What is the placebo worth?

The doctor-patient relationship is a crucial part of its value

RESEARCH, p 999

David Spiegel Willson professor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA 94305-5718, USA
dspiegel@stanford.edu
Anne Harrington professor and chair, Department of the History of Science, Science Center 371, Harvard University, Cambridge, MA 02138, USA

Competing interests: None declared.
Provenance and peer review: Commissioned; not peer reviewed.

BMJ 2008;336:967-8
doi: 10.1136/bmj.39535.344201.BE

George Bernard Shaw described a miracle as "an event that creates faith." Belief is a powerful tool, and many factors influence it. A recent study testing pain relief from analgesics showed that merely telling people that a novel form of codeine they were taking (actually a placebo) was worth $2.50 (£1.25; €1.58) rather than 10 cents increased the proportion of people who reported pain relief from 61% to 85.4%.[1] When the "price" of the placebo was reduced, so was the pain relief. A meta-analysis of decades of clinical trials proposed that the placebo effect was more hype than reality.[2] However, the resulting backlash against it has had the implicit effect of clarifying what is best practice with regard to the placebo.[3]

Hovering over much of the research is a practical question for clinicians—what does all this mean for patient care? In the accompanying randomised controlled trial, Kaptchuk and colleagues undertake a dismantling approach to the examination of placebo effects.[4] In 262 adults with irritable bowel syndrome, they examined the effects of placebo acupuncture in circumstances that involved observation only (evaluating a "Hawthorne effect"), sham acupuncture alone, and an enriched relationship with the treating doctor along with the sham procedure. The proportion of patients who reported moderate or substantial improvement on the irritable bowel syndrome global improvement scale was 3% in the observation group, 20% in the procedure alone group, and 37% in the augmented intervention group ($P<0.001$ for trend).

Clearly the group with the greatest relief of symptoms was the one that received not only sham acupuncture but 45 minutes of quality contact with a clinician. This contact involved questions about the patient's symptoms and beliefs about them, a "warm, friendly manner," empathy, and communication of confidence and positive expectations. In contrast, the doctor-patient relationship in the sham acupuncture only group sounds like a caricature of procedure based medicine practised under strict time limitations: the practitioners explained that this was "a scientific study" and they had been instructed not to talk about it with patients.

Global improvement scores were higher and quality of life and amelioration in symptom severity were almost doubled in people receiving augmented care, which raises some interesting questions. Perhaps the ratcheting down of the time that doctors spend with patients and our modern overemphasis on drugs and procedures is "penny wise and pound foolish." Patients might respond better to real as well as placebo interventions if they were associated with a good doctor-patient relationship. Although the increased time and concern may enhance the effects of the placebo, it also changes the context of associations with the treatment—the doctor may enhance the effect of the sham needle, but the needle also becomes a reminder of the enriched relationship.

That this study chose to evaluate placebo effects associated with an unconventional treatment raises further interesting questions. It is already widely assumed by sceptics that most if not all of the benefit of "alternative" or integrative medicine comes from the placebo effect. It is then assumed that demonstration of a powerful placebo effect, without proving a specific effect, is enough to consign the treatment to the realm of quackery.

But what if we asked a different question? Is it possible that the alternative medical community has tended historically to understand something important about the experience of illness and the ritual of doctor-patient interactions that the rest of medicine might do well to hear? Many people may be drawn to alternative practitioners because of the holistic concern for their wellbeing they are likely to experience, and many may also experience appreciable placebo responses. Why shouldn't we try to understand what alternative practitioners know and do, as this may help explain why so many patients are prepared to pay to be treated by them, even when many of the treatments are unproven?[5]

In seeking such understanding we should think about the conditions for which patients often seek alternative treatment, and what that might teach us. Patients with irritable bowel syndrome have a chronic condition that can deeply affect their quality of life. They usually have a story to tell about their suffering and want it to be heard, and an empathetic ear may be just what they need. Both the emotional and physical needs of a patient needing emergency surgery, however, might be very different. Such patients might well have a strong placebo response to a calm, orderly, high-tech hospital environment and a kind but focused doctor who does not stop long to chat but instead brings his or her full attention to the pressing business at hand.

