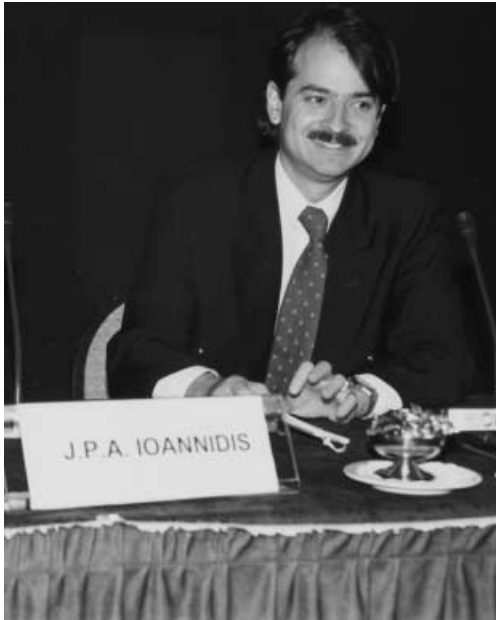


The ESCI Award for Excellence in Clinical Science 2007



Professor John Ioannidis was born in New York, USA, in 1965 and graduated from Athens College in Greece as the Valedictorian of the class of 1984, the highest honour of this prestigious school. He won several awards, including first prize in the Greek National Mathematics Competition and the John Vakis Award for Natural Sciences. He graduated at the top of his class from the School of Medicine, University of Athens, Greece, in 1990. He then did his residency training in Internal Medicine at Harvard's New England Deaconess Hospital, currently Beth Israel Deaconess Medical Centre, USA. In 1993–1996, he pursued fellowship training at the Division of Infectious Diseases and the Division of Clinical Care Research at Tufts-New England Medical Centre, Boston, MA, USA. After completing his fellowship, he moved to the National Institutes of Health, in a career position at the National Institute of Allergy and Infectious Diseases, Bethesda, MD, USA, with a joint appointment at Johns Hopkins Hospital and Johns Hopkins University School of Medicine, Baltimore, MD, USA. He has been Chairman of the Department of Hygiene and Epidemiology at the School of Medicine, University of Ioannina, Ioannina, Greece, since 1999 (as tenured Professor since 2003). He has also served in the adjunct faculty of Tufts University School of Medicine since 1996, with the rank of Professor

since 2002. He is a member of the editorial board of 12 international journals, including the *Lancet*.

Professor Ioannidis's publication record includes approximately 300 papers, which have received over 8000 citations. His work combines skills in novel clinical research methodology and evidence-based medicine with the challenges of current molecular medicine. He has made major contributions in the following areas:

1 Leading empirical research on clinical research evidence and its limitations; the concordance of various forms of evidence and small studies versus . large-scale evidence.

2 Making sense of evidence in the revolution of human genomics (large-scale evidence on genetic susceptibility for complex diseases, replication validity, racial descent differences, microarray research and discovery-oriented research, human genome epidemiology network and the roadmap for human genome epidemiology).

3 Evaluations of bias and credibility in research findings (time-lag bias, publication bias, reporting biases, empirical evidence on replication and contradiction of clinical research, challenges in translational medicine and modelling the credibility of research findings).

4 Development and leadership of international consortia of investigators in clinical research, especially in human genomics and beyond; meta-analyses of individual participant data; prospective meta-analysis.

5 Highlighting the importance of reporting and appraising harms of medical interventions (reporting of harms in clinical trials, large-scale evidence on harms and lead author of CONSORT for harms).

6 Extensive applications of evidence-based medicine across a wide-range of medical specialties, including meta-analyses of randomized trials, epidemiological studies, diagnostic evaluations and prognostic factors; randomized trials in AIDS treatment, antiretroviral resistance testing, anaemia in kidney disease, autoimmune diseases and antibiotic use.

Professor John Ioannidis has been an outstanding researcher of the methodology of biomedical research. He has both broad and deep knowledge of basic and clinical bioscience and an immensely analytical and deductive mind that has helped analyse and synthesize data creatively from diverse studies. His contribution to biomedical research is both unique and major. The ESCI is proud to bestow upon him the 2007 Clinical Investigator Award.

Prof. George P. Chrousos, MD, PhD,
Athens, Greece
E-mail: chrousge@med.uoa.gr

ESCI Award 2007

Molecular evidence-based medicine

Evolution and integration of information in the genomic era

J. P. A. Ioannidis

University of Ioannina School of Medicine, Greece; Biomedical Research Institute-Foundation for Research and Technology-Hellas, Ioannina, Greece; Tufts University School of Medicine, Boston, MA, USA

Abstract

Evidence-based medicine and molecular medicine have both been influential in biomedical research in the last 15 years. Despite following largely parallel routes to date, the goals and principles of evidence-based and molecular medicine are complementary and they should be converging. I define molecular evidence-based medicine as the study of medical information that makes sense of the advances of molecular biological disciplines and where errors and biases are properly appreciated and placed in context. Biomedical measurement capacity improves very rapidly. The exponentially growing mass of hypotheses being tested requires a new approach to both statistical and biological inference. Multidimensional biology requires careful exact replication of research findings, but indirect corroboration is often all that is achieved at best. Besides random error, bias remains a major threat. It is often difficult to separate bias from the spirit of scientific inquiry to force data into coherent and ‘significant’ biological stories. Transparency and public availability of protocols, data, analyses and results may be crucial to make sense of the complex biology of human disease and avoid being flooded by spurious research findings. Research efforts should be integrated across teams in an open, sharing environment. Most research in the future may be designed, performed, and integrated in the public cyberspace.

Keywords Evidence-based medicine, molecular medicine, replication, research, translation.
Eur J Clin Invest 2007; 37 (5): 340–349

Evidence-based medicine and molecular medicine

Both ‘evidence-based medicine’ and ‘molecular medicine’ are widely circulating terms in the literature. Both have their

This article is based on the lecture given at the European Society of Clinical Investigation Annual Meeting in Uppsala, Sweden, in April 2007 on the occasion of the 2007 ESCI Award for Excellence in Clinical Science.

Clinical and Molecular Epidemiology Unit and Evidence-Based Medicine and Clinical Trials Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Greece; Biomedical Research Institute-Foundation for Research and Technology-Hellas, Ioannina, Greece; Department of Medicine, Tufts University School of Medicine, Boston, USA (J. P. A. Ioannidis).

