Editor's choice

Editor's Choice: Subtle signs
Rajendra Kale
BMJ 2007;335, doi:10.1136/bmj.39371.537755.47

Editorials

Boxing and the risk of chronic brain injury
Paul McCrory
BMJ 2007;335:781-782, doi:10.1136/bmj.39352.454792.80 (published 4 October 2007)

Breast feeding and the risk of allergy and asthma
Sheila Gahagan
BMJ 2007;335:782-783, doi:10.1136/bmj.39349.658993.80

Reporting of observational studies
Peter M Rothwell, Meena Bhatia
BMJ 2007;335:783-784, doi:10.1136/bmj.39351.581366.BE

Ototoxicity caused by aminoglycosides
Maria Bitner-Glindzicz, Shamima Rahman
BMJ 2007;335:784-785, doi:10.1136/bmj.39301.680266.AE

Management of chronic knee pain
Rob Herbert, Marlene Fransen

Letters

This week's letters
Detecting LVH: QRS voltage criteria can be useful
Peter J Bourdillon
BMJ 2007;335:787, doi:10.1136/bmj.39367.394155.BE
Detecting LVH: Ethnicity is relevant
Andrew P Vanezis, Raj Bhopal
BMJ 2007;335:787, doi:10.1136/bmj.39367.371736.BE

GPs' 24 hour responsibility: Summary of responses
Birte Twisselmann
BMJ 2007;335:787-788, doi:10.1136/bmj.39367.698125.BE

US liability: Damned if you treat, damned if you don't
Sean T Hudson
BMJ 2007;335:788, doi:10.1136/bmj.39367.432755.BE

Bismark v Beveridge: Unfair comparison
Iain S Fraser
BMJ 2007;335:788, doi:10.1136/bmj.39367.453519.BE

Vive la difference?: Training and careers in France
Herv Maisonneuve, Yves Matillon
BMJ 2007;335:788, doi:10.1136/bmj.39367.484213.BE

News

Experts clash as committee debates reducing 24 week abortion limit
Clare Dyer
BMJ 2007;335:789, doi:10.1136/bmj.39371.423924.DB

Obesity report signals start of government action
Lisa Hitchen

In Brief: News
BMJ 2007;335:790, doi:10.1136/bmj.39370.690845.13

Doctors need a "sea change" in their attitude to C difficile
Adrian O'Dowd
BMJ 2007;335:790, doi:10.1136/bmj.39371.650266.DB

West Yorkshire has most casualties from road traffic incidents
Roger Dobson
BMJ 2007;335:790, doi:10.1136/bmj.39371.471817.13

NICE guidelines create ethical dilemmas in care of elderly people, says report
Anne Gulland
BMJ 2007;335:791, doi:10.1136/bmj.39371.570509.DB

COX 2 inhibitor rejected in North America but retained in Europe
Bob Burton
Surgeons get modern training facility
Zosia Kmietowicz
BMJ 2007;335:792-793, doi:10.1136/bmj.39371.435718.13

74 year old GP refuses to take part in deputising service
Tony Sheldon
BMJ 2007;335:792, doi:10.1136/bmj.39367.610058.DB

Cancer surgeon is suspended over "unconventional donations"
Annette Tuffs
BMJ 2007;335:793, doi:10.1136/bmj.39367.571782.DB

NHS gets 4% more funding a year until 2011—but spend it wisely, says BMA
Helen Mooney
BMJ 2007;335:793, doi:10.1136/bmj.39367.545220.DB

Financial ties common between US medical schools and drug companies
Janice Hopkins Tanne
BMJ 2007;335:793, doi:10.1136/bmj.39370.485880.13

German hospitals demand extra cash to alleviate financial problems
Annette Tuffs
BMJ 2007;335:794, doi:10.1136/bmj.39367.502269.DB

Government should restrict advertising of baby milk products, say campaigners
Peter Moszynski
BMJ 2007;335:794, doi:10.1136/bmj.39367.616782.DB

Backpedalling from Blair’s privatisation agenda
Nicholas Timmins
BMJ 2007;335:794-795, doi:10.1136/bmj.39370.404294.DB

Royal colleges call for more staff on maternity wards
Andrew Cole
BMJ 2007;335:795, doi:10.1136/bmj.39367.603947.DB

NHS launches scheme to encourage diversity in top management posts
Jacqui Wise
BMJ 2007;335:795, doi:10.1136/bmj.39371.434803.DB

Scientists challenge companies' dubious marketing claims
Adrian O'Dowd
BMJ 2007;335:795, doi:10.1136/bmj.39367.694294.4E

Oral antibiotics work as well as intravenous treatment in severe urinary tract infection
Susan Mayor
Health of indigenous people suffers modern world ills
Peter Moszynski
BMJ 2007;335:795, doi:10.1136/bmj.39367.494757.DB

Wider use of colonoscopy could improve outcomes in bowel cancer
Zosia Kmietowicz

Shortcuts from other journals: Stress at work is bad for your heart
BMJ 2007;335:796, doi:10.1136/bmj.39367.567662.BE

Shortcuts from other journals: Children in the US receive inadequate outpatient care
BMJ 2007;335:796, doi:10.1136/bmj.335.7624.796-a

Shortcuts from other journals: Childhood obesity is a chronic disease
BMJ 2007;335:796, doi:10.1136/bmj.335.7624.796-b

Shortcuts from other journals: Bacterial colonisation of the airway linked to wheeze in young children
BMJ 2007;335:796, doi:10.1136/bmj.335.7624.796-c

Shortcuts from other journals: Topiramate helps alcohol dependent adults drink less
BMJ 2007;335:797, doi:10.1136/bmj.335.7624.797