Whatever the specifics, the take home message is clear. We treat patients in a social and psychophysiological context that can either improve or, alas, worsen outcome. The meanings and expectations created by the interactions of doctors and patients matter physically, not

just subjectively. Recent brain imaging research on pain and the placebo effect has shown functional connectivity between specific brain regions that process attention (the anterior cingulate gyrus) and pain (periaqueductal grey), involving endogenous opiate receptors.[6] Techniques such as hypnosis improve a range of objective symptoms of irritable bowel syndrome and produce subjective reductions in distress.[7 8] The word "placebo" is Latin for "I will please." On the basis of these and related studies, it seems fair to conclude that a good doctor-patient relationship can tangibly improve patients' responses to treatment, placebo or otherwise.

1 Waber RL, Shiv B, Carmon Z, Ariely D. Commercial features of placebo and therapeutic efficacy. *JAMA* 2008;299:1016-7.
2 Hrobjartsson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* 2001;344:1594-602.
3 Spiegel D, Kraemer H, Carlson R. Letter to the editor. Is the placebo powerless? *N Engl J Med* 2001;345:1276.
4 Kaptchuk TJ, Kelley JM, Conboy LA, Davis RB, Kerr CE, Jacobson EE, et al. Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ* 2008; doi: 10.1136/bmj.39524.439618.25.
5 Eisenberg DM, Davis RB, Ettner SL, Appel S, Wilkey S, Van Rompay M, et al. Trends in alternative medicine use in the United States, 1990-1997: results of a follow-up national survey. *JAMA* 1998;280:1569-75.
6 Wager TD, Scott DJ, Zubieta JK. Placebo effects on human mu-opioid activity during pain. *Proc Natl Acad Sci USA* 2007;104:11056-61.
7 Gonsalkorale WM, Miller V, Afzal A, Whorwell PJ. Long term benefits of hypnotherapy for irritable bowel syndrome. *Gut* 2003;52:1623-9.
8 Whorwell PJ. Hypnotherapy in irritable bowel syndrome. *Lancet* 1989;1:622.

Assessment of proteinuria in pregnancy

Urinary spot protein:creatinine ratio can reliably rule out proteinuria in pregnancy

IAN HOOTON/SCIENCE PHOTO LIBRARY

RESEARCH, p 1003

Lucy C Chappell senior lecturer in maternal and fetal medicine
lucy.chappell@kcl.ac.uk

Andrew H Shennan professor of obstetrics, Division of Reproduction and Endocrinology, King's College, London SE1 7EH

Competing interests: None declared.

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

BMJ 2008;336:968-9
doi: 10.1136/bmj.39540.657928.BE

Pre-eclampsia is a global problem—it affects 2-8% of pregnancies, and an estimated 8.3 million women develop the disease each year. For developing countries, the priority is preventing maternal deaths from multiorgan complications of the disease. The difference in case fatality rates from eclampsia between developing countries and developed countries (5.2% *v* 0.72%) suggests that mortality is easily avoidable.[1] In developed countries where death is rarer, research is directed towards improving prediction and prevention of pre-eclampsia and minimising morbidity. Accurate diagnosis is needed to accomplish this. In the accompanying systematic review, Côté and colleagues assess urinary spot protein:creatinine and albumin:creatinine ratios as diagnostic tests for significant proteinuria in women with hypertension in pregnancy.[2]

Pre-eclampsia is a multiorgan syndrome, the clinical characteristics of which may include kidney, liver, and cerebral damage, an altered coagulant state, and fetal growth restriction.[3] It is defined by two imperfect measures of end organ involvement—hypertension and proteinuria.[4] Early blood markers of the disease can now be identified many weeks before these clinical manifestations,[5] but antenatal diagnosis still relies on measuring blood pressure and urine dipstick inspection. In the United Kingdom alone, 660 000 women each year will have at least 7-10 such antenatal checks during each pregnancy, according to the schedule recommended by the National Institute for Health and Clinical Excellence.[6] In other developed countries, women are seen even more often.

The classic clinical presentation of pre-eclampsia is initial hypertension and subsequent proteinuria, which is a cue to seek other manifestations of the disease. The likelihood of adverse pregnancy outcomes escalates in women with pre-eclampsia rather than gestational hypertension alone, so proteinuria often dictates the need to admit women to hospital or deliver the baby.[7] Screening for proteinuria has traditionally been by urine dipstick, but this method is prone to considerable error.