Correspondence to: John P.A. Ioannidis, Professor and Chairman, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece. Tel.: +302651097807; fax: +302651097867; e-mail: jioannid@cc.uoi.gr

Received 9 January 2007; accepted 23 January 2007

strong proponents and occasional, equally strong critics. A PubMed search (as of 27 December 2006) retrieves 23 957 items with ‘evidence-based medicine’, 32 751 with ‘evidence-based’ (it could be ‘evidence-based X’, i.e. evidence-based ‘anything’), and 660 853 with ‘evidence’. ‘Molecular medicine’ retrieves 18 025 items and ‘molecular’ alone retrieves way over a million (1 251 859). Clearly these are two powerful currents of thinking in the biomedical sciences.

The two currents nevertheless have had little overlap to date and they have been promoted by largely different circles. A search of the *Journal of Clinical Investigation*, *Cell*, *Journal of Experimental Medicine* and *Journal of Biological Chemistry* does not yield a single article with ‘evidence-based medicine’, while 1273 articles are retrieved in these same journals for ‘molecular medicine’. A search of the *British Medical Journal* and the *Journal of the American Medical Association* conversely yields 707 articles for ‘evidence-based medicine’ and only 15 for ‘molecular medicine’.

Both terms have conceptual roots that reach the distant past [1,2], but their mainstream emergence in the biomedical literature, in the way that we conceive them today, is only

Table 1 'Evidence-based X': titles of articles where *evidence-based* is used to accompany diverse entities

Analysis, appraisal, approach, assessment of the literature, clinical practice, clinical update, complementary medicine, conceptual framework, consensus process, data, decision making, dentistry, dermatology, design, discussion, emergency medicine, guidelines, health care architecture, health information, health information provision, health policy information, interventions, knowledge, management, medical education, medicines, model, nursing, pain management, patient choice, politics, practice, prevention programmes, problem solutions, psychotherapies, public health system, radiology, research, review, surgery, surgical rebuttal, systematic review, testing, therapies, update

Based on screening of the 100 latest entries in PubMed retrieved with a search for 'evidence-based' on December 27, 2006

about 15 years old. 'Evidence-based' medicine was first used as a term by Gordon Guyatt and the McMaster team in a *JAMA* paper in 1992 [3]. Interestingly, the same journal used the term 'molecular medicine' in a 1993 review [4] – actually an even earlier 'molecular medicine' title had appeared in *BMJ* in 1987 [5], well before wide spread to basic/translational biomedical journals.

'Evidence-based' has recently been invoked to accompany almost anything (Table 1) that seeks a touch of scientific justification and prestige. It has replaced authoritarian experts in the role of accredited guarantor of merit; although the term is often applied no more judiciously than expert dogma had been applied in the past [6]. The current paper, as most papers that support evidence-based medicine, is also unfortunately written by an expert – of sorts. Surprisingly, or even disappointingly, 'evidence-based molecular medicine' has not been used in the literature – PubMed does not recognize this phrase. Perhaps proponents of molecular medicine did not feel that the 'evidence-based' seal could enhance the credibility of their efforts. PubMed similarly does not recognize the phrase 'molecular evidence-based medicine'. Perhaps those who trusted evidence-based medicine did not feel at home with the complex deeds of molecular medicine. So why bring these two currents together then?

Wikipedia [7] claims that 'Evidence-based medicine applies the scientific method to medical practice. According to the Centre for Evidence-Based Medicine, "Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients"' [8]. But then, if this is about applying the scientific method and best evidence, can this be done without using, exploiting or integrating what the biomedical sciences are working on currently? I could not find a Wikipedia entry for 'molecular medicine', but arguably it encompasses all the efforts to apply to medicine the scientific principles and insights from advances in 'basic' molecular biological sciences. If this is so, then evidence-based medicine and molecular medicine have complementary aims.

What is the current best evidence?

In the 1990s, evidence-based medicine tried to develop explicit procedures and schemes ('hierarchies') for rating evidence. Despite variability [9–12], these hierarchies typically placed meta-analyses and randomized trials at the top. Non-randomized studies, even worse, uncontrolled studies, and isolated observations were placed at lower levels. Expert opinion was always at the bottom or not listed at all. Classic empirical studies also showed that experts were indeed not to be trusted [13]. Biological or molecular mechanisms were typically not mentioned at all, as if they had nothing to do with evidence.

It was soon realized that the type of design alone could not guarantee the credibility of a specific piece of clinical investigation [6]. While randomized studies are more reliable and more protected from bias compared with non-randomized studies, some randomized trials may be worse than respective non-randomized studies. When the two designs disagree, it is not always certain that randomized trials have found the right answer [14–16]. It also became apparent that appraisal of what is a good study can be considerably subjective, especially when studies are appraised after their completion [17]. This led to some strong controversies, as exemplified by the heated mammography debate [18,19].

Currently proposed grading systems focus more appropriately on a composite appraisal of the credibility and protection from bias in the accumulated evidence [20,21]. However, as we move from the level of interpreting the results of a single study or set to studies towards making a decision about clinical use, the consensus decreases steeply even when very knowledgeable clinician investigators and methodologists are involved [21]. Moreover, for each question of interest, what constitutes the best study design and best type of evidence may be different. There are many important questions that cannot be addressed with randomized trials – in fact, perhaps most questions of scientific and clinical interest cannot be addressed with randomized trials [6].

Unfortunately, this fruitful debate was disconnected from what was happening in the molecular medicine side. Molecular medicine was largely redefining the questions of interest in clinical investigation. A constant argument against evidence-based medicine has been that the average evidence, as derived from one or more clinical studies, cannot be applied to the individual patient. Clinicians want to treat individuals, not an average abstract phantom. Individuals vary enormously in the risk of disease, outcomes and treatment responses [22–26]. Molecular medicine in the past decade has taken the route of trying to achieve this individualization, pursuing an ideal of 'personalized medicine' [27,28]. New technologies stemming from the genomics revolution have made major promises in this regard [29–31]. However, it was not clearly realized that a main obstacle in obtaining reliable evidence from these new technologies was to tackle the typical errors and biases that evidence-based medicine was so sensitized to [32]. Not surprisingly, personalized medicine is not here yet [33].

The situation seems almost schizophrenic. While evidence-based medicine has been questioning the validity of well-designed mega-trials and large-scale meta-analyses of several thousand subjects [34,35], molecular medicine has been making certain promises from studies based on a few dozen samples where noise overwhelms true biological signals [36,37]. Conversely, while molecular medicine has entered the level of complexity where tens and hundreds of thousands of biological factors are measured, evidence-based medicine is still preoccupied with painfully appraising interventions, concepts and approaches that perhaps should have been abandoned long ago based on the wealth of information acquired in the biological sciences.