Shortcuts from other journals: A fifth of pregnancies end in abortion
BMJ 2007;335:797, doi:10.1136/bmj.335.7624.797-a

Shortcuts from other journals: Maternal mortality is falling too slowly to meet international targets
BMJ 2007;335:797, doi:10.1136/bmj.335.7624.797-b

Shortcuts from other journals: Pravastatin has enduring effects on risk of heart disease in men
BMJ 2007;335:797, doi:10.1136/bmj.335.7624.797-c

Feature
Health in developing countries: Who cares for baby?
Hannah Brown
BMJ 2007;335:798-800, doi:10.1136/bmj.39365.579016.94

How can we regulate medicines better?
Silvio Garattini, Vittorio Bertele'
BMJ 2007;335:803-805, doi:10.1136/bmj.39281.615706.94

Observations
Media watch: Beware of mentioning psychosocial factors
Ben Goldacre
Response: Fiona Patterson replies to Parashkev Nachev
Fiona Patterson
BMJ 2007;335:802, doi:10.1136/bmj.39360.727535.59

Analysis
Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies
Erik von Elm, Douglas G Altman, Matthias Egger, Stuart J Pocock, Peter C Gøtzsche, Jan P Vandenbroucke, STROBE Initiative
BMJ 2007;335:806-808, doi:10.1136/bmj.39335.541782.AD

Research
Amateur boxing and risk of chronic traumatic brain injury: systematic review of observational studies
Mike Loosemore, Charles H Knowles, Greg P Whyte
BMJ 2007;335:809, doi:10.1136/bmj.39342.690220.55 (published 4 October 2007)

Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials
Catherine J Minns Lowe, Karen L Barker, Michael Dewey, Catherine M Sackley

Effect of prolonged and exclusive breast feeding on risk of allergy and asthma: cluster randomised trial
Michael S Kramer, Lidia Matush, Irina Vanilovich, Robert Platt, Natalia Bogdanovich, Zinaida Sevkovskaya, Irina Dzikovich, Gyorgy Shishko, Bruce Mazer, the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group

Clinical review
Uterine prolapse
Anjum Doshani, Roderick E C Teo, Christopher J Mayne, Douglas G Tincello
BMJ 2007;335:819-823, doi:10.1136/bmj.39356.604074.BE

Practice
A patient's journey: Acromegaly
Jon Danzig
BMJ 2007;335:824-825, doi:10.1136/bmj.39253.602141.AD

Views & reviews
Personal views: How to become an MP
Steven Duncan Ford
BMJ 2007;335:826, doi:10.1136/bmj.39353.702535.59
Review of the week: Reader, I didn't blush
Colin Martin
BMJ 2007;335:827, doi:10.1136/bmj.39370.480856.59

From the frontline: MTAS in GP land
Des Spence
BMJ 2007;335:828, doi:10.1136/bmj.39370.454884.59

Past Caring: A man who called hospitals to account
Wendy Moore
BMJ 2007;335:828, doi:10.1136/bmj.39366.616563.59

Between the lines: Miscellaneous truths
Theodore Dalrymple
BMJ 2007;335:829, doi:10.1136/bmj.39351.627199.59

Medical Classic: The Cunning Man
Alex Paton
BMJ 2007;335:829, doi:10.1136/bmj.39366.673333.59

Obituaries
This week's obituaries

John Michael Newsom-Davis
Caroline Richmond
BMJ 2007;335:830, doi:10.1136/bmj.39358.519248.BE

Stanley Norman Cole
Malcolm Read
BMJ 2007;335:831, doi:10.1136/bmj.39350.808183.BE

Nigel Henry Harris
Andrew Harris
BMJ 2007;335:831, doi:10.1136/bmj.39357.734282.BE

Michael Stuart Brighton Hollinrake
Aileen (Hollinrake) Jennings, Patricia Hollinrake, Sarah Hollinrake
BMJ 2007;335:831, doi:10.1136/bmj.39358.564664.BE

Colin Proteroe
H A McClelland, T A Kerr
BMJ 2007;335:831, doi:10.1136/bmj.39358.574850.BE

Geoffrey Sheers
Roger Sheers
BMJ 2007;335:831, doi:10.1136/bmj.39351.707106.BE

Julian Edmund Christopher Tower
Monica Baynes, June Tower
BMJ 2007;335:831, doi:10.1136/bmj.39358.595856.BE
John Ernest Michael Whitehead
Stephen Whitehead
BMJ 2007;335:831, doi:10.1136/bmj.39357.670023.BE

Minerva
Minerva
BMJ 2007;335:832, doi:10.1136/bmj.39367.622060.BD1

Minerva
S K Chan, P M Todd, N P Burrows
BMJ 2007;335:832, doi:10.1136/bmj.39367.622060.BD

Fillers
Clint's favourite question
Hannah Kuper
BMJ 2007;335:818, doi:10.1136/bmj.39290.613206.DE

Body heat
Bikki Gautam
BMJ 2007;335:823, doi:10.1136/bmj.39290.699074.DE

Corrections
Participation in mammography screening
BMJ 2007;335, doi:10.1136/bmj.39371.633032.AD
Boxing and the risk of chronic brain injury

Evidence is inconclusive but the absolute risk in modern day boxing is still low . . .

In this week's *BMJ*, a systematic review of observational studies by Loosemore and colleagues assesses the risk of chronic traumatic brain injury with amateur boxing. It finds that the quality of evidence is too poor to come to any definite conclusions. So, do we need to worry about the health of modern boxers, amateur or professional?