A previous systematic review reported the pooled negative likelihood ratio for 1+ protein or greater on urine dipstick for predicting 300 mg/24 h proteinuria as 0.6 (95% confidence interval 0.45 to 0.8), with a pooled positive likelihood ratio of 3.48 (1.66 to 7.27).[8] This implies that the test often misleads the healthcare professional but, as no easy alternative is available at point of care, it continues to be used widely in clinical practice.

Comparison of urine dipstick with 24 hour collection is complicated by 24 hour collection not being a precise gold standard. Healthcare professionals and the women rarely follow the correct procedure at the start and end of collection, and because women micturate so often during pregnancy they may forget to add all the samples to the collection. Different laboratory assays for measuring protein give varying results,[9] and the threshold of 300 mg/24 h is not based on reaching a distinct clinicopathological state, but on the 95th centile for normal pregnant women. Proteinuria is key to management, but current clinical assessment is far from perfect. Can we improve on this situation?

In their systematic review, Côté and colleagues report a pooled negative likelihood ratio for a cut-off of 30 mg/mmol of 0.21 (0.13 to 0.31).[2] They conclude that the test is useful for ruling out significant proteinuria but do not advocate its use for quantification as the positive likelihood ratio was only poor to fair (3.53, 2.83 to 4.49). However, assessment of the value of the test must include comparison with its alternatives and the implications of both a positive and missed diagnosis, which are considerable.

Quantification is not crucial to obstetricians, because once a threshold is breached the diagnosis is made. Unlike blood pressure (which is related to risk of a cerebrovascular event), serial assessment of proteinuria after the diagnosis of pre-eclampsia rarely influences management because the rate of increase is not an important

predictor of adverse pregnancy outcomes.[10] Ruling out proteinuria is key, and the protein:creatinine ratio performs well in this respect. Detection of proteinuria above the threshold in a pregnant woman with hypertension differentiates between relatively innocuous gestational hypertension and pre-eclampsia and dictates a considerable step-up in surveillance, often including admission. The social and financial repercussions of this for the woman and the economic consequences for the healthcare system are considerable.

The protein:creatinine ratio is most likely to improve the initial screening of proteinuria by dipstick assessment. It may have a more limited role in precluding the need for 24 hour collection. When the protein:creatinine ratio is above 30 mg/mmol, the 24 hour collection will still be used in many units to quantify proteinuria and confirm the diagnosis, but when the ratio is below the threshold it could replace 24 hour collection in excluding significant proteinuria. Although it is currently a laboratory test, it is quick; obstetric day care units could use it to improve accuracy and speed of diagnosis. Fully quantitative point of care devices exist for measuring albumin:creatinine ratio[11]; a similar one for protein:creatinine ratio may be an easy and more accurate alternative to current urinalysis screening in the future.

Côté and colleagues rightly highlight the need for future research into the validity of the 30 mg/mmol threshold for the protein:creatinine ratio in predicting adverse pregnancy outcome.[2] Further work is also needed to compare the use of urinary spot protein:creatinine ratio with current urine dipstick analysis in the management of pregnant women with hypertension, particularly with respect to accuracy of diagnosis and the effect on inappropriate admissions and discharges. The 24 hour collection system will still be needed for those women who require accurate quantification of proteinuria. New technology in the assessment of proteinuria that incorporates the protein:creatinine ratio has the potential to improve the lives of many pregnant women, while reducing wrong diagnoses and therefore enhancing safety.