Molecular evidence-based medicine: information, errors, and biases

I define here molecular evidence-based medicine as the study of medical information that makes sense of the advances of molecular biological disciplines and where errors and biases are properly appreciated and placed in context. I prefer the term 'information', because 'evidence' already implies an appraisal of the strengths and weaknesses of the information, while 'knowledge' and 'science' are even more remote goals.

Information is accumulating exponentially in current research efforts, biomedical and beyond. Less than half a century ago, Bradford Hill, a father of modern epidemiology, claimed that no scientific paper was satisfactory unless an independent reader could check the results on the back of an envelope [38]. Currently, the databases that we can amass are stretching beyond the handling limits of software such as MS Excel, Mathematica or MATLAB. Raw data in some sciences have already reached the levels of petabytes (2 to the 50th power bytes of information) and there is no reason to believe that the pace of expansion will slow down [39,40]. For several problems in the physical sciences, splitting problems into many thousands of pieces and asking interested citizens to run them in their personal computers is the only viable solution. This resorting to amateur citizen-scientists may soon be true for the biomedical sciences as well.

There is no reason to worry about having so much data. Information is great news. If anything, the current information mass provides an obvious proof of how amazingly little we knew in the past, and how much we have to try to learn in the future, if the expansion continues. Large disciplines of the past become obsolete and we can only look back with a smile upon some of the claims we made even not so long ago [41]. However, the availability of so much data creates a challenge for meeting some of the basic prerequisites of the scientific endeavour. These are the need for transparency and availability of the information and the ability to understand and measure the errors and biases that are potentially embedded in it.

Many people think that science is about making discoveries. However, this discovery gold rush has resulted in a new

culture where 'publish or perish' has been replaced and/or enhanced by 'patent and prosper' [42]. Despite the extensive discoveries and geometrically increasing number of patents, nevertheless, really new and useful medical interventions and products for therapeutic, preventive, diagnostic, or predictive purposes are few [43–45]. Of 101 promises for clinical use of discoveries that were made in the top basic science journals between 1979 and 1983, only five were in clinical use and only one had made a major clinical impact 25 years later [43]. Most basic research and even clinical investigation seems to get nowhere, despite efforts to strengthen the translational interface [41,46]. Empirical evidence has shown that animal research has led to no successes in some fields such as acute stroke [47] or has often been refuted by clinical research in others [48]. On the clinical trials side, some of the best proponents of evidence-based medicine were justly lamenting recently why we have so few useful trials [49]. We have conducted half a million trials in the last half century, but empirical evaluations show that they remain laden with errors and biases and many of them tell us very little, or give wrong messages [50–52].

I have argued that in the current era, we have so many postulated discoveries that the main priority is to get rid of false discoveries, replicate and validate the few true ones, and move these to translation to benefit individual patients and the population at large [41,53,54]. This shifts attention from the production of still more tentative discoveries to the understanding of the errors and biases inherent in the research process. Understanding errors and biases is critical for making the right choices to discard the false and promote true research findings.

Mass of hypotheses and complex information: implications for errors

Let us use the traditional epidemiological terminology for chance (random) error that has no systematic component, i.e. results are not tilted more in one direction rather than the other, on average. Until recently, we were content to use statistical techniques that would make sure that 5% of our experiments might give false-positive significant results simply by chance. That was probably acceptable when there were few scientists and few experimental hypotheses being tested. It was also particularly acceptable when the hypotheses being tested had strong support from other lines of evidence, i.e. we were not searching in the dark, but simply reinforcing our appreciation for some research finding that we had strong reasons to believe in. Given this background, when we get a significant result in an experiment, the chances that it is truly significant rather than a spurious chance finding are high.

Two major changes have gradually occurred in biomedical investigation, with accelerated speed over the last decade. First, the number of scientific hypotheses that we can test has increased at an exponential pace and this increase has been further compounded by a rapid increase in the number of teams of researchers proposing and testing hypotheses.

To take one field alone, complex disease genetics, we are currently aware of 12 000 000 variants in the human genome. Testing half a million is a matter of routine already, while testing all of them, and more, will soon also be routine. The actual number of possible hypotheses in a single experiment where these polymorphisms are involved is not just 12 000 000. If we consider all the possible combinations among 12 000 000 binary variables, the total number of possible hypotheses in a single experiment is $2^{12\,000\,000}$, i.e. 4×10^{2085} . If I try to write this number in full expansion with all its zeros, it will take about the space of two double-spaced pages in a Word document on a PC and about half a page in a print journal.

Second, we have become far less selective in our choice of hypotheses to test. In many biomedical fields, we have even adopted techniques and methods that simply test en masse everything that can be tested rather than try to select upfront a few hypotheses that are more likely to yield fruit. After a long series of high-profile refutations in many fields ranging from molecular genetics to influential clinical trials [55–57], we realize that much of the epidemiology and pathophysiology inference machinery to select and filter hypotheses has probably not been working adequately – often it is not working at all. Causal inference and pathophysiological thinking may work for some very clear-cut situations, e.g. Mendelian inheritance mutations where the mutation causes the disease and the disease cannot exist without the mutation [58]; or major acquired disease risks such as smoking for lung cancer, where the risk calculations can indeed be made on the back of an envelope [59]. However, most of the biology that underlies human health and disease is likely to be extremely complex, multifactorial, and laden with weak effects [60,61]. The growth of systems biology reminds us of this complexity even when we try to assemble the pieces of the biological machinery, let alone see them work over time in a dynamical interactive fashion [62–64]. Empirical evaluations also suggest that when it comes to complex disease pathogenesis, biological plausibility does not square very well with epidemiological and clinical data [65,66].

The failure of many/most translational efforts to date may reflect that basic/preclinical investigations until now have made simplistic assumptions, as compared with the complex biology of human health and disease [67]. Studying one or a few biological parameters in experimental systems under-appreciates complex biological pathways. This may also underlie the failures of several ‘simple’ biological surrogate markers as clinical trial endpoints [68,69] and the relative dearth of evidence-based diagnostic and prognostic markers in the molecular era [70–73]. Now that we are flooded with millions of single biological factors, our insisting on one or another of them in the recent past seems unbelievably implausible, if not naive.

