Evidence based medicine: a movement in crisis?

Trisha Greenhalgh and colleagues argue that, although evidence based medicine has had many benefits, it has also had some negative unintended consequences. They offer a preliminary agenda for the movement’s renaissance, refocusing on providing useable evidence that can be combined with context and professional expertise so that individual patients get optimal treatment.

Two decades of enthusiasm and funding have produced numerous successes for evidence based medicine. An early example was the British Thoracic Society’s 1990 asthma guidelines, developed through consensus but based on a combination of randomised trials and observational studies. Subsequently, the use of personal care plans and step wise prescription of inhaled steroids for asthma increased, and morbidity and mortality fell. More recently, uptake of the UK National Institute for Health and Care Excellence guidelines for prevention of venous thromboembolism after surgery has produced significant reductions in thromboembolic complications.

Despite these and many other successes, wide variation in implementing evidence based practice remains a problem. For example, the incidence of arthroscopic washout of the knee joint, whose benefits are unproved except when there is a known loose body, varies from 3 to 48 per 100 000 in England. More fundamentally, many who support evidence based medicine in principle have argued that the movement is now facing a serious crisis (box 1).

Below we set out the problems and suggest some solutions.

Distortion of the evidence based brand

The first problem is that the evidence based “quality mark” has been misappropriated and distorted by vested interests. In particular, the drug and medical devices industries increasingly set the research agenda. They define what counts as disease (for example, female sexual arousal disorder, treatable with sildenafil) and predisease “risk states” (such as low bone density, treatable with alendronate). They also decide which tests and treatments will...
be compared in empirical studies and choose (often surrogate) outcome measures for establishing “efficacy.”

Furthermore, by overpowering trials to ensure that small differences will be statistically significant, setting inclusion criteria to select those most likely to respond to treatment, manipulating the dose of both intervention and control drugs, using surrogate endpoints, and selectively publishing positive studies, industry may manage to publish its outputs as “unbiased” studies in leading peer reviewed journals. Use of these kinds of tactic in studies of psychiatric drugs sponsored by their respective manufacturers enabled them to show that drug A outperformed drug B, which outperformed drug C, which in turn outperformed drug A. One review of industry sponsored trials of antidepressants showed that 37 of 38 with positive findings, but only 14 of 36 with negative findings, were published.

Evidence based medicine’s quality checklists and risk of bias tools may be unable to detect the increasingly subtle biases in industry sponsored studies. Some so called evidence based policies (such as dementia case finding for the over 75s and universal health checks for the over 40s in the UK) seem to be based largely on political conviction. Critics have condemned the role of the drug industry in influencing the policy makers who introduced them.

**Too much evidence**

The second aspect of evidence based medicine’s crisis (and yet, ironically, also a measure of its success) is the sheer volume of evidence available. In particular, the number of clinical guidelines is now both unmanageable and unfathomable. One 2005 audit of a 24 hour medical take in an acute hospital, for example, included 18 patients with 44 diagnoses and identified 3679 pages of national guidelines (an estimated 122 hours of reading) relevant to their immediate care.

**Marginal gains and a shift from disease to risk**

Evidence based medicine is, increasingly, a science of marginal gains—since the low hanging fruit (interventions that promise big improvements) for many conditions were picked long ago. After the early big gains of highly active antiretroviral therapy for HIV and triple therapy for Helicobacter pylori positive peptic ulcer, contemporary research questions focus on the marginal gains of whether these drug combinations should be given in series or in parallel and how to increase the proportion of patients who take their complex medication regimen as directed.

Large trials designed to achieve marginal gains in a near saturated therapeutic field typically overestimate potential benefits (because trial samples are unrepresentative and, if the trial is overpowered, effects may be statistically but not clinically significant) and underestimate harms (because adverse events tend to be underdetected or underreported). The 74 year old who is put on a high dose statin because the clinician applies a fragment of a guideline uncritically and who, as a result, develops muscle pains that interfere with her hobbies and ability to exercise, is a good example of the evidence based tail wagging the clinical dog. In such scenarios, the focus of clinical care shifts insidiously from the patient (this 74 year old woman) to the population subgroup (women aged 70 to 75) and from ends (what is the goal of investigation or treatment in this patient?) to means (how can we ensure that everyone in a defined denominator population is taking statins?).