1 Villar J, Say L, Gulmezoglu AM, Merialdi M, Lindheimer MD, Betran AP, et al. Eclampsia and preeclampsia: a worldwide health problem for 2000 years. In: Critchley H, MacLean A, Poston L, Walker J, eds. *Preeclampsia.* London: RCOG Press, 2003.
2 Côté AM, Brown MA, Lam E, von Dadelszen P, Firoz T, Liston RM, et al. Diagnostic accuracy of urinary spot protein:creatinine ratio for proteinuria in hypertensive pregnant women: systematic review. *BMJ* 2008; doi: 10.1136/bmj.39532.543947.BE.
3 Roberts JM, Gammill HS. Preeclampsia: recent insights. *Hypertension* 2005;46:1243-9.
4 Brown MA, Lindheimer MD, de Swiet M, Van Assche A, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens Pregnancy* 2001;20:IX-XIV.
5 Tjoa ML, Oudejans CB, van Vugt JM, Blankenstein MA, van Wijk IJ. Markers for presymptomatic prediction of preeclampsia and intrauterine growth restriction. *Hypertens Pregnancy* 2004;23:171-89.
6 National Institute for Health and Clinical Excellence. *Antenatal care: routine care for the healthy pregnant woman.* 2008. www.nice.org.uk/CG62.
7 Ferrazzani S, Caruso A, De Carolis S, Martino IV, Mancuso S. Proteinuria and outcome of 444 pregnancies complicated by hypertension. *Am J Obstet Gynecol* 1990;162:366-71.
8 Waugh JJ, Clark TJ, Divakaran TG, Khan KS, Kilby MD. Accuracy of urinalysis dipstick techniques in predicting significant proteinuria in pregnancy. *Obstet Gynecol* 2004;103:769-77.
9 Waugh J, Bell SC, Kilby MD, Lambert P, Shennan A, Halligan A. Urine protein estimation in hypertensive pregnancy: which thresholds and laboratory assay best predict clinical outcome? *Hypertens Pregnancy* 2005;24:291-302.
10 Schiff E, Friedman SA, Kao L, Sibai BM. The importance of urinary protein excretion during conservative management of severe preeclampsia. *Am J Obstet Gynecol* 1996;175:1313-6.
11 Waugh JJ, Bell SC, Kilby MD, Blackwell CN, Seed P, Shennan AH, et al. Optimal bedside urinalysis for the detection of proteinuria in hypertensive pregnancy: a study of diagnostic accuracy. *Br J Obstet Gynecol* 2005;112:412-7.

Corticosteroids for acute respiratory distress syndrome

Avoid corticosteroids for prophylaxis; possibly use them for treatment

RESEARCH, p 1006

Neill K J Adhikari lecturer
neill.adhikari@utoronto.ca

Damon C Scales assistant professor, Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, ON, Canada M4N 3M5

Competing interests: None declared.

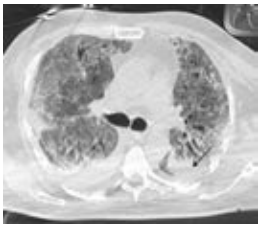
Provenance and peer review: Commissioned; not externally peer reviewed.

BMJ 2008;336:969-70
doi: 10.1136/bmj.39553.408924.80

Acute respiratory distress syndrome (ARDS) is a major health problem, with the annual incidence in the United States approaching 200 000 cases.[1] In the accompanying paper, Peter and colleagues present a systematic review and meta-analysis of nine randomised controlled trials (1073 patients) of corticosteroids for prevention and treatment of ARDS.[2] ARDS is a form of severe respiratory failure resulting from direct pulmonary insults (for example, aspiration or pneumonia) or indirect systemic causes (for example, sepsis or trauma).[3] This syndrome often has devastating consequences, such as the prolonged need for mechanical ventilation, a high probability of death, and long term physical and psychological sequelae in survivors. Treatment is unlikely to be successful in least developed countries because of limited critical care resources.[4]

No effective drug treatments are available for ARDS,[5] but corticosteroids have attracted attention because they have anti-inflammatory properties that are relevant to ARDS pathology. They reduce both leakage of fluid through the alveolar-capillary membrane and the adhesion of neutrophils to the capillary endothelium, and they modulate the balance between proinflammatory and anti-inflammatory genes.[6] However, these physiological benefits are tempered by concerns about side effects, including infections and neuromuscular weakness.

Peter and colleagues analysed outcomes using Bayesian methods, which may be unfamiliar to some readers. Standard frequentist analysis of observed data generates a 95% confidence interval. This implies that if the study were repeated many times, 95% of the confidence intervals generated should contain the true (but unknown) population effect. Unlike these standard confidence intervals, which despite their name do not tell readers how "confident" to be in the observed results, Bayesian methods produce a 95% "credible interval," which has a 95% probability of containing the population treatment



The most compelling reason for additional research may simply be that clinicians' beliefs after the systematic review still depend largely on their prior beliefs

effect.[7] These methods also estimate the probability that the population treatment effect exceeds a specified threshold or falls within a specified interval.