Current multidimensional (‘-omics’) approaches may address this criticism by diminishing the bulk of hypotheses-at-hand to a viable small number where composite systems of biological factors are seen as a package. Thousands and millions of biological factors can again be streamlined to relatively few packages. However, handling such complex

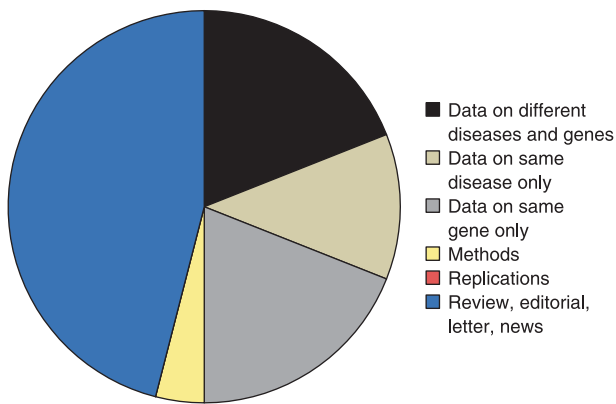
packages, avoiding biases, and translating them for practical purposes is very challenging [74–76]. Despite extreme interest, to date only one modern multidimensional application (tests for gene expression profiling for node-negative breast cancer) has moved into clinical practice for predictive purposes; even this one probably had less optimal independent validation than thought [77–79] and prospective results from clinical trials are still not available. Some promising multidimensional biological applications with microarrays or proteomics have been refuted on rigorous scrutiny [80–82]. Other multidimensional approaches such as genome-wide association studies still discover biological factors one at a time despite testing thousands thereof [83,84].

The extreme number of possible hypotheses suggests that in many fields perhaps the only viable way forward is to proceed with en masse testing, without making any effort of prefiltering with biological or other plausibility filters. As we do this, conventional levels of statistical significance make no sense [85]. Trying to correct for the number of hypotheses is also very difficult, since we often cannot even count very well the number of hypotheses that have been tested, and we still do not know how many other scientists are working on the same or similar hypotheses. So, no matter how low we set the *P*-value threshold, we may never be fully certain about the truth of a research finding.

Evolution of the scientific information: exact replication versus corroboration

Regardless of the complexity involved, if many teams run the same exact experiments and studies, then lack of replication will force the abandonment of the spurious false-positive claims. This is often true: rigorous, exact replication is a way to make safe progress. However, a prerequisite would be that exact replication does not also copy the errors that were possibly made in the first study [86]. Moreover, subgroup differences [87], experimental peculiarities, subtle modifications in a study may be invoked to transform lack of replication into spurious corroboration. Approximate corroboration is very frequently invoked in the biological sciences. Sometimes exact replication is indeed very difficult. In a recent *Nature* article, a famous scientist was claiming, perhaps rightly, that his results may be invalidated simply by moving to a different laboratory where the water in the pipelines would be different [88]. Generalizability and even more transposability of research findings in different experimental conditions and settings remains an open question from the basic sciences to large-scale pragmatic clinical trials [89,90].

Often exact replication has even been discredited as ‘me too’ poor-quality research. This is a misconception. In fact the problem with ‘me too’ research is not so much that people try to do the same thing as a previous team of scientists has done. Running a rigorous replication study can be a very demanding effort. The problem may be mostly that the replicating scientists are forced to convince their peers that they have done something different, thus new. In addition to statistical significance at all cost [91], novelty at all cost

(a) *TNFA* and malaria

(b) Genome-wide association findings for Parkinson disease

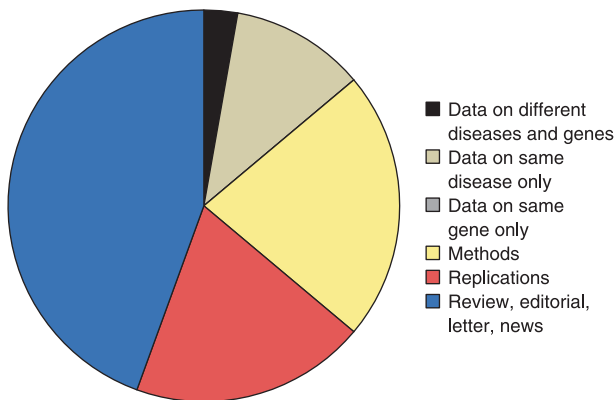


Figure 1 Types of articles citing (a) a *Nature* paper that had proposed in 1994 that a tumour necrosis factor alpha (*TNFA*) promoter variant confers susceptibility to cerebral malaria (100 first citations of the 792 received as of the end of 2006) and (b) an *American Journal of Human Genetics* paper that presented the results of a genome-wide association study in 2005 (first 36 citations as of late 2006).

is often considered a prerequisite for publication in many journals, especially the most prestigious ones [92]. Investigators may be forced to distort their analyses, outcomes, reporting, findings, or highlight spurious subgroups, unfounded interactions or peculiar exceptions simply to show that they have something novel to report.

There are many examples where exact replication has been left aside in favour of novelty-seeking. One of the first proposed high-profile associations between a polymorphism and a disease outcome was a 1994 *Nature* publication on a *TNFA* promoter variant conferring susceptibility to cerebral malaria [93]. By the end of 2006, this article had been cited 792 times (per Web of Science). One would expect that this would reflect extensive replication of the proposed association. However, an analysis of the first 100 citations that this paper received (covering citations up to late 1996) shows that not a single one of them tried to probe again the association of *TNFA* genetic variability and cerebral malaria. Of the 100

citing articles (Fig. 1a), 50 had no new data: they were reviews, hypotheses, editorials and letters. Another 19 dealt neither with malaria nor with *TNFA* genetic variability, 12 addressed malaria but not *TNFA* genetic variability, and 19 probed associations of *TNFA* genetic variability with various other conditions and phenotypes. These included in order of appearance type II diabetes mellitus, toxoplasma cyst burden and encephalitis, early onset pauciarticular juvenile chronic arthritis, mucocutaneous leishmaniasis, X-linked adrenoleukodystrophy phenotype diversity, multiple sclerosis, severe sepsis, differential tumour necrosis factor alpha production, insulin-dependent diabetes mellitus, inflammatory bowel disease, rheumatoid arthritis, coeliac disease, death from meningococcal disease and systemic lupus erythematosus – some of them studied more than once in various aspects and with 12 of these 19 studies proposing significant associations.

The proposing team subsequently also published on a different *TNFA* polymorphism that would modulate malarial outcomes [94], and also claimed that different alleles conferred susceptibility to severe anaemia from malaria versus cerebral malaria [95]. Independent teams then found no association with the original proposed polymorphism with either cerebral malaria or severe anaemia [96,97]. Thus, what was probably a false-positive finding, not only got entrenched in the literature, but it also lent citation support for probably over 100 other proposed associations, many/most of which are likely to also be spurious. Several other examples exist in the literature where lack of adherence to exact replication has created literature bodies where outcomes, biological factors, or both are not standardized [98,99] or the replication process has been incomplete or spurious [100,101].

