic use in geriatric patients. *Curr Drug Met*. 2011;12:635-646.

69. de Leon J, Susce, MT, Johnson M, et al. DNA microarray technology in the clinical environment: the AmpliChip CYP 450 Test for CYP2D6 and CYP2C19 genotyping. *CNS Spectrums*. 2009;14:19-34.

70. Johnson M, Markham-Abedi C, Susce MT, et al. A poor metabolizer for both Cytochrome P450 2D6 and 2C19 (CYP2D6 and CYP2C19): a case report on antidepressant treatment. *CNS Spectrums*. 2006;11:757-760.

71. Hirschfeld RM. History and evolution of the monoamine hypothesis of depression. *J Clin Psychiatry*. 2000;6(Suppl 6):4-6.

72. Bertilsson L, Asberg M, Thorén P. Differential effect of chlorimipramine and nortriptyline on cerebrospinal fluid metabolites of serotonin and noradrenaline in depression. *Eur J Clin Pharmacol.* 1974;7:365-368.

73. Maas JW. Biogenic amines and depression. Biochemical and pharmacological separation of two types of depression. *Arch Gen Psychiatry*. 1975;32:1357-1361.

74. Stern SL, Rush AJ, Mendels J. Toward a rational pharmacotherapy of depression. *Am J Psychiatry*. 1980;137:545-552.

75. Carroll BJ, Curtis GC, Mendels J. Neuroendocrine regulation in depression. I. Limbic system-adrenocortical dysfunction. *Arch Gen Psychiatry*. 1976;33:1039-1044.

76. Carroll BJ, Curtis GC, Mendels J. Neuroendocrine regulation in depression. II. Discrimination of depressed from nondepressed patients. *Arch Gen Psychiatry*. 1976;33:1051-1058.

77. The APA Task Force on Laboratory Tests in Psychiatry. The dexamethasone suppression test: an overview of its current status in psychiatry. *Am J Psychiatry*. 1987;144:1253-1262.

78. Nierenberg AA, Feinstein AR. How to evaluate a diagnostic marker test. Lessons from the rise and fall of dexamethasone suppression test. *JAMA*. 1988;259:1699-1702.

79. American Psychiatric Association. Appendix. Practice guideline development process. American Psychiatric Association, 2006. Available at: http://www.psych.org/Departments/QIPS/Downloads/appendix_development%20process_04-28-06.aspx Accessed on August 15, 2011.

80. Swen JJ, Huizinga TW, Gelderblom H, et al. Translating pharmacogenomics: challenges on the road to the clinic. *PLoS Med*. 2007;4:1317-1324.

81. Kirchheiner J, Fuhr U, Brockmöller J. Pharmacogenetics-based therapeutic recommendations--ready for clinical practice? *Nat Rev Drug Discov*. 2005;4:639-647.

82. Stingl JC, Brockmöller J. Why, when, and how should pharmacogenetics be applied in clinical studies? Current and future approaches to study designs. *Clin Pharmacol Ther*. 2011;89,198-209.

83. Ozdemir V, Aklillu E, Mee S, et al. Pharmacogenetics for off-patent antipsychotics: reframing the risk for tardive dyskinesia and access to essential medicines. *Expert Opin Pharmacother*. 2006;7:119-133.

84. Shah RR. Can pharmacogenetics help rescue drugs withdrawn from the market? *Pharmacogenomics*. 2006;7:889-908.

85. Rahmioğlu N, Ahmadi KR. Classical twin design in modern pharmacogenomics studies. *Pharmacogenomics*. 2010;11:215-226.

86. Tan Q, Ohm Kyvik K, Kruse TA, et al. Dissecting complex phenotypes using the genomics of twins. *Funct Integr Genomics*. 2010;10:321-327.

87. Vesell ES. Pharmacogenetic perspectives gained from twin and family studies. *Pharmacol Ther*. 1989;41:535–552.

88. Angst MS, Phillips NG, Drover DR, et al. Opioid pharmacogenomics using a twin study paradigm: methods and procedures for determining familial aggregation and heritability. *Twin Res Hum Genet*. 2010;13:412-425.

89. Birkenfeld AL, Jordan J, Hofmann U, et al. Genetic influences on the pharmacokinetics of orally and intravenously administered digoxin as exhibited by monozygotic twins. *Clin Pharmacol Ther*. 2009;86:605-608.

90. Ozdemir V, Kalow W, Tothfalusi L, et al. Multigenic control of drug response and regulatory decision-making in pharmacogenomics: the need for an upper-bound estimate of genetic contributions. *Curr Pharmacogenomics Person Med*. 2005;3:53–71.

91. Dorne JL. Impact of inter-individual differences in drug metabolism and pharmacokinetics on safety evaluation. *Fundam Clin Pharmacol.* 2004;18:609-620.