The Bayesian approach thus helps readers quantify the statistical uncertainty of the results, but at the cost of requiring a prespecified range of prior beliefs about the treatment effect.[8] Clinicians' prior beliefs—ranging from "enthusiastic" to "non-informative" (no prior belief) to "sceptical"—stem from clinical experience and interpretation of existing studies. In practice, Bayesian analysis of a clinical trial using non-informative priors gives similar results to standard statistical methods.[7] However, the advantage of Bayesian results is that their interpretation is more intuitive than the confidence interval.[9]

The systematic review by Peter and colleagues included trials of moderate quality; steroids had opposite effects depending on the clinical scenario. When prescribed to patients at risk before the onset of ARDS in four trials, steroids increased both the odds of developing ARDS (odds ratio 1.55, 95% credible interval 0.58 to 4.05) and subsequently dying from ARDS (1.52, 0.30 to 5.94). Both credible intervals spanned 1 and thus could not exclude a null effect.

Bayesian analysis then answers a question not dealt with by frequentist statistics—what is the probability of harm given the available data? Neutral or sceptical clinicians would conclude that the probability of prophylactic corticosteroids causing ARDS (86.6%) and increasing mortality from ARDS (72.8%) is high. Such clinicians will probably decide that steroids are harmful and that they should be avoided, while even steroid enthusiasts may become less convinced.

Conversely, when used for established ARDS, corticosteroids seemed to decrease mortality (five trials; 0.62, 0.23 to 1.26) and increase days both alive and off mechanical ventilation (three trials). Reassuringly, they did not lead to new infections, except possibly at high doses. Although a null effect on mortality could not be excluded, the probability of a positive effect was high (93.2%).

Many neutral doctors will conclude that corticosteroids are probably helpful for established ARDS, although it would also be useful to know the probability that the reduction in odds ratio exceeds a clinically important threshold. Sceptics will correctly observe that while Bayesian meta-analysis expresses statistical uncertainty more clearly it does not deal with concerns related to bias or other limitations of the primary trials. Indeed, trials were few and clinically heterogeneous, with limited data on other side effects.

Finally, although steroid enthusiasts may be reassured, optimal delivery of this intervention remains unclear. Peter and colleagues found no strong evidence for associations between treatment effect and variables at the study level of total corticosteroid dose or duration of ARDS before treatment, although in the largest trial corticosteroids seemed to increase mortality when prescribed more than 14 days after the development of ARDS.[10] In addition, trials differed in precautions to reduce steroid related complications, including actively checking for infection, limiting use of muscle relaxants, slow tapering of the steroid dose, and avoiding hyperglycaemia.

Consequently, for most doctors without strong prior beliefs, this systematic review provides moderately strong evidence for avoiding prophylactic corticosteroids in ARDS, and weak evidence for their therapeutic use. Such doctors will view these results as hypothesis generating and will await results of additional trials. In contrast, previously enthusiastic doctors will probably remain so, at least for the treatment of ARDS. In practice, clinicians have largely abandoned steroids for prevention and early treatment, but many still prescribe them for non-resolving ARDS.[11]

Given current data, should additional randomised trials of corticosteroids to treat ARDS be conducted? We think so, because relatively few patients have been studied; the corticosteroid regimen (dosing, timing of initiation and discontinuation, monitoring) is unclear; and these drugs are inexpensive and therefore useful in low resource settings if effective. Such trials should be informed by further pilot work to refine the study question. Sample size calculations should consider plausible treatment effects in ARDS (for example, ~9% absolute reduction in mortality with tidal volume limitation)[12] and the minimum difference worth detecting that would be important to clinicians and patients[13]; Bayesian methods may also be used. In addition, clinical predictors of responsiveness to corticosteroids in patients with ARDS would also be desirable, given the side effects. Stated otherwise, the most compelling reason for additional research may simply be that clinicians' beliefs after the systematic review still depend largely on their prior beliefs.