Hopefully, the lack of exact replication is not the rule and it is possible that recognition of the need to replicate findings is increasing. As a comparative example, Fig. 1(b) shows the split of the first 36 citations received by the article on the first large-scale genome-wide association study of Parkinson's disease [102]. Within a few months of its publication, a series of studies were published trying to replicate, without success, its findings; within less than a year, a large collaborative replication effort was also published that did not replicate any of the proposed findings [103]. As shown (Fig. 1b), reviews and editorials still have their lion's share in shaping the literature, but real replications also have a prominent place. Another welcome emergence is the considerable proportion of methodology papers, in this specific example at least.

Bias and the spirit of scientific inquiry

I have only indirectly touched upon bias until now. In a world without bias, we simply have to deal with our chance and multiple testing problems. Tough as they might be, they might be manageable eventually. However, bias is a whole different story. Deviating a bit from conventional epidemiology nomenclature, I define bias here as anything, beyond chance

error, that can cause the appearance of significant research findings, while these do not really exist. Reverse bias is also possible, i.e. some true findings may be missed because of bias. However, there is probably a preponderance of bias over reverse bias in the research endeavour at large. Consciously, subconsciously, and unconsciously the quest for discovery is a quest for significant findings and thus bias is a way to get to this goal earlier and easier [104].

Much of the basic and clinical investigation of the recent past has made a principle of trying to read meaningful biological stories in the data. Understanding mechanisms and processes has required an abstractive mode of thinking where a biological story would emerge linking different aspects in the data. Scientific thinking was purposefully trained to remove selectively what did not fit in the picture and find, isolate, strengthen and highlight links and associations. If this is a good description of scientific thinking, then scientific thinking is trained par excellence to generate bias and find links and associations not only where they do exist, but also where they do not exist.

The increasing complexity of modern biomedical databases may be creating an intimidation barrier to anyone who wants to perform linking exercises with bare hands and bare brains on the back of an envelope. However, computer power has supplemented the required means for continuing these exercises in the face of increasing complexity. This includes text mining and connectivity approaches that try to isolate new and more comprehensive biological links in large-scale information [105–107]. Clearly, whatever emerges out of such complex processing requires further independent replication. However, given the complexity of the derived patterns, exact replication becomes a major challenge and approximate corroboration may be what can best be achieved with all the limitations discussed above.

Transparency and integration of information

Bias can occur within single studies and also in scientific fields at large. Measurements, data, analyses and results may be guided, or distorted, towards a postulated research finding in a single study. This has happened in the past and will continue to happen unavoidably. While some study designs are proud of their rigorous, inflexible adherence to protocols, others are trained by default upon data dredging and data mining [108]. Data dredging and data mining are often absolutely appropriate. What matters is that (1) they are acknowledged as such; (2) the data dredging and mining is transparent and other scientists can track the process; (3) the field contains complete and transparent information on other studies where the products of the data dredging and data mining can be tested for replication; and, ideally, (4) the complete information accumulated by all teams working on the same topic can be nonselectively integrated.

In the last decade, we have witnessed many efforts that try to maximize transparency on what exactly is being done within a research project. These efforts cover both evidence-based medicine (e.g. CONSORT for clinical trials

[109,110], STARD for diagnostic tests [111], REMARK for prognostic marker studies [112], and QUOROM for meta-analyses [113]); and molecular medicine initiatives (e.g. MIAME for microarrays experiments [114] and many other efforts in systems biology [115]). Most of these efforts have focused on comprehensive reporting of the study methods and results. However, after the fact, one may be unwilling to report the deficiencies of one's study design, data and analyses thereof. Forcing investigators to report on each aspect of their study may force white and not so white lies.

This problem can be bypassed only if the full protocols of scientific research are transparently available upfront before any experiments or measurements are done. For some other sciences, this sounds almost self-evident. For example, a spacecraft sent out to explore the galaxy beyond the solar system follows a specifically laid out plan on what information it is to collect and how and what the backup plans are in case of adversities and system failures. Moreover, unavoidable adjustments to the original plans are also thoroughly recorded with minute attention as the experiment unfolds. In biomedical research, this has not been so clear cut. Perhaps such upfront transparency seems to contradict the individualistic spirit of scientific discovery, where any bright person can surprise the establishment with his or her fresh ideas and results. One has to be very careful not to stifle creativity, independence and spontaneity. Research is not about bureaucrats who simply keep good records. It is about creative and imaginative people who, nevertheless, should still keep good records.

For a lot of biomedical research, upfront registration of protocols should be feasible. For example, there is no reason why this cannot happen for all randomized trials. Upfront trial registration is a very important initiative in this regard [116,117]. Nevertheless, 2 years after its adoption by the most influential medical journals, the majority of clinical trials are still not registered upfront. Even for those that are registered, registration does not mean that full and exact details are provided for the outcomes and anticipated analyses. Therefore, while simple registration obviates to some extent the problem of publication bias, it still leaves considerable room for selective analysis and outcome reporting [118,119]. Moreover, if the preference for significant results continues to guide the research process, the problem of time-lag bias remains unchallenged despite registration [120]. Early published results may be inflated [121–125] and one should be cautious to wait until the more complete picture emerges.

Many fields in molecular medicine have made major progress in realizing the importance of making all the protocol details and data available in public, even if this is done after the fact. The microarrays field is one such example, where sophisticated databases are already available for this purpose [126–129] and similar initiatives are also being launched for other multidimensional biological research. The availability of all methods and data can help integrate evidence efficiently and minimize selection bias across different studies on the same research question. However, the complexity of the methods and data make such efforts increasingly difficult [130,131]. Transparency is struggling to match complexity.

Table 2 Typical papers in the past, present, and near future

| Topic | Journal and year | Length | Methods (statistical) | Additional data and protocol available |
|--------------|---|-----------------------|-----------------------|--|
| DNA helix | <i>Nature</i> 1953 | 1 page | None (none) | None |
| Statin trial | <i>New England Journal of Medicine</i> 1996 | 9 pages | 0.5 page (0.1 page) | Not in public (selectively submitted to regulatory agencies) |
| Microarrays | <i>PLoS Medicine</i> 2006 | 10 pages | 1.5 pages (0.75 page) | Data and platforms fully available in public (over 1 gigabyte) Data on 177 samples, each one about 10 megabytes Data on 10 platforms, each one about 3 megabytes |
| Near future | <i>Soon</i> 20xx | Modified Over time | Very detailed | Data and platforms available in public Protocol available upfront Size of the data in the petabyte scale and larger |

Table 2 shows comparatively the amount of information on methods and data available in public, using representative examples of landmark papers. The discovery of the DNA double helix in 1953 is often applauded for its exceptional brevity: one single page sufficed to present the greatest discovery of the past century [132]. However, viewed from a different angle, that paper is devoid of methods, certainly has no statistics, and there are no data deposited for public view anywhere. The discovery seems like a sudden stroke of genius, and it is mostly indirectly from memoirs that we know about the intense interchange of data and information between Watson and Crick and the teams of Franklin and Pauling at that time – even to the point of gossip arising on who really had the major contribution to the discovery. More than four decades later, the landmark paper of the CARE trial that propelled the benefits of cholesterol reduction with a statin in myocardial infarction even in patients with average cholesterol [133] was nine pages long, but the methods ran at half a page and only a tenth of the page in this small print section was devoted to statistical methods. Now, compare a recent paper presenting the results of microarrays experimentation for a clinical predictive purpose [134]. The paper is ten pages long, but its methods run into much greater detail and it is also accompanied by a couple gigabytes of online-accessed, publicly available data.