As the examples above show, evidence based medicine has drifted in recent years from investigating and managing established disease to detecting and intervening in non-diseases. Risk assessment using “evidence based” scores and algorithms (for heart disease, diabetes, cancer, and osteoporosis, for example) now occurs on an industrial scale, with scant attention to the opportunity costs or unintended human and financial consequences.

**Overemphasis on following algorithmic rules**

Well intentioned efforts to automate use of evidence through computerised decision support systems, structured templates, and point of care prompts can crowd out the local, individualised, and patient initiated elements of the clinical consultation. For example, when a clinician is following a template driven diabetes check-up, serious non-diabetes related symptoms that the patient mentions in passing may not be documented or acted on. Inexperienced clinicians may (partly through fear of litigation) engage mechanically and defensively with decision support technologies, stifling the development of a more nuanced clinical expertise that embraces accumulated practical experience, tolerance of uncertainty, and the ability to apply practical and ethical judgment in a unique case.

Templates and point of care prompts also contribute to the creeping managerialism and politicisation of clinical practice. As Harrison and Checkland observe: “As the language of EBM becomes ever more embedded in medical practice, and as bureaucratic rules become the accepted way to implement ‘the best’ evidence, its requirements for evidence are quietly attenuated in favour of an emphasis on rules.”

For example, the Quality and Outcomes Framework (QOF) in UK general practice is incentivised by financial “quality points” and administered largely by non-clinical staff who generate these points by recalling patients for structured reviews and checks. QOF has been associated with significant improvements in blood pressure control, especially in deprived populations. But its downside is an audit driven, technocratic exercise in which few patients are offered personalised shared decision making with a senior clinician before having the recommended tests and treatments, and in which clinical consultations are continually interrupted by pop-up point of care prompts.
Poor fit for multimorbidity

Finally, as the population ages and the prevalence of chronic degenerative diseases increases, the patient with a single condition that maps unproblematically to a single evidence based guideline is becoming a rarity. Even when primary studies were designed to include participants with multiple conditions, applying their findings to patients with particular comorbidities remains problematic. Multimorbidity (a single condition only in name) affects every person differently and seems to defy efforts to produce or apply objective scores, metrics, interventions, or guidelines. Increasingly, the evidence based management of one disease or risk state may cause or exacerbate another—most commonly through the perils of polypharmacy in the older patient.

Return to real evidence based medicine

To address the above concerns, we believe it is time to launch a campaign for real evidence based medicine (box 2).

Individualised for the patient

Real evidence based medicine has the care of individual patients as its top priority, asking, “what is the best course of action for this patient, in these circumstances, at this point in their illness or condition?” It consciously and reflexively refuses to let process (doing tests, prescribing medicines) dominate outcomes (the agreed goal of management in an individual case). It engages with an ethical and existential agenda (how should we live? when should we accept death?) and with that goal in mind, carefully distinguishes between whether to investigate, treat, or screen and how to do so.

To support such an approach, evidence must be individualised for the patient. This requires that research findings be expressed in ways that most people will understand (such as the number needed to treat, number needed to harm, and number needed to screen) and that practitioners, together with their patients, are free to make appropriate care decisions that may not match what “best (average) evidence” seems to suggest.

Importantly, real shared decision making is not the same as taking the patient through a series of if-then decision options. Rather, it involves finding out what matters to the patient—what is at stake for them—and making judicious use of professional knowledge and status (to what extent, and in what ways, does this person want to be “empowered?”) and introducing research evidence in a way that informs a dialogue about what best to do, how, and why. This is a simple concept but by no means easy to deliver. Tools that contain quantitative estimates of risk and benefit are needed, but they must be designed to support conversations not climb probability trees.