1 Rubenfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, et al. Incidence and outcomes of acute lung injury. *N Engl J Med* 2005;353:1685-93.
2 Peter JV, John P, Graham PL, Moran JL, George IA, Bersten A. Corticosteroids in the prevention and treatment of acute respiratory distress syndrome (ARDS) in adults: meta-analysis. *BMJ* 2008; doi: 10.1136/bmj.39537.939039.BE.
3 Leaver SK, Evans TW. Acute respiratory distress syndrome. *BMJ* 2007;335:389-94.
4 Dunser MW, Baelani I, Ganbold L. A review and analysis of intensive care medicine in the least developed countries. *Crit Care Med* 2006;34:1234-42.
5 Adhikari N, Burns KE, Meade MO. Pharmacologic therapies for adults with acute lung injury and acute respiratory distress syndrome. *Cochrane Database Syst Rev* 2004;(4):CD004477.
6 Rhen T, Cidlowski JA. Antiinflammatory action of glucocorticoids—new mechanisms for old drugs. *N Engl J Med* 2005;353:1711-23.
7 Sterne JA, Davey SG. Sifting the evidence—what's wrong with significance tests? *BMJ* 2001;322:226-31.
8 Diamond GA, Kaul S. Prior convictions: Bayesian approaches to the analysis and interpretation of clinical megatrials. *J Am Coll Cardiol* 2004;43:1929-39.
9 Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. *Ann Intern Med* 1999;130:995-1004.
10 Steinberg KP, Hudson LD, Goodman RB, Hough CL, Lanken PN, Hyzy RC, et al; National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome. *N Engl J Med* 2006;354:1671-84.
11 Meade MO, Jacka MJ, Cook DJ, Dodek P, Griffith L, Guyatt GH, et al. Survey of interventions for the prevention and treatment of acute respiratory distress syndrome. *Crit Care Med* 2004;32:946-54.
12 Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. *N Engl J Med* 2000;342:1301-8.
13 Scales DC, Rubenfeld GD. Estimating sample size in critical care clinical trials. *J Crit Care* 2005;20:6-11.

Problem based learning

Time to stop arguing about the process and examine the outcomes

Diana F Wood director of medical education and clinical dean, University of Cambridge School of Clinical Medicine, Addenbrooke's Hospital, Cambridge CB2 2SP
dfw23@medschl.cam.ac.uk

Competing interests: None declared.

Provenance and peer review: Commissioned; not externally peer reviewed.

BMJ 2008;336:971
doi: 10.1136/bmj.39546.716053.80

Problem based learning was developed in the late 1960s and has been the most influential innovation in medical education during the past 40 years. Essentially, problem based learning is a small group teaching method that combines the acquisition of knowledge with the development of generic skills. Educationally, it is theoretically grounded in adult learning theory and constructivism and is predicted to produce a better learning environment and improved graduate knowledge, skills, and attitudes.

Implementation of problem based learning requires fundamental changes in the way educators conceive, design, deliver, and assess the curriculum.[1] Despite the cost and resource implications, problem based learning has been introduced to varying degrees throughout the world—for example, it is used in most medical schools in the United States and many new medical schools in developing countries. Given this wide scale adoption, why is problem based learning still a controversial topic? The answer lies in the lack of convincing evidence for its superiority over other teaching methods in terms of graduate outcomes. This is not for want of trying—the medical education literature abounds with publications on problem based learning, which have produced lively debate.[2-6] But the question remains—does problem based learning produce better doctors?

A recent systematic review assessed how problem based learning during medical school affected the competence of doctors after graduation.[7] Only publications that included a control group of graduates from a "traditional" curriculum were included. The study population ranged from first year graduates to doctors who had been in practice for up to 20 years. Most of the studies were surveys, and an important feature in the final analysis was that doctors' self assessments of the competencies in question and assessments by independent observers were considered separately. The level of evidence in favour of problem based learning over traditional learning was derived from previously published data coupled with the research team's scoring system, which increased weighting for randomisation, sample size, objective assessment, and response rate. Thirteen studies were included and 38 competencies were categorised into eight dimensions—overall, technical, social, cognitive, managerial, research, teaching, and knowledge. Of these, the social dimension showed the strongest evidence in favour of problem based learning. In line with previous data,[8] little correlation was seen between self assessed and observer assessed competency. When both self reported and independently observed assessments were combined, four competencies had moderate to strong evidence in favour of problem based learning—coping with uncertainty (strong), appreciation of legal and ethical aspects of health care (strong), communication skills (moderate (self assessed), strong (observed)), and self directed learning (moderate). Self assessment showed a strong level of evidence against problem based learning for possession of medical knowledge, but this was not confirmed by independent observation. The authors conclude that problem based learning has positive effects on graduate competencies in important social and cognitive domains.