Concern over injury to fighters has been a persistent theme throughout the history of boxing. Although boxing was popular in early Rome, the practice was banned by Caesar Augustus, supposedly because of the high rates of injury in Roman legionnaires. The sport resurfaced in England during the 17th century in the form of bare knuckle boxing or prize fighting. The most famous of the rules introduced to protect the injured or incapacitated boxer were the 1867 Queensberry rules, which dictated that fights should be "a fair stand-up boxing match." Each fighter was given a 10 second count if he was knocked down and the length of bouts was time limited. Gloves of a "fair size" were introduced, which changed the nature of the sport, as bouts became longer and more strategic, with greater importance attached to defensive manoeuvres such as slipping, bobbing, countering, and angling.

As the changing nature of the sport and the use of protective equipment reduced acute injuries, concern began to develop about the chronic neurological risks of boxing. This was fuelled by a study published in 1928, which introduced the lay term “punch drunk” into medical terminology; this term has since become synonymous with impaired boxers. Surprisingly, the only clinical case examined in that study concerned idiopathic Parkinson’s disease.

Chronic traumatic brain injury has since been described in more detail. In the early stages of the condition, symptoms reflect lesions affecting the pyramidal, cerebellar, and extrapyramidal systems. In the later stages, cognitive and behavioural impairment predominate. About one third of cases are progressive. The pathological features of the condition are similar to Alzheimer’s disease, although some specific differences exist.

The crucial risk factor for chronic traumatic brain injury is exposure to head impact. The largest and best of the neuropathological studies included 15 ex-boxers, 12 of whom were professionals. These boxers had fought in the period 1900-40, and eight of them were national champions or world champions in their weight division. Although the study had methodological flaws—for example, all demographic and boxing exposure data were collected retrospectively—the most striking feature was the fighters’ high exposure to boxing. The number of career fights ranged from 400 to 700. Many boxers also worked in fairground boxing booths and had up to 30 or 40 fights each day over several years. The pathological features described have become the essential diagnostic criteria for chronic traumatic brain injury.

These injuries are unlikely to be seen in boxers today because of their relatively short careers. More recent studies of professional boxers find that 95% of registered boxers have fewer than three fights in their careers, and that the theoretical risk of concussive injury from sparring is almost non-existent.

The other major risk factor for chronic traumatic brain injury is genetic. Recent studies show that boxers with the apolipoprotein E4 (apoE4) allele are susceptible to chronic neurological deficits. Male boxers who have 12 or more professional fights, as well as the ApoE4 allele are 16 times more likely to have neurological deficits than those without the allele. The ApoE4 allele has also been linked to poor neurological outcome after traumatic brain injury from any cause.

The precise incidence of chronic traumatic brain injury is difficult to measure, and it may largely be a condition of historical interest. Few prospective epidemiological studies have been performed in boxers, and often they do not distinguish between amateur boxing and professional boxing. A dose-response effect has been suggested, whereby professional boxers have a higher rate of chronic traumatic brain injury than amateurs because of greater exposure to head impacts—bouts are longer (12 rounds versus three rounds) and they do not wear protective headgear. However, this has never been formally tested. Given the quality of the published literature, it is not surprising that Loosemore and colleagues find little conclusive evidence for chronic traumatic brain injury in amateur boxing.

The difficulty with extrapolating early studies to today’s sport is that the nature of the sport has changed substantially. In the 1930s to 1950s, boxers’ careers generally lasted 10-20 years, started in childhood, and involved up to 1000 professional fights. Many boxers also became professional sparring partners or boxers in tents and booths, where they fought up to 30-40 unsupervised bouts each day. Fighters were not matched by skill or weight, they had no medical supervision, and they fought with 6 oz gloves. Bouts were often not stopped even when a boxer was overmatched, and bouts lasted longer (up to 20 rounds of two minutes...
each). There was no mandatory exclusion after a knock-out or head injury. Because of the depression in the 1930s, financial reasons kept many boxers competing, despite the onset of neurological symptoms.

No compelling evidence is available to suggest that regular magnetic resonance imaging of the brain, rigorous medical supervision, or currently practised safety measures will influence or prevent the development of chronic traumatic brain injury. However, because today’s boxers have shorter careers and reduced exposure to repetitive head trauma, the likelihood of this condition developing is probably low. Whether governing bodies should recommend or mandate genetic testing for the ApoE4 allele in prospective boxers is an ethical question that needs to be debated. One of the reasons for doing so would be to provide an opportunity to counsel boxers about their risk of injury.

Breast feeding and the risk of allergy and asthma

New trial shows no reduction in risk

The possibility that breast feeding might protect against allergy and asthma has generated interest for 70 years. In this week’s BMJ, a cluster randomised trial by Kramer and colleagues assesses whether exclusive and prolonged breast feeding reduces the risk of allergy and asthma at 6 years of age. It found no significant difference in allergy and asthma symptoms reported by parents or the results of allergy skin prick tests.

Hospitals in Belarus were randomised to promote breast feeding or usual care, and mothers intending to breast feed were eligible. The intervention increased the total duration of breast feeding and exclusive breast feeding in the intervention group. Six years later, parents answered seven questions about wheezing, hay fever, itchy rash, and whether their child had ever had asthma or eczema. The children also had skin prick tests to determine hypersensitivity to five airborne allergens. Overall, 10% of parents reported that their child ever wheezed, 5% that they ever had symptoms of hay fever, and 1% that they ever had asthma, with no significant difference between intervention and control groups. Positive skin prick tests were more common, with 27% of children having more than one positive test, but again there was no significant difference between the two groups.