 Judgment not rules

Real evidence based medicine is not bound by rules. The Dreyfus brothers have described five levels of learning, beginning with the novice who learns the basic rules and applies them mechanically with no attention to context. The next two stages involve increasing depth of knowledge and sensitivity to context when applying rules. In the fourth and fifth stages, rule following gives way to expert judgments, characterised by rapid, intuitive reasoning informed by imagination, common sense, and judiciously selected research evidence and other rules.

In clinical diagnosis, for example, the novice clinician works methodically and slowly through a long and standardised history, exhaustive physical examination, and (often numerous) diagnostic tests. The expert, in contrast, makes a rapid initial differential diagnosis through intuition, then uses a more selective history, examination, and set of tests to rule in or rule out particular possibilities. To equate “quality” in clinical care with strict adherence to guidelines or protocols, however robust these rules may be, is to overlook the evidence on the more sophisticated process of advanced expertise.

Aligned with professional, relationship based care

Real evidence based medicine builds (ideally) on a strong interpersonal relationship between patient and clinician. It values continuity of care and empathetic listening, especially for people who are seriously and incurably sick. Research evidence may still be key to making the right decision—but it does not determine that decision. Clinicians may provide information, but they are also trained to make ethical and technical judgments, and they hold a socially recognised role to care, comfort, and bear witness to suffering. The challenges of self management in severe chronic illness, for example, are not merely about making treatment choices but about the practical and emotional work of implementing those choices.

As serious illness is lived, evidence based guidelines may become irrelevant, absurd, or even harmful (most obviously, in terminal illness).

Public health dimension

Although we have focused on individual clinical care, there is also an important evidence base relating to population level interventions aimed at improving public health (such as pricing and labelling of consumables, fluoridation of water, and sex education). These are often complex, multifaceted programmes with important ethical and practical dimensions, but the same principles apply as in clinical care. Success of interventions depends on local feasibility, acceptability, and fit with context—and hence on informed, shared decision making with and by local communities, using summaries and visualisations of population level metrics.

Delivering real evidence based medicine

To deliver real evidence based medicine, the movement’s stakeholders must be proactive and persistent. Patients (for whose care the movement exists) must demand better evidence, better presented, better explained, and applied in a more personalised way with sensitivity to context and individual goals. There are already some models of good practice here. In arthritis, for example, patient advocacy groups that emphasise the importance of experiential evidence and patient centred strategies have existed for over 30 years and have influenced the choice of outcome measures used in comparative effectiveness studies. Patient input has refocused several NICE guidelines (for example, on psoriasis).

Third sector advisory and advocacy groups such as the UK’s Consumer Association (www.which.co.uk), Picker Institute (www.pickereurope.org), and Sense About Science (www.senseaboutscience.org) have a crucial role in educating citizens and contributing to public debate about the use and abuse of evidence. The James Lind Alliance (www.lindalliance.org) brings patients, carers, and clinicians together to prioritise research questions. Such groups must remain, as far as possible, independent of vested interests and aware of the distorting influence of tied funding.
Critical appraisal skills—including basic numeracy, electronic database searching, and the ability systematically to ask questions of a research study—are prerequisites for competence in evidence based medicine. But clinicians need to be able to apply them to real case examples.

Training must be reoriented from rule following

Evidence users include clinicians and patients of varying statistical literacy, many of whom have limited time or inclination for the small print. Different approaches such as brief, plain language summaries for the non-expert (as offered by NICE), visualisations, infographics, option grids, and other decision aids should be routinely offered and widely used. Yet currently, only a fraction of the available evidence is presented in usable form, and few clinicians are aware that such usable shared decision aids exist.

Publishers must raise the bar

Evidence users include clinicians and patients of varying statistical literacy, many of whom have limited time or inclination for the small print. Different approaches such as brief, plain language summaries for the non-expert (as offered by NICE), visualisations, infographics, option grids, and other decision aids should be routinely offered and widely used. Yet currently, only a fraction of the available evidence is presented in usable form, and few clinicians are aware that such usable shared decision aids exist.