This review confirms what most educators have come to believe on the basis of hundreds of less rigorous reports—that, compared with traditional learning, problem based learning has beneficial effects on some psychosocial outcomes of undergraduate medical education. However, one important factor not acknowledged here or elsewhere in the literature is the lack of definition of the "control traditional curriculum."

The student cohorts in Koh and colleagues review date from the 1980s and 1990s, when traditional control curriculums were probably based on a rigid divide between preclinical and clinical education, entirely lecture based programmes, and didactic clinical teaching. Since then, outcome based frameworks for medical education have focused on the competencies expected of graduates to meet the demands of patients in modern society. Crucially, the emphasis in medical education has moved from the process to the product.[9 10] Features previously associated with problem based learning such as fewer lectures and smaller groups are now found in most undergraduate curriculums. Teaching and learning in communication skills and the psychosocial domains can be achieved in many ways, and working in small groups—coupled with timely and constructive feedback—may be just as effective as problem based learning.

Performing outcomes based research in education is difficult because of the large range of confounding factors. What is clear, however, is that graduates from different medical schools perform very differently in postgraduate examinations, and some of this variance is due to the undergraduate teaching programme.[3 11] Surely it is time to stop arguing about the process and ensure that diversity in undergraduate educational provision is related to declared graduate outcomes and delivers doctors who have the required competencies for good medical practice.

1 Wood DF. Problem based learning. *BMJ* 2003;326:328-30.
2 Albanese M. Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Med Educ* 2000;34:729-38.
3 Colliver JA. Effectiveness of problem-based learning curricula: research and theory. *Acad Med* 2000;75:259-66.
4 Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Med Educ* 2000;34:721-8.
5 Dolmans D. The effectiveness of PBL: the debate continues. Some concerns about the BEME movement. *Med Educ* 2003;37:1129-30.
6 Farrow R, Norman GR. The effectiveness of PBL: the debate continues. Is meta-analysis helpful? *Med Educ* 2003;37:1131-2.
7 Koh GC-H, Khoo HE, Wong ML, Koh D. The effects of problem-based learning during medical school on physician competency: a systematic review. *CMAJ* 2008;178:34-41.
8 Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence. A systematic review. *JAMA* 2006;296:1094-102.
9 Harden RM. Outcome-based education: the future is today. *Med Teach* 2007;29:625-9.
10 General Medical Council. *Tomorrow's doctors*. 2002. www.gmc-uk.org/education/undergraduate/GMC_tomorrows_doctors.pdf.
11 McManus IC, Elder AT, de Champlain A, Dacre JE, Mollon J, Chis L. Graduates of different UK medical schools show substantial differences in performance on MRCP(UK) part 1, part 2 and PACES examinations. *BMC Med* 2008;6:5.

Opium production in Afghanistan

Legalising production for medical usage is neither feasible nor desirable

AHMAD MASOOD/REUTERS

Mark Malloch-Brown minister of state for Africa, Asia, and the UN, Foreign and Commonwealth Office, London SW1A 2AH
Stelios.Kyriakides@fco.gov.uk

Competing interests: None declared.
Provenance and peer review: Commissioned; not externally peer reviewed.

BMJ 2008;336:972
doi: 10.1136/bmj.39554.402199.BE

The case for allowing Afghanistan to cultivate poppies to combat the alleged shortage of opiate medicines has been reported in the *BMJ*.[1] Around 90% of the world's opium comes from Afghanistan, and most of it is destined for the illicit market. Counter narcotics is at the heart of the United Kingdom's involvement in Afghanistan, not just because it is a major source of finance for the Taliban, but also because of its impact on our streets and neighbourhoods. The UK has repeatedly considered legal poppy cultivation—the first time in a report by David Mansfield that was commissioned in 2001.[2] The findings of the Mansfield report remain true today; unless the Afghan government can control the size of the crop, legal opium will supplement and not substitute illegal opium production.