The trial overcomes many of the challenges inherent in studying the influence of breast feeding on health outcomes. Assigning mothers to breastfeeding promotion or usual care eliminates the confounding inherent in observational studies. The cluster design allows better estimation of effects within each intervention group. Furthermore, the design includes prospective collection of high quality data on feeding when the children were 3, 6, 9, and 12 months, with standardised definitions for exclusive and any breast feeding.

The limitations of this study include a highly selected sample, comparison of two relatively similar breastfeeding groups, and the validity of the outcome measures. It is appropriate to select mothers intending to breast feed when testing the efficacy of a programme to promote breast feeding as this improves the duration of total and exclusive breast feeding. However, it limits external validity, because women who choose to breast feed may differ from those who do not in characteristics related to allergy and asthma outcomes, such as geography and socioeconomic status.

Although large differences were seen between the duration of breast feeding in the two groups, all women started breast feeding, and even in the control group 36% were still breast feeding at 6 months. Only 6.4% of the control group were exclusively breast feeding at 3 months compared with 44.3% of the intervention group, but many more may have been exclusively breast feeding at an earlier time point, such as 6-8 weeks. Hypothetically, exclusive breast feeding in the early weeks might be protective. It is possible that the groups were not divergent enough to answer the question of whether breast feeding protects against allergy and asthma.

The outcome measures also need to be considered. The reported prevalence of asthma was five times lower than the expected rate in the United Kingdom or the United States. Possible explanations include a lower prevalence of childhood asthma in this sample from Belarus compared with the UK and US; under-reporting or underdiagnosis of asthma in this sample; or lower prevalence of asthma in both the intervention group and the control group related to a.
common factor, such as the high initial breastfeeding rate. The second outcome, positive skin prick tests, is also problematic. Skin prick tests are better negative predictors than positive predictors and in clinical practice are recommended only as confirmatory tests for people with symptoms. A test with a positive predictive value of 11.9% for hay fever may not have adequate specificity to determine if breast feeding is associated with allergy.

The finding that promoting breast feeding did not reduce hay fever, eczema, or asthma reported by parents or result in fewer positive skin prick tests despite large increases in the duration of exclusive breast feeding calls into question previous findings of associations between breast feeding and decreased risk of allergy and asthma. Although this study must be interpreted cautiously—taking into account its limitations—previous work on this question is conflicting. For the moment, promotion of breast feeding should include evidence that it reduces the incidence of a wide range of infectious diseases, including diarrhoeal diseases and lower respiratory tract infections. Evidence that it reduces the incidence of other conditions including diabetes, obesity, and some cancers is emerging. Furthermore, breast feeding has health benefits for the mother. Therefore, there is already ample evidence to promote breast feeding as a public health measure. None the less, the claim that breast feeding reduces the risk of allergy and asthma is not supported by evidence.

In this week’s BMJ, von Elm and colleagues report the STROBE (strengthening the reporting of observational studies in epidemiology) statement, which recommends what should be included in an accurate and complete report of an analytical observational study.

Observational epidemiology has made an immense contribution to our understanding of the causes and treatment of disease. Numerous causal associations between risk factors and disease have been identified (see box on bmj.com). Most of these observations have led to substantial improvements in public health by causing changes in policy or by leading to the development of effective treatments.

Observational studies are also essential for effective clinical practice. Cohort studies allow us to improve the reliability of diagnosis; to understand prognosis; to develop and validate risk scores to target treatment appropriately; to monitor the safety of treatments in routine practice; to identify treatment effects (adverse or beneficial) that are not reliably detected in trials (perhaps because they are too rare, have too long a latency, or are confined to people excluded from trials); and to estimate the effects of interventions in circumstances in which randomised trials are not feasible.

To make the most of the enormous potential of observational epidemiology to transform clinical practice and improve public health, studies must be designed and reported as rigorously as possible. However, as with other areas of research, including laboratory sciences and randomised controlled trials, the design and reporting of epidemiological studies can be poor, with consequences for the reliability of results.

Quality control is unlikely to improve in the near future, given the ever increasing number of medical journals, and the consequently reduced influence of peer review on the likelihood that poor quality research will be published. The STROBE guidelines on the reporting of epidemiological studies are therefore welcome. The summary paper published in this week’s journal will be backed up by a more detailed document, which will explain the background and justification for each guideline. Such guidelines inevitably have limitations, and there is always a risk that poorly designed studies will be made more difficult to spot by superficial improvements in the way they are reported. However, experience with similar guidelines for reporting randomised trials and

Reporting of observational studies
New recommendations should help researchers, journal editors, and readers


ANALYSIS, p 806

Peter M Rothwell professor of clinical neurology, Meena Bhatia research fellow, Stroke Prevention Research Unit, University Department of Clinical Neurology, Oxford OX2 6HA peter.rothwell@clneuro.ox.ac.uk

Competing interests: None declared.

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

BMJ 2007;335:783-4. doi:10.1136/bmj.39351.581366.BE

EDITORIALS
Ototoxicity caused by aminoglycosides

Is severe and permanent in genetically susceptible people

Aminoglycoside antibiotics are widely used for the treatment of Gram negative sepsis. It is well known that they can cause dose related renal toxicity and ototoxicity, which occur in almost everyone who receives a sufficiently toxic dose.\(^1\) It is less well known that some people have an inherited predisposition that renders them highly sensitive to the ototoxic effects of these antibiotics: aminoglycosides taken at levels that are well within the therapeutic range can result in rapid, profound, and irreversible hearing loss. Even a single dose in a predisposed individual can result in permanent hearing loss.\(^5\)

In countries that use aminoglycosides widely, a quarter of people with hearing loss induced by aminoglycosides have maternal relatives who also have deafness related to drug induced ototoxicity.\(^1\) In the familial cases of hearing loss, individuals received antibiotics for a much shorter period than those without a family history of ototoxicity, suggesting the presence of an inherited predisposing mutation. The most common predisposing mutation is now known as m.1555A>G, a mitochondrial DNA mutation. The most common predisposing mutation is now known as m.1555A>G, a mitochondrial DNA mutation.