92. Mant D. Can randomised trials inform clinical decisions about individual patients? *Lancet.* 1999;353:743-746.

93. Hampton JR. Evidence-based medicine, opinion-based medicine, and real-world medicine. *Perspect Biol Med.* 2002;45:549-568.

94. Smith GD, Egger M. Incommunicable knowledge? Interpreting and applying the results of clinical trials and meta-analyses. *J Clin Epidemiol.* 1998;51:289–295.

95. Yatham LN, Grossman F, Augustyns I, et al. Mood stabilisers plus risperidone or placebo in the treatment of acute mania. International, double-blind, randomised controlled trial. *Br J Psychiatry*. 2003;182:141-147.

96. de Leon J, Bork JA. Risperidone and the cytochrome P4503A (letter). *J Clin Psychiatry*. 1997;58:450.

97. Bork J, Rogers T, Wedlund P, et al. A pilot study of risperidone metabolism: the role of cytochrome P450 2D6 and 3A. *J Clin Psychiatry*. 1999;60:469-476.

98. Spina E, Avenoso A, Facciolà G, et al: Plasma concentrations of risperidone and 9-hydroxyrisperidone: effect of comedication with carbamazepine or valproate. *Ther Drug Monit*. 2000;22:481-485.

99. Fang J, Bourin M, Baker GB. Metabolism of risperidone to 9-hydroxyrisperidone by human cytochromes P450 2D6 and 3A4. *Naunyn Schmiedebergs Arch Pharmacol*. 1999;359:147-151.

100. Yasui-Furukori N, Hidestrand M, Spina E, et al. Different enantioselective 9-hydroxylation of risperidone by the two human CYP2D6 and CYP3A4 enzymes. *Drug Metab Dispos*. 2001;29:1263-1268.

101. Ebell MH, Siwek J, Weiss BD, et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *J Am Board Fam Pract*. 2004;17:59-67.

102. Kirchheiner J, Nickchen K, Bauer M, et al. Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol Psychiatry*. 2004;9:442-473.

103. Swen JJ, Nijenhuis M, de Boer A, et al. Pharmacogenetics: from bench to byte--an update of guidelines. *Clin Pharmacol Ther*. 2011;89:662-673.

104. Swen JJ, Wilting I, de Goede AL, et al. Pharmacogenetics: from bench to byte. *Clin Pharmacol Ther*. 2008;83:781-787.

105. Altar CA, Amakye D, Bounos D, et al. A prototypical process for creating evidentiary standards for biomarkers and diagnostics. *Clin Pharmacol Ther*. 2008;83:368-371.

106. Vandenbroucke JP. Observational research, randomised trials and two views of medical science. *PLoS Med*. 2008;5:339-343.

107. Vandenbroucke JP. What is the best evidence for determining harms of medical treatment? *CMAJ*. 2006;174:645-646.

108. Vandenbroucke JP, Psaty BM. Benefits and risks of drug treatments: how to combine the best evidence on benefits with the best data about adverse effects. *JAMA*. 2008;300:2417-2419.

109. Vandenbroucke JP. In defense of case reports and case series. *Ann Intern Med*. 2001;134:330-334.

110. Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. *Arch Intern Med*. 2009;169:1737-1739.

111. Issa AM, Phillips KA, Van Bebber S, et al. Drug withdrawals in the United States: a systematic review of the evidence and analysis of trends. *Curr Drug Saf*. 2007;2:177-185.

112. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Q.* 2004;82:661-687.

113. Cartwright N. Are RCTs the gold standard. *BioSocieties.* 2007;2:11-20.

114. Ioannidis JP. Why most published research findings are false. *PLoS Med.* 2005;2:e124.

115. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*. 2005;294:218-228.

116. Young NS, Ioannidis JP, Al-Ubaydli O. Why current publication practices may distort science. *PLoS Med*. 2008;5:e201.

117. Vandenbroucke JP, de Craen AJM. Alternative medicine: a "mirror image" for scientific reasoning in conventional medicine. *Ann Intern Med*. 2001;135:507-513.

118. Senn SJ. Overstating the evidence: double counting in meta-analysis and related problems. *BMC Med Res Methodol*. 2009;9:10.

119. Feist, GJ. *The psychology of science and the origins of the scientific mind*. New Haven: Yale University Press, 2006.

120. Matchar DB, Thakur ME, Grossman I, et al. Testing for cytochrome P450 polymorphisms in adults with non-psychotic depression treated with selective serotonin reuptake inhibitors (SSRIs). Evidence Report/Technology Assessment No. 146. (Prepared by the Duke Evidence-based

Practice Center under Contract No. 290-02-0025.) AHRQ Publication No. 07-E002. Rockville, MD: Agency for Healthcare Research and Quality. November 2006.

121. Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. *Genet Med*. 2007;9:819-825.