The Afghan government lacks the necessary resources, institutional capacity, and control mechanisms to guarantee that opium is only purchased legally. Those cultivating and purchasing opium for medical usage would be in direct competition with illegal traffickers, which could drive up the price of opium and encourage increased cultivation. Farmers who do not currently grow poppies would abandon legal crops to meet the market's demand. Ultimately, the area of land under poppy cultivation could increase. Quite simply, farmers would grow more to supply an additional purchaser.

Demand for legal opium is therefore better met by established sources of production from countries like Turkey and Australia. These countries have far greater security, stability, and government capacity. With opium production up to six times cheaper in these countries than in Afghanistan, supply is more economically viable too.

The UK government's approach in Afghanistan is based on the recognition that poverty and instability continue to drive farmers to produce opium. The insurgency in the south prevents access to agricultural commodity and labour markets, so crops other than opium are not practical to produce.[3] With a swift turnover and high returns, opium gets farmers the best credit, access, and influence. But it is precisely because of these benefits that opium production continues to drive instability, and therefore insurgency. Opium production fuels corruption and undermines the rule of law. And only by breaking the vicious cycle, by reducing the flow of illegal opium, will the goals of establishing security and good governance in Afghanistan be met. As the G8's partner nation leading on counter narcotics, the UK provided more than £290m (€360m; $570m) between 2005 and 2008 to support the Afghan government's national drug control strategy to try and break this cycle.

Half of the UK's approach uses the stick. Prohibition actions are being intensified against drug traders, opium processing plants, and people in government who are associated with the opium industry. Institutions of law enforcement are being developed. The tireless efforts of our troops on the ground are fighting the insurgency that is so intertwined with opium production.

The other half is supplying the carrots. The challenge is to create development initiatives and economic incentives that provide attractive alternatives for farmers—ones that allow them to pursue sustainable and legal livelihoods. Infrastructure and local government capacity must be further developed so that farmers' access to markets, land, water, credit, food security, and employment can be improved. Where these conditions are in place, some real successes have been reported—poppy-free provinces doubled from six to 13 in 2007. In 2008, the Afghan government aims to increase this to half of all provinces.

The arguments regarding legal opium production stretch beyond Afghanistan's borders. The question of demand is key. The International Narcotics Control Board, the independent body that monitors the implementation of international drug control conventions, recently reported that global demand for opiates for medical purposes is fully satisfied. Over the past five years, total supply (production and stocks) of opiate raw materials has exceeded the total amount of opiates needed for medical and scientific purposes.[4] However, the opposite may be true—that because of poor medical infrastructure and awareness in many countries, the global market for opiates for medical usage is underdeveloped.

Ultimately, however, the solution to the global problem of illicit opium production lies not just in Afghanistan but in every country that suffers the blight of heroin usage, including the UK. The Home Office drugs strategy 2008-18 aims to restrict the supply of illegal drugs to the UK and reduce demand. It will do so through tackling drug supply and drug related crime, preventing harm to children and families affected by drug misuse, delivering new approaches to drug treatment and social reintegration, and delivering targeted public information campaigns. But the UK is responsible for around only 4% of global demand for heroin. Only by reducing demand on the streets everywhere will the producers and traffickers on the streets of Afghanistan be given the best reason to follow their alternative livelihoods.

1 Dyer O. Afghan farmers should be licensed to grow poppies for morphine, Senlis Council says. *BMJ* 2007;334:1343
2 Mansfield D. *The displacement of opium poppy cultivation: a shift in the regional threat?* 2001. www.davidmansfield.org/data/Policy_Advice/UK/Disp_Report.doc.
3 Ward C, Mansfield D, Oldham P, Byrd W. *Afghanistan: economic incentives and development initiatives to reduce opium production*. Department for International Development/World Bank, 2008. www.dfid.gov.uk/pubs/files/Afghan-Opium-Incentives.pdf.
4 International Narcotics Control Board. Press release. *Current supply of legal opium adequate to meet world demand, says INCB president*. 2007. www.incb.org/incb/en/press_release_2007-11-12_01.html.