\(^2\) Apfalter P, Reischl U, Hammerschlag MR. In-house nucleic acid amplification assays in research: how much quality control is needed before one can rely upon the results? J Clin Microbiol 2005;43:5835-41.


Mutation at position 1555 of human mitochondrial DNA makes the human mitochondrial ribosome even more similar to the bacterial one, which facilitates aminoglycoside binding. Once bound, aminoglycosides have a long half life in the hair cells of the inner ear (several months), which increases the risk of ototoxicity. How common is the m.1555A>G mutation? To date, no large prevalence studies have been performed and data can only be extrapolated from small studies. In the US state of Texas, screening of blood spots from 1161 newborns found one positive case, and in New Zealand there was one positive case among 206 random blood samples screened (0.48%; 95% confidence interval 0.01 to 2.75).6 This prevalence is much higher than previously suspected from calculations of its contribution to childhood deafness.

In the United Kingdom about 1 in 1000 children are born deaf; half of these cases have a genetic cause, with about 80-85% caused by recessive genes, 10% by dominant genes, and 2-5% caused by the m.1555A>G mutation.7 This indicates a prevalence of the m.1555A>G mutation of 1 in 40000. The discrepancy between this and the prevalence in New Zealand and Texas implies that either the prevalence in the UK is very much lower or penetrance of the mutation is very low, meaning that more people have the mutation but are not deaf. As aminoglycosides in the UK are used only in hospitals, penetrance is likely to be low in the absence of exposure to aminoglycosides. A genuine population frequency of between 1 in 206 and 1 in 1161 would have substantial implications for clinical practice in terms of the numbers of people at risk of ototoxicity.

Even in the absence of exposure to aminoglycosides, some families carrying this mutation may also develop deafness, albeit at a later age and with a lower penetrance. The variable penetrance of the m.1555A>G mutation may be attributable partly to the presence of a modifying nuclear genetic mutation.8 In some populations, the m.1555A>G mutation seems to be a common cause of deafness. In Spain, 27% (19/70) of families with at least two deaf individuals were positive for this mutation.9 Everyone with the mutation who was exposed to aminoglycosides became deaf. The probability of becoming deaf by the age of 30 years if an individual had received such antibiotics was 96.5% compared with 39.9% if they had never been treated. Thus aminoglycosides are a major environmental modifier of the m.1555A>G mutation. Because penetrance of the mutation is very low in some families (0-18%), exposure to aminoglycosides may cause drug-induced deafness that may be erroneously categorised as sporadic.10 11

Is it cost effective to screen for this mutation before aminoglycosides are given? Cost effectiveness is determined by the cost of a screening test and the prevalence of the mutation versus the cost of not screening. The current cost of testing for this mutation in the UK is about £35 (£52; $71) per test, based on a small number being performed (generally in those who have already lost their hearing after aminoglycoside administration). However demand for more tests would reduce the unit costs, and single nucleotide genotyping in the commercial sector costs pennies per genotype. Conversely, the cost to the health service of providing a cochlear implant for a child who becomes deaf before acquiring language and of maintaining the implant for 15 years is estimated to be about £47000 per child, rising to £61000 over a child’s lifetime.12 Educational costs for a profoundly deaf child with a cochlear implant are estimated at about £18000 a year.13 However, the cost of not providing a cochlear implant to a profoundly deaf child is even greater in terms of educational costs and eventual earning power. In the US, the total lifetime cost to society for a child with prelingual onset of profound deafness has been estimated to exceed $1m.14

Hearing loss induced by aminoglycosides in individuals with the m.1555A>G mutation is in theory preventable. The mutation is well known among doctors who see patients who already have hearing loss. However, the general medical community is not aware of this susceptibility and that mutation testing is available through regional genetics centres. We recommend that the true prevalence of the mutation in the UK be ascertained to determine the cost effectiveness of screening everyone prescribed aminoglycoside antibiotics. In the meantime, patients who are likely to receive multiple courses of aminoglycosides—for example, patients with leukaemia and newborns admitted to special care baby units—should be screened. Genetic testing needs to be turned around rapidly, and consideration should be given to using an alternative antibiotic until the result of genetic testing is known.

Management of chronic knee pain

Acupuncture has no additional benefit in people taking a course of exercise

Two papers have recently been published on bmj.com on the treatment of osteoarthritis of the knee.1 2 The first is a randomised trial of adding acupuncture to a course of advice and exercise delivered by physiotherapists;3 the second, which is also published in this week’s BMJ, is a systematic review of the effectiveness of physiotherapy after elective total knee arthroplasty in people with osteoarthritis.2

Clinical trials conducted over the past decade have helped to define the role of acupuncture in various clinical conditions. A particular focus of these trials has been the use of acupuncture for chronic knee pain or osteoarthritis of the knee.3

The findings of randomised trials of acupuncture have caused much debate. Positive trials have been criticised because of inadequate blinding. Negative trials have been criticised because the intervention was not administered by properly trained practitioners or because control interventions may have had analgesic effects. However, a systematic review of high quality randomised controlled trials suggests that acupuncture can reduce pain and disability in people with chronic pain.4

Despite this evidence the role of acupuncture in the management of chronic knee pain is still unclear. Foster and colleagues1 argue that acupuncture is useful only if it adds to the benefits of the first line treatments of exercise and advice. They investigated whether acupuncture is useful for people receiving exercise and advice by randomising 352 adults with osteoarthritis of the knee to advice and exercise, advice and exercise plus acupuncture, and advice plus sham acupuncture.