The same table also shows an imaginary future paper. The paper may exist only online and may be updated over time with comments, response to criticism, and corrections [135]. The amount of information may be well beyond the capability of a current typical PC to handle. Data will be fully publicly available on the Web; access, data inspection, calculations and efforts to integrate these databases with other similar databases may be performed from a distance without any need to transfer the data to a particular computer terminal. Clinical and laboratory experiment output may be processed with sophisticated quality control systems and deposited in the public online database. ‘Running’ new studies may become largely a job for robots – the really important scientific activities are to meticulously design these studies and to find ways to integrate the information across many studies conducted worldwide. Protocols and data would be in public view with every process automatically recorded in the system. Such recording systems that capture

anything that is being done in the data or analyses are already available for some genomic organizations, so this is not really science fiction.

Does a transparent global integration system mean the demise of personal ingenuity and investigative talent? Until now, much scientific innovation has seemingly come out of nowhere, and this surprise part has been a great excitement about research: ‘unregistered’ scientists challenging and refuting old dogmas and opening new avenues. I think that integration and transparency of the scientific endeavour can only facilitate, not stifle, this ingenuity and creative spontaneity in the research process. This process resembles world history at large, where major events were often triggered and acted by people and factors unpredictable to outsiders of established civilizations that mistakenly thought that the world could not extend beyond their own. Now the whole world is known, isn’t it? – so we cannot be really surprised, perhaps history has ended [136]. Again and again, we have seen that this belief has been refuted and history has continued to surprise us. No doubt, research will also continue to surprise and fascinate us.

References

- Dickersin K, Straus SE, Bero LA. Evidence based medicine: increasing, not dictating, choice. *Br Med J* 2007;**334** (Suppl. 1):s10.
- Weatherall DJ. Towards molecular medicine; reminiscences of the haemoglobin field 1960–2000. *Br J Haematol* 2001;**115**:729–38.
- Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;**268**:2420–5.
- Caskey CT. Molecular medicine. A spin-off from the helix. *JAMA* 1993;**269**:1986–92.
- Craig RK. Methods in molecular medicine. *Br Med J* 1987;**295**:646–9.
- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn’t. *Br Med J* 1996;**312**:71–2.
- Wikipedia, ‘Evidence-based medicine’. Accessed at: http://en.wikipedia.org/wiki/Evidence-based_medicine. Last accessed December 30, 2006.

- 8 Oxford Centre for Evidence-Based Medicine, Glossary. Accessed at: <http://www.cebm.net/glossary.asp>. Last accessed December 20, 2006.
- 9 US Department of Health and Human Services, Public Health Service, Agency Health Care Policy and Research. Acute pain management: operative or medical procedures and trauma. AHCPR Publications, 92-0038. Rockville, MD: Agency for Health Care Policy and Research Publications;1992.
- 10 Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ *et al.* Users' guide to the medical literature IX: a method for grading health care recommendations. *JAMA* 1995;274:1800-4.
- 11 Scottish Intercollegiate Guidelines Network (SIGN). Forming guideline recommendations. In: *A Guideline Developers' Handbook*. Edinburgh: SIGN;2001: (Publication No. 50.).
- 12 US Preventive Services Task Force. *Guide to Clinical Preventive Services*, 2nd edn. Baltimore: Williams & Wilkins;1996.pp.xxxix-iv.
- 13 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 1992;268:240-8.
- 14 Ioannidis JP, Haidich AB, Lau J. Any casualties in the clash of randomised and observational evidence? *Br Med J* 2001;322:879-80.
- 15 Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG *et al.* Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001;286:821-30.
- 16 Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006;174:635-41.
- 17 Goodman SN. The mammography dilemma: a crisis for evidence-based medicine? *Ann Intern Med* 2002;137:363-5.
- 18 Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129-34.
- 19 Duffy SW, Tabar L, Smith RA. The mammographic screening trials: commentary on the recent work by Olsen and Gotzsche. *CA Cancer J Clin* 2002;52:68-71.
- 20 Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S *et al.* Grading quality of evidence and strength of recommendations. *Br Med J* 2004;328:1490.
- 21 Atkins D, Briss PA, Eccles M, Flottorp S, Guyatt GH, Harbour RT *et al.* Systems for grading the quality of evidence and the strength of recommendations II: pilot study of a new system. *BMC Health Serv Res* 2005;5:25.
- 22 Ioannidis JP, Lau J. The impact of high-risk patients on the results of clinical trials. *J Clin Epidemiol* 1997;50:1089-98.
- 23 Glasziou PP, Irwig LM. An evidence based approach to individualising treatment. *Br Med J* 1995;311:1356-9.
- 24 Ioannidis JP, Lau J. Heterogeneity of the baseline risk within patient populations of clinical trials: a proposed evaluation algorithm. *Am J Epidemiol* 1998;148:1117-26.
- 25 Trikalinos TA, Ioannidis JP. Predictive modeling and heterogeneity of baseline risk in meta-analysis of individual patient data. *J Clin Epidemiol* 2001;54:245-52.
- 26 Lau J, Ioannidis JP, Schmid CH. Summing up evidence: one answer is not always enough. *Lancet* 1998;351:123-7.
- 27 Liotta LA, Kohn EC, Petricoin EF. Clinical proteomics: personalized molecular medicine. *JAMA* 2001;286:2211-4.
- 28 Collins FS. Shattuck lecture - medical and societal consequences of the Human Genome Project. *N Engl J Med* 1999;341:28-37.
- 29 Guttmacher AE, Collins FS. Welcome to the genomic era. *N Engl J Med* 2003;349:996-8.
- 30 Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research. *Nature* 2003;422:835-47.
- 31 Collins FS, Guttmacher AE. Genetics moves into the medical mainstream. *JAMA* 2001;286:2322-4.
- 32 Ransohoff DF. Bias as a threat to the validity of cancer molecular-marker research. *Nat Rev Cancer* 2005;5:142-9.
- 33 Haga SB, Khoury MJ, Burke W. Genomic profiling to promote a healthy lifestyle: not ready for prime time. *Nat Genet* 2003;34:347-50.
- 34 Ioannidis JP, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. *JAMA* 1998;279:1089-93.
- 35 Furukawa TA, Streiner DL, Hori S. Discrepancies among megatrials. *J Clin Epidemiol* 2000;53:1193-9.
- 36 Frantz S. An array of problems. *Nat Rev Drug Discov* 2005;4:362-3.
- 37 Ioannidis JP. Microarrays and molecular research: noise discovery? *Lancet* 2005;365:454-5.
- 38 Shapiro S. Looking to the 21st century: have we learned from our mistakes, or are we doomed to compound them? *Pharmacoepidemiol Drug Saf* 2004;13:257-65.
- 39 Brent R, Bruck J. 2020 computing: can computers help to explain biology? *Nature* 2006;440:416-7.
- 40 Muggleton SH. 2020 computing: exceeding human limits. *Nature* 2006;440:409-10.
- 41 Ioannidis JP. Evolution and translation of research findings: from bench to where? *PLoS Clin Trials* 2006;1:e36.
- 42 Schachman HK. From 'publish or perish' to 'patent and prosper'. *J Biol Chem* 2006;281:6889-903.
- 43 Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. *Am J Med* 2003;114:477-84.
- 44 Crowley WF Jr, Sherwood L, Salber P, Scheinberg D, Slavkin H, Tilson H *et al.* Clinical research in the United States at a crossroads: proposal for a novel public-private partnership to establish a national clinical research enterprise. *JAMA* 2004;291:1120-6.
- 45 Cuatrecasas P. Drug discovery in jeopardy. *J Clin Invest* 2006;116:2837-42.
- 46 Zerhouni EA. Translational and clinical science - time for a new vision. *N Engl J Med* 2005;353:1621-3.
- 47 O'Collins VE, Macleod MR, Donnan GA, Horkey LL, van der Worp BH, Howells DW. 1026 experimental treatments in acute stroke. *Ann Neurol* 2006;59:467-77.
- 48 Hackam DG, Redelmeier DA. Translation of research evidence from animals to humans. *JAMA* 2006;296:1731-2.
- 49 Zwarenstein M, Oxman A; Pragmatic Trials in Health Care Systems (PRACTIHC). Why are so few randomized trials useful, and what can we do about it? *J Clin Epidemiol* 2006;59:1125-6.
- 50 Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. *JAMA* 1998;279:319-20.
- 51 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;365:1159-62.
- 52 Gluud LL. Bias in clinical intervention research. *Am J Epidemiol* 2006;163:493-501.
- 53 Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2:e124.
- 54 Ioannidis JP. Contradictions in highly cited clinical research. *JAMA* 2005;294:2696.
- 55 Ioannidis JP. Genetic associations: false or true? *Trends Mol Med* 2003;9:135-8.