122. Teagarden JR. Warfarin and pharmacogenomic testing: what would Pascal do? *Pharmacotherapy.* 2009;29:245-247.

123. Pierce CS. *Essays in the philosophy of science*. New York: The Liberal Art Press, 1957.

124. Haack S. Trial and error: the Supreme Court's philosophy of science. *Am J Public Health*. 2005;95:S66–S73.

125. Polanyi M. *Personal knowledge: towards a post-critical philosophy*. Chicago: The University of Chicago Press, 1962.

126. Koshland DE. The cha-cha-cha theory of scientific discovery. *Science*. 2007;317:761-762.

127. Epstein RS, Teagarden JR. Comparative effectiveness research and PM: catalyzing or colliding? *Pharmacoeconomics*. 2010;28:905-913.

128. Moonesinghe R, Khoury MJ, Janssens AC. Most published research findings are false-but a little replication goes a long way. *PLoS Med*. 2007;4:e28.

129. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet.* 1995;345:840-842.

130. Waldrop M. *Complexity: the emerging science at the edge of order and chaos.* New York: Touchstone Books, 1992.

131. Mitchell M. *Complexity: a guided tour.* New York: Oxford University Press, 2009.

132. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet*. 2011;12:56-68.

133. McEachin RC, Saccone NL, Saccone SF, et al. Modeling complex genetic and environmental influences on comorbid bipolar disorder with tobacco use disorder. *BMC Med Genet*. 2010;11:14.

134. Lee SSJ. Defining statistical race and phenotypic race and their implications for health disparities. *Curr Pharmacogenomics Person Med*. 2009;7:238-242.

135. Baye TM. Inter-chromosomal variation in the pattern of human population genetic structure. *Hum Genomics*. 2011;5:220-240.

136. Badcott D. Scientific contribution. Some causal limitations of pharmacogenetic concepts. *Med Health Care Philos*. 2006;9:307-316.

137. Chakraborty B. Dynamic treatment regimes for managing chronic health conditions: a statistical perspective. *Am J Public Health.* 2011;101:40-45.

138. Bereiter C, Scardamalia M. *Surpassing ourselves. An inquiry in to the nature and implications expertise.* Chicago and La Salle, IL: Open Court, 1993.

139. Berrios GE, Markova IS. Conceptual issues. In D'haenen H, den Boer JA, Willner P eds. *Biological Psychiatry.* Chichester, John Wiley & Sons, Ltd., 2002: 1-24.

140. Hampton JR. Evidence-based medicine, opinion-based medicine, and real-world medicine. *Perspect Biol Med.* 2002;45:549-568.

141. Hirsch JE. An index to quantify an individual's scientific research output. *PNAS*. 2005;102:16569-16572.

142. Goldman, Gilbert M. The tacit dimension of clinical judgment. *Yale J Biol Med.* 1990;63:47-61.

143. Epstein RS, Teagarden JR. Comparative effectiveness research and personalized medicine: catalyzing or colliding? *Pharmacoeconomics*. 2010;28:905-913.

144. de Leon J, Greenlee B, Barber J, et al. Practical guidelines for the use of new generation antipsychotic drugs (except clozapine) in adult individuals with intellectual disabilities. *Res Dev Disabil.* 2009;30:613-669.

145. Garber AM, Tunis SR. Does comparative-effectiveness research threaten PM? *N Engl J Med*. 2009;360:1925-1927.

146. Frueh FW. Back to the future: why randomized controlled trials cannot be the answer to pharmacogenomics and personalized medicine. *Pharmacogenomics*. 2009;10:1077-1081.

147. Khoury MJ, Rich EC, Randhawa G, et al. Comparative effectiveness research and genomic medicine: an evolving partnership for 21st century medicine. *Genet Med*. 2009;11:707-711.

148. Woodcock J. The human genome and translational research: how much evidence is enough? *Health Aff*. 2008;27:1616-1618.

149. Miles A, Loughlin M, Polychronis A. Evidence-based healthcare, clinical knowledge and the rise of personalised medicine. *J Eval Clin Pract.* 2008;14:621-649.

150. Haynes B. ACP Journal Club. Editorial: new and underutilized features of ACP Journal Club PLUS and ACPJC.org: stellar articles, searches, and succinct synopses of the principles and practice of "personalized" evidence-based medicine. *Ann Intern Med*. 2009;151:JC6-2,JC6-3.

Table 1. Statistical assumptions on drug response underlying EBM and PM

	EBM	PM
Statistical homogeneity	Assumed to be present	Assumed to be absent
Statistical heterogeneity	Ignored	Assumed
Mean	Represents sample well	Does not represent sample well
Outliers	Ignored	Crucial
Stratification before randomization	Not needed	Ideal for studying outliers
Analyses for HTE	Not needed	Needed if stratification was not done

HTE: heterogeneity in treatment effects