The trial found that acupuncture did not significantly reduce pain (measured on the Western Ontario and McMaster Universities Osteoarthritis index [WOMAC] subscale) at six months compared with sham acupuncture when combined with advice and exercise. This finding agrees with another large well designed trial that compared acupuncture with sham acupuncture given in addition to exercise in people with knee osteoarthritis.4 A pooled estimate from these two studies shows that acupuncture does not significantly reduce pain compared with sham acupuncture (reduction in pain score on the 10 point WOMAC subscale 0.1 points, 95% confidence interval 0 to 0.2).

A systematic review by Minns Lowe and colleagues,9 published in this week’s BMJ, assesses the effects of physiotherapy exercise programmes given after total knee replacement surgery in people with osteoarthritis. The review found a small to moderate effect of functional exercise on joint motion and quality of life at three to four months after surgery, but the effect was not sustained at one year.

The findings should be considered provisional at best. In four of the six included trials, all study participants received an exercise or physiotherapy programme after discharge from the acute hospital3 4; these trials cannot tell us about the effectiveness of such programmes because the control groups also received an exercise intervention. The two remaining trials9 10 focused on the effects of outpatient programmes on the range of knee flexion and found little or no effect on this outcome. Most of the trials evaluated low intensity exercise programmes provided soon after surgery. More lengthy and intensive physiotherapy exercise programmes may be needed to overcome the considerable deficits in muscle strength and endurance that are evident in these patients.

What conclusions can be drawn from these studies? The findings of the trial by Foster and colleagues suggest there is little point in recommending acupuncture to people with chronic knee pain who are already undertaking a course of exercise.1 Acupuncture might be recommended to people who do not exercise.3 It is difficult to make clinical recommendations on the basis of Minns Lowe and colleagues’ review, although it does highlight the lack of research into the effectiveness of physiotherapy exercise programmes after total knee replacement.2

also increased the number of diagnoses positive criterion for a diagnosis of LVH

The alternative of combining ECG measurements to generate a test with a relatively low sensitivity but a high specificity is a pragmatic one.

Peter J Bourdillon honorary senior lecturer ECG Department, Hammersmith Hospital, London W12 OHS pbourdillon@msn.com

Ethnicity is relevant

Pewson et al highlight the danger of using electrocardiography (ECG) for detecting left ventricular hypertrophy (LVH), particularly as it has low sensitivity. They conclude that no criteria are superior to the Sokolow-Lyon criteria. Our recent review supports the first, but not the second, conclusion.

Bourdillon (previous letter) emphasises the need to take into account age, sex, and ethnicity. In a systematic review of the literature, we identified five studies comparing the sensitivity and specificity of ECG (using the Sokolow-Lyon and Cornell criteria) for detecting LVH in white and black (African origin) populations.\(^3\,^5\)

Specificity was high using both sets of criteria in white populations (Cornell 87.4%, Sokolow-Lyon 88.9%) but was much lower in black groups using the Sokolow-Lyon criteria (72.1%). Specificity was higher in black groups using the Cornell criteria (86.2%). Some evidence suggested that Cornell criteria were more sensitive than Sokolow-Lyon criteria in black populations.

Our evidence favours the Cornell criteria over the Sokolow-Lyon criteria. While we agree with Pewson et al that ECG is not sufficient for diagnosing LVH, we emphasise that it is not equally valid across ethnic groups.

Andrew P Vanezis foundation doctor, Colchester General Hospital, Colchester CO4 5JL
Bruce and John Usher professor of Public Health Raj Bhupal, University of Edinburgh Medical School, Edinburgh EH8 9AG andrewvanezis@doctors.org.uk

Competing interests: None declared.


2 Sokolow M, Lyon TP. The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. Am Heart J 1949;37:161-3.


GPs’ 24 HOUR RESPONSIBILITY

Summary of responses

Slightly aggrieved at the suggestion that general practitioners’ jobs are daytime only, many of the respondents to the head to head on whether GPs should resume 24 hour responsibility for their patients remind us that GPs do still provide out of hours care, albeit in different organisational set-ups.\(^1\,^2\)

Most think that...
Extending GPs’ working hours back to those before the 2004 contract is neither feasible nor desirable—mainly because of increased workloads, doctors’ and patients’ safety, and a total lack of financial incentives. Few think that GPs should be expected to have to “opt in” again.

Respondents are indignant at the government and primary care trusts for not fully accepting that organising out of hours care is their responsibility under the new contract; for feeding the public perception (via the media) of GPs as overpaid, greedy, and lazy; for creating unrealistic expectations in patients as healthcare “consumers” entitled to have their demands met at all times; and for expecting GPs to do more without adequate remuneration and compensation.

GPs working for out of hours providers are as well qualified and experienced as any others, they argue, and NHS complaints have increased in total, not especially for out of hours care. Many might consider providing out of hours care with the right “package.” Others do so successfully in local cooperatives staffed by doctors and other specialists, and calls have gone down.

And the way forward? One recommendation is specialists in primary care out of hours services, a separate, defined specialty with recognised qualifications and bespoke or mandatory training.