- 56 Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001;**29**:306–9.
- 57 Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005;**294**:218–28.
- 58 Antonarakis SE, Beckmann JS. Mendelian disorders deserve more attention. *Nat Rev Genet* 2006;**7**:277–82.
- 59 Doll R. Tobacco: a medical history. *J Urban Health* 1999;**76**:289–313.
- 60 Khoury MJ, Little J, Gwinn M, Ioannidis JP. On the synthesis and interpretation of consistent but weak gene–disease associations in the era of genome-wide association studies. *Int J Epidemiol* 2006 December 20 [Epub ahead of print].
- 61 Ioannidis JP. Commentary: grading the credibility of molecular evidence for complex diseases. *Int J Epidemiol* 2006;**35**:572–8.
- 62 Kitano H. Computational systems biology. *Nature* 2002;**420**:206–10.
- 63 Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. *Science* 2004;**306**:640–3.
- 64 Pennisi E. How will big pictures emerge from a sea of biological data? *Science* 2005;**309**:94.
- 65 Ioannidis JP, Kavvoura FK. Concordance of functional in vitro biological data with epidemiological associations for complex diseases. *Genet Med* 2006;**8**:583–93.
- 66 Jais PH. How frequent is altered gene expression among susceptibility genes to human complex disorders? *Genet Med* 2005;**7**:83–96.
- 67 Ioannidis JP. Materializing research promises: opportunities, priorities and conflicts in translational medicine. *J Transl Med* 2004;**2**:5.
- 68 Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;**125**:605–13.
- 69 De Gruttola V, Fleming T, Lin DY, Coombs R. Perspective: validating surrogate markers – are we being naive? *J Infect Dis* 1997;**175**:237–46.
- 70 Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer* 2005;**5**:845–56.
- 71 Dalton WS, Friend SH. Cancer biomarkers-an invitation to the table. *Science* 2006;**312**:1165–8.
- 72 Bossuyt PM. The quality of reporting in diagnostic test research: getting better, still not optimal. *Clin Chem* 2004;**50**:465–6.
- 73 Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6.
- 74 Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. *Nat Rev Cancer* 2004;**4**:309–14.
- 75 Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 2006;**7**:55–65.
- 76 Boguski MS, McIntosh MW. Biomedical informatics for proteomics. *Nature* 2003;**422**:233–7.
- 77 Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M *et al*. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;**351**:2817–26.
- 78 Paik S, Tang G, Shak S, Kim C, Baker J, Kim W *et al*. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;**24**:3726–34.
- 79 Ioannidis JPA. Gene expression profiling for individualized breast cancer chemotherapy: success – or not? *Nat Clin Pract Oncol* 2006;**3**:538–9.
- 80 Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;**365**:488–92.
- 81 Ioannidis JPA. Is molecular profiling ready for clinical decision making? *Oncologist* 2007.
- 82 Baggerly KA, Morris JS, Edmonson SR, Coombes KR. Signal in noise: evaluating reported reproducibility of serum proteomic tests for ovarian cancer. *J Natl Cancer Inst* 2005;**97**:307–9.
- 83 Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C *et al*. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;**308**:385–9.
- 84 Todd JA. Statistical false positive or true disease pathway? *Nat Genet* 2006;**38**:731–3.
- 85 Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004;**96**:434–42.
- 86 Rosenbaum PM. Replication effects and biases. *Am Statistician* 2001;**55**:223–7.
- 87 Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992;**116**:78–84.
- 88 Giles J. The trouble with replication. *Nature* 2006;**442**:344–7.
- 89 Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials* 2006;**1**:e9.
- 90 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**:515–24.
- 91 Ioannidis JPA. Journals should publish all 'null' results and should sparingly publish 'positive' results. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:186.
- 92 Ioannidis JPA. Limitations are not properly acknowledged in the scientific literature. *J Clin Epidemiol* 2007;**60**:324–9.
- 93 McGuire W, Hill AV, Allsopp CE, Greenwood BM, Kwiatkowski D. Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria. *Nature* 1994;**371**:508–10.
- 94 Knight JC, Udalova I, Hill AV, Greenwood BM, Peshu N, Marsh K *et al*. A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. *Nat Genet* 1999;**22**:145–50.
- 95 McGuire W, Knight JC, Hill AV, Allsopp CE, Greenwood BM, Kwiatkowski D. Severe malarial anemia and cerebral malaria are associated with different tumor necrosis factor promoter alleles. *J Infect Dis* 1999;**179**:287–90.
- 96 Cabantous S, Doumbo O, Ranque S, Poudiougou B, Traore A, Hou X *et al*. Alleles 308A and 238A in the tumor necrosis factor alpha gene promoter do not increase the risk of severe malaria in children with *Plasmodium falciparum* infection in Mali. *Infect Immun* 2006;**74**:7040–2.
- 97 Hananantachai H, Patarapotikul J, Looareesuwan S, Ohashi J, Naka I, Tokunaga K. Lack of association of –308A/G TNFA promoter and 196R/MTNFR2 polymorphisms with disease severity in Thai adult malaria patients. *Am J Med Genet* 2001;**102**:391–2.
- 98 Contopoulos-Ioannidis DG, Alexiou GA, Gouvas TC, Ioannidis JP. An empirical evaluation of multifarious outcomes in pharmacogenetics: beta-2 adrenoceptor gene polymorphisms in asthma treatment. *Pharmacogenet Genomics* 2006;**16**:705–11.
- 99 Mutsuddi M, Morris DW, Waggoner SG, Daly MJ, Scolnick EM, Sklar P. Analysis of high-resolution HapMap