Evidence must be usable as well as robust

Another precondition for real evidence based medicine is that those who produce and summarise research evidence must attend more closely to the needs of those who might use it. Lengthy and expensive reviews that are “methodologically robust” but unusable in practice often fail to inform, inspire, or influence. A recent systematic review of diabetes risk scores revealed that the authors of most studies were primarily concerned with the intellectual concept of improving the predictive value of the score but had given little or no thought to how their score might be used, by whom, or for what—nor what the implications would be for real people who would be designated “at risk” by the score.

Evidence users include clinicians and patients of varying statistical literacy, many of whom have limited time or inclination for the small print. Different approaches such as brief, plain language summaries for the non-expert (as offered by NICE), visualisations, infographics, option grids, and other decision aids should be routinely offered and widely used. Yet currently, only a fraction of the available evidence is presented in usable form, and few clinicians are aware that such usable shared decision aids exist.

Research must transcend conflicts of interest

To support real evidence based medicine, and in particular to reassure policy makers, clinicians, and the public that research and the guidance derived from it can be trusted, the infrastructure for research and guideline development must show the highest standards of probity. Independent funding of national bodies for medical research is crucial.
Broader, more imaginative research is needed

The research agenda for real evidence based medicine is much broader than critical appraisal and draws on a wider range of underpinning disciplines. For example, it should include the study of the patient’s experience of illness and the real life clinical encounter for different conditions and in different circumstances. The field would be enriched, for example, by qualitative research to elucidate the logic of care—that is, the numerous elements of good illness management that are complementary to the application of research evidence.4

We need to gain a better understanding (perhaps beginning with a synthesis of the cognitive psychology literature) of how clinicians and patients find, interpret, and evaluate evidence from research studies, and how (and if) these processes feed into clinical communication, exploration of diagnostic options, and shared decision making.3 Deeper study is also needed into the less algorithmic components of clinical method such as intuition and heuristic reasoning, and how evidence may be incorporated into such reasoning.5

In relation to producing usable evidence, we need to identify how to balance gold standard systematic reviews with pragmatic, rapid reviews that gain in timeliness and accessibility what they lose in depth and detail.6 In the same vein, we need research on how and in what circumstances to trade detail for brevity in developing guidelines. We need to develop decision aids that support clinicians and patients to clarify the goals of care, raise and answer questions about the quality and completeness of evidence, and understand and contextualize estimates of benefit and harm. We also need to improve both the usefulness and ease of use of these and other evidence based tools (models, scores, algorithms, and so on) including the intellectual, social, and temporal demands they make on users and the resource implications for the healthcare organisation and system.

In the educational field, it is time we extended the evidence base for integrated curriculums that promote reflection and case discussion alongside the application of evidence.6 Discussions on how to interpret and apply evidence to real cases, and the sharing of collective knowledge and expertise in the form of “mindlines” among clinicians3 or within illness communities5 may provide useful data sources for such studies. It is by studying these more sophisticated forms of knowing that we are likely to determine how best to produce expert clinicians and expert patients, and to prevent the harms that arise from overdiagnosis, overtreatment, and overscreening.13

In relation to effectiveness, we need greater attention to postmarketing research in day to day hospital and primary care settings to confirm that subsequent experience replicates the results of licensing trials. This will allow gold standard tests and their cut-off points for ruling out diagnoses and treatments to be revised to minimise overdiagnosis or underdiagnosis.43

Finally, in relation to the collective effort to prevent the misappropriation of the evidence based quality mark, a key research priority remains the study of hidden biases in sponsored research—for example, by refining the statistical techniques for challenging findings that appear too good to be true.

Conclusion

Much progress has been made and lives have been saved through the systematic collation, synthesis, and application of high quality empirical evidence. However, evidence based medicine has not resolved the problems it set out to address (especially evidence biases and the hidden hand of vested interests), which have become subtler and harder to detect. Furthermore, contemporary healthcare’s complex economic, political, technological and commercial context has tended to steer the evidence based agenda towards populations, statistics, risk, and spurious certainty. Despite lip service to shared decision making, patients can be left confused and even tyrannised when their clinical management is inappropriately driven by algorithmic protocols, top-down directives and population targets.