Birte Twisselmann assistant editor, bmj.com
BMJ, London WC1H 9JR btwisselmann@bmj.com
Competing interests: None declared.
1 Jones R. Should general practitioners resume 24 hour responsibility for their patients? Yes. BMJ 2007;335:696. (6 October.)
2 Herbert H. Should general practitioners resume 24 hour responsibility for their patients? No. BMJ 2007;335:697. (6 October.)

US LIABILITY

Damned if you treat, damned if you don’t

The UK government recognised the need for standardisation in the adventure travel market1 in this year’s BS8848 document, which covered the need to provide medical support by a recognised medical practitioner. All was well until the implications of providing medical cover to Americans, among others, was highlighted. Defence unions advise doctors that although they are covered to treat Americans, they are not covered for court cases that arise in North America.

(Americans can sue a doctor in America, independent of where the transgression occurred.) The General Medical Council (GMC) advises doctors not to participate in activities without appropriate cover.

However, if a doctor does not treat an American on an expedition, the American participant can sue for racial discrimination and report the doctor to the GMC for improper conduct. The defence unions escape culpability in the eyes of the racial discrimination board as they are not discriminating against Americans but against legal action taken in the US, whether by an American or a UK citizen. What about Americans living permanently in the UK? Or dual nationals who retain their American status but are also UK citizens? They fall into the same bracket, so the advice to the doctor is the same.

How can this untenable situation be rectified? Could something as simple as a signed legal waiver work, or will the defence unions have to accept the risk? Whatever the solution the situation cannot be allowed to continue as it is, with doctors risking being sued for treating or being sued for refusing to treat.

Sean T Hudson medical director, expedition medicine
St Paul’s Medical Centre, Carlisle, Cumbria CA1 1AA
sean@expeditionmedicine.co.uk
Competing interests: None declared.

BISMARK v BEVERIDGE

Unfair comparison

The Euro Health Consumer Index 2007 is being cited as evidence that the Bismark system “delivers better value” than the Beveridge system.1 One might, however, pause to consider World Health Statistics 2007 from the World Health Organization (www.who.int) as shown in the table.

Surely it is premature to draw conclusions on the merits of one system over another when the playing field is far from level?

Iain S Fraser general practitioner
Manchester M12 5HL iands@doctors.org.uk
Competing interests: None declared.
1 Watson R. UK does well on giving information to patients but poorly on access to new treatments. BMJ 2007;335:666. (6 October.)

VIVE LA DIFFERENCE?

Training and careers in France

The United Kingdom is not alone in its uncertainty about doctors’ training and careers.1 Other European countries, and particularly France, are struggling with a shortage of junior doctors and an uneven distribution of doctors across the country, leading to potential issues regarding the immigration of doctors to fill the gaps.

In France, medical schools admitted 8000 students in 1975, 6000 in 1980, 4000 in 1990, and 3500 in 2000, and it was urgently decided to train 7100 students in 2007. Thus we are still following short term reasoning, without considering other factors such as the feminisation of the workforce, the decrease in working hours, the increasing gap between graduation and beginning professional activity, early retirement, the quest for a better quality of life, the place of other health care professionals (specifically nurses), and the migration of doctors around the world.

The distribution of doctors across France is also a subject of debate as there are discrepancies between regions, with more doctors per capita in the south than in the north. Many villages in the countryside have no doctors, and too many specialists are competing in large cities.

In hospitals, vacant positions are filled by poorly paid foreign doctors. In 2007 there are thousands of doctors who qualified abroad and are employed in hospitals without having passed any serious selection process.

We need to assess competencies of all doctors throughout their career, irrespective of the country where they qualified. As in many countries, in France, this reflection on competency assessment started in 2002,2 after the Bristol affair.

We should allocate funds to organise conferences and observe the immigration of doctors. This should avoid making short term opinion based decisions, and allow for long term decisions to be taken based on research data.

Hervé Maisonneuve director, continuing medical education Pfizer, 75014 Paris, France
Yves Matillon head, health professions competencies evaluation task force, Health Ministry, Paris
hervemasonneuve@gmail.com
We thank Chloe Brown for editing this paper.
Competing interests: None declared.
1 Godlee F. Editor’s choice. Training our doctors. BMJ 2007;335:0. (22 September.)

Public healthcare statistics from the WHO

<table>
<thead>
<tr>
<th>Measure</th>
<th>Germany</th>
<th>United Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per capita total expenditure($)</td>
<td>3521.4</td>
<td>2899.7</td>
</tr>
<tr>
<td>Hospital beds/100 000 people</td>
<td>844.49</td>
<td>389.79</td>
</tr>
<tr>
<td>Doctors/100 000 people</td>
<td>340.20</td>
<td>389.79</td>
</tr>
<tr>
<td>TB=£0.5=€0.7.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experts clash over reducing abortion limit

Clare Dyer BMJ

Health experts clashed this week over the long term health risks for women who have terminations as the House of Commons science and technology committee launched its inquiry into whether the abortion laws in England and Wales need updating in the light of current knowledge.

Doctors who gave evidence at the committee’s first session agreed that the research showed that women who have induced abortions are more likely to have premature births in future—and the more terminations they have, the greater the risk.

But there was disagreement about whether abortion increases the risk of mental health problems in later life and whether women who have terminations are more prone to breast cancer.

The MPs are considering whether medical research indicates that the 24 week normal cut-off point for abortion should be reduced.

The Pro-Life Alliance wants the upper limit cut to 20 weeks, but the BMA says that the number surviving at 24 weeks is still “extremely small” and argues that the limit should stay as it is.

There were 193,000 abortions in England and Wales last year, 89% of them performed in the first 13 weeks.