- of DTNBP1 (dysbindin) suggests no consistency between reported common variant associations and schizophrenia. *Am J Hum Genet* 2006;**79**:903–9.
- 100 Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;**95**:14–8.
 - 101 Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;**19**:453–73.
 - 102 Maraganore DM, de Andrade M, Lesnick TG, Strain KJ, Farrer MJ, Rocca WA *et al.* High resolution whole-genome association study of Parkinson disease. *Am J Hum Genet* 2005;**77**:685–93.
 - 103 Elbaz A, Nelson LM, Payami H, Ioannidis JP, Fiske BK, Annesi G *et al.* Lack of replication of thirteen single-nucleotide polymorphisms implicated in Parkinson's disease: a large-scale international study. *Lancet Neurol* 2006;**5**:917–23.
 - 104 Ioannidis JP. Molecular bias. *Eur J Epidemiol* 2005;**20**:739–45.
 - 105 Jensen LJ, Saric J, Bork P. Literature mining for the biologist: from information retrieval to biological discovery. *Nat Rev Genet* 2006;**7**:119–29.
 - 106 Muller HM, Kenny EE, Sternberg PW. Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biol* 2004;**2**:e309.
 - 107 Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;**313**:1929–35.
 - 108 Michels KB, Rosner BA. Data trawling: to fish or not to fish. *Lancet* 1996;**348**:1152–3.
 - 109 Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D *et al.* The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;**134**:663–94.
 - 110 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K *et al.* Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;**141**:781–8.
 - 111 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM *et al.* The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003;**49**:7–18.
 - 112 McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst* 2005;**97**:1180–4.
 - 113 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* 1999;**354**:1896–900.
 - 114 Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C *et al.* Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001;**29**:365–71.
 - 115 Brazma A, Krestyaninova M, Sarkans U. Standards for systems biology. *Nat Rev Genet* 2006;**7**:593–605.
 - 116 De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R *et al.* Clinical trial registration: a Statement from the International Committee of Medical Journal Editors. *Lancet* 2004;**364**:911–2.
 - 117 Rennie D. Trial registration: a great idea switches from ignored to irresistible. *JAMA* 2004;**292**:1359–62.
 - 118 Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *Br Med J* 2005;**330**:753.
 - 119 Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;**291**:2457–65.
 - 120 Ioannidis JPA. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 1998;**279**:281–6.
 - 121 Jennions MD, Moeller AP. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proc R Soc Lond B* 2002;**269**:43–8.
 - 122 Hauben M, Reich L, Van Puijenbroek EP, Gerrits CM, Patadia VK. Data mining in pharmacovigilance: lessons from phantom ships. *Eur J Clin Pharmacol* 2006;**62**:967–70.
 - 123 Trikalinos TA, Churchill R, Ferri M, Leucht S, Tuunainen A, Wahlbeck K *et al.* Effect sizes in cumulative meta-analyses of mental health randomized trials evolved over time. *J Clin Epidemiol* 2004;**57**:1124–30.
 - 124 Goring HH, Terwilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 2001;**69**:1357–69.
 - 125 Ioannidis JP, Trikalinos TA. Early extreme contradictory estimates may appear in published research: the Proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol* 2005;**58**:543–9.
 - 126 Ball CA, Brazma A, Causton H, Chervitz S, Edgar R, Hingamp P *et al.* Submission of microarray data to public repositories. *PLoS Biol* 2004;**2**:e317.
 - 127 Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization assay repository. *Nucl Acids Res* 2002;**30**:207–10.
 - 128 Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N *et al.* Array Express – a public repository for microarray gene expression data at the EBI. *Nucl Acids Res* 2003;**31**:68–71.
 - 129 Philippi S, Kohler J. Addressing the problems with life-science databases for traditional uses and systems biology. *Nat Rev Genet* 2006;**7**:482–8.
 - 130 Larsson O, Sandberg R. Lack of correct data format and comparability limits future integrative microarray research. *Nat Biotechnol* 2006;**24**:1322–3.
 - 131 Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D *et al.* Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* 2004;**101**:9309–14.
 - 132 Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 1953;**171**:737–8.
 - 133 Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG *et al.* The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N Engl J Med* 1996;**335**:1001–9.
 - 134 Zhao H, Ljungberg B, Grankvist K, Rasmuson T, Tibshirani R, Brooks JD. Gene expression profiling predicts survival in conventional renal cell carcinoma. *PLoS Med* 2006;**3**:e13.
 - 135 Shear M, Cave R, Uman R, Aquino S, Brown E, Burkes D *et al.* PLoS ONE sandbox: a place to learn and play. *PLoS ONE* 2006;**1**:e0.
 - 136 Fukuyama F. *The End of History and the Last Man*. New York: Free Press; 1992.