Such problems have led some to argue for the rejection of evidence based medicine as a failed model. Instead we argue for a return to the movement’s founding principles—to individualise evidence and share decisions through meaningful conversations in the context of a humanistic and professional clinician-patient relationship (box 2). To deliver this agenda, evidence based medicine’s many stakeholders—patients, clinicians, educators, producers and publishers of evidence, policy makers, research funders, and researchers from a range of academic disciplines—must work together. Many of the ideas in this paper are not new, and a number of cross sector campaigns with similar goals have already begun (box 3). We hope that our call for a campaign for real evidence based medicine will open up debate and invite readers to contribute (for example, by posting rapid responses on bmj.com).

We thank Ruth Davis and Sarah Hardy for administrative support. The workshop was funded by a National Institute for Health Research senior investigator award for TG and by the Oxford Centre for Evidence Based Medicine. We also thank Bernard Higgins for advice on sections of the manuscript and Michael Rawlins and Bmj editors for helpful comments. Contributors and sources: Other members of the Evidence Based Medicine Renaissance Group are: Jon Brassey, Drum Burch, Martin Burton, Hasok Chang, Paul Glassiou, Iona Heath, Carl Heneghan, Michael P Kelly, Richard Lehman, Huw Llewelyn, Margaret McCartney, Ruairidh Milne, and Des Spence. This essay is partly the product of a workshop held at Kellogg College, Oxford, on 13-14 January 2014 and organised collaboratively by the Centre for Evidence Based Medicine, University of Oxford and the Centre for Primary Care and Public Health, Barts and the London School of Medicine and Dentistry. TG, in correspondence with CH, conceptualised the idea for a reappraisal and renaissance of evidence based medicine and developed the plan to hold a workshop to progress this idea. TG and JH facilitated the workshop. All authors attended the workshop and contributed to the development of key themes for the paper. TG wrote the first draft of the paper and refined it with important intellectual contribution from JH and NM. All authors provided some feedback on earlier drafts of the paper and approved the final manuscript.

Competing interests: All authors have read and understood BMJ policy on declaration of interests and declare the following interests: NM and MK have senior roles in the National Institute for Health and Care Excellence; RM is employed by, and TG and JH’s salaries are part funded by, the National Institute for Health Research; MB is director of the UK Cochrane Centre, JB is a director and shareholder in the TRIP database, which offers a rapid search service, RL is a founder member, and TG a member of the AI trials campaign. IH is chairing the scientific committee for the preventing overdiagnosis conference in September 2014, and PG and HL have also been active in the preventing overdiagnosis campaign. RL is currently a consultant to the Yale Open Data Access project, which receives funding from Johnson and Johnson to develop methods to promote clinical trial data sharing; the project has also received funding from Medtronic.

Provenance and peer review: Not commissioned; externally peer reviewed.


Box 3: Campaigns aligned with real evidence based medicine

Too much medicine—A rapidly growing movement, led jointly by clinicians, academics and patients, aims to reduce harm from overdiagnosis, overscreening, overtreatment. The campaign will hold an annual “preventing overdiagnosis” conference to be held in Oxford in September 2014 (www.preventingoverdiagnosis.net)

All trials (www.alltrials.net)—an international initiative to ensure that all clinical trials are registered at inception and no findings are withheld from publication.

Reducing waste and increasing value in medical research (www.thelancet.com/series/research)—A recent Lancet series highlighting the waste and loss of value caused by poor research design, is weighed down by bureaucracy, or is so badly or inappropriately reported that practitioners and policymakers simply cannot apply it.

Improving publishing standards (www.icmje.org/urm_main.html)—A campaign by the International Committee of Medical Journal Editors to improve the quality and transparency of medical publishing by discouraging ghost-writing and raising the standards for declarations of conflicts of interest.

Integrated medical education—Campaign to strengthen the integration of the different components of the curriculum by developing bedside clinical skills, understanding and applying research evidence, and reflecting and deliberating about complex cases.24


Cite this as: BMJ 2014;348:g3725
© BMJ Publishing Group Ltd 2014