The MPs were told that a new study of the most premature babies, EPI Cure2, was expected to find no evidence of recent improvement in chances of survival. The British Association of Perinatal Medicine said that initial findings from the study did not support arguments for the upper limit to be reduced.

Rates of survival for babies born below 24 weeks’ gestation in 2006 were 10-15%, the same as in 1995, the association said in a written submission.

But John Wyatt, professor of neonatal paediatrics at University College London, said “substantial numbers” now survived at 23 weeks at his specialist unit.

Obesity report provokes government action

Lisa Hitchen LONDON

The government has signalled its support for a report on obesity after furious criticism at a National Obesity Forum conference this week for years of inaction over the epidemic.

The public health minister Dawn Primarolo said that the Department of Health would work with Foresight, the government’s futures think tank, to try to translate the report’s findings into action.

The independent report, Foresight Tackling Obesities: Future Choices, which is part of the group’s work for government, predicts that 40% of Britons will be obese by 2025, with the country being mainly obese by 2050.

Health problems caused by obesity will place an “intolerable burden” on the Treasury in terms of health costs; on employers through staff absence; and on families having to cope with long term disability, says Foresight.

Obesity will cost society £46bn (€65bn; $92bn) a year by 2050, the report estimates.

The obesity epidemic has many causes, said Peter Kopelman, one of the report’s authors and director of the faculty of health at the University of East Anglia.

“It is because biological mechanisms are poorly adapted to the 21st century. The pace of the technological revolution has outstripped human evolution,” he said.

Changes have meant transport, work, and food patterns have altered so radically that solutions to the problem would require a “paradigm shift in thinking” by government, individuals, and society, the report says.

Although drugs and devices provide some help with obesity, they will never be full solutions. Neither will changing behaviour at an individual level, it says.

The report calls for change at an environmental and organisational level as well as in the “obesogenic environment” to reduce obesity at a population level.

Susan Jebb, one of the report’s authors and head of nutrition research for the Medical Research Council, said, “Until now there has been far too much emphasis on a headline grabbing initiative here and there, but now the penny has dropped that it has got to be more systemic than that.”

Ms Primarolo pointed out that the responsibility for tackling obesity did not lie with one government department but with all of them working together. “It means having thinking from education to the built environment, from planning to transport to what we eat and why we eat it,” she said.

The report is available at www.foresight.gov.uk.
Doctors need a “sea change” in their attitude to C difficile

Adrian O’Dowd MARGATE

An NHS trust could face criminal charges over failures in its infection control measures that contributed to the deaths of 90 patients.

Every NHS trust in England will now get a copy of a damning report into failures at the Maidstone and Tunbridge Wells NHS Trust in Kent as a precautionary measure to highlight mistakes that occurred there.

The Healthcare Commission, England’s health watchdog, has just published its report after an investigation into outbreaks of *Clostridium difficile* at the trust’s hospitals between April 2004 and September 2006.

The Healthcare Commission said that there were “significant failings” at the trust and that of 345 deaths it considered as part of its review 90 were “definitely or probably” caused by *C difficile*. The police and the Health and Safety Executive are now considering whether criminal charges should be brought against the trust over the deaths.

During the period under study more than 1170 patients were infected across the trust’s three hospitals.

The report details several contributory factors, including:

- Failure of the trust board to deal with problems raised consistently by patients and staff
- Shortages of nurses, poor care of patients, and poor processes for managing movement of patients from one ward to another
- Average rates of daily bed occupancy of more than 90% (and in some cases more than 100%) on medical wards where patients were not receiving surgery
- High turnover of patients, limiting the time available to clean beds between patients.

The report’s author, Heather Wood, said, “One of the things we would like to think the medical profession might pick up on particularly is our concern about the standards of basic medical care and management there, as well as the regular review of patients.”

Dr Wood said she had concerns about the general approach and attitude of staff to *C difficile*. “I would hope that this report would bring about a degree of sea change in the way in which doctors regard *C difficile*,” she said.

*Investigation into Outbreaks of Clostridium difficile at Maidstone and Tunbridge Wells NHS Trust is at www.healthcarecommission.org.uk.*

West Yorkshire has most road crash casualties

Roger Dobson ABERGAVERN

An analysis of deaths and injuries from road traffic incidents has identified West Yorkshire and the West Midlands as the counties with highest risk of injury and the Isle of Wight and parts of Wales as the safest (*Health and Place* 2007; doi: 10.1016/j.healthplace.2007.10.001).

Between 1995 and 2000 there were 15 797 deaths, 191 870 serious injuries that needed hospital treatment, and 1 282 563 minor injuries from road traffic incidents in England and Wales.

For the NHS funded study researchers looked at geographical variations in mortality and morbidity throughout England and Wales.

The results were presented in two ways for each county. The first, the null model, was the total number of deaths and injuries. The second took into account differences in population, vehicle movements, car ownership, road length and curvature, and other factors, which made some differences to the ranking orders.

The results for the null model show that for deaths West Yorkshire was top, followed by West Midlands, South Yorkshire, and Lincolnshire. The counties with the lowest mortality were the Isle of Wight, West Glamorgan, and Gwynedd.

For serious injuries, the West Midlands was ranked in first place followed by West Yorkshire, Greater London, and South Yorkshire. The lowest risk was in West Glamorgan followed by neighbouring Mid Glamorgan.

The authors say that fatalities from road traffic incidents in Britain have averaged more than 3000 a year since 1998.

NICE guidelines create ethical dilemmas in care of elderly people, says international report

Anne Gulland LONDON

A report that looks at problems affecting elderly people around the world has warned that the United Kingdom’s policy of prescribing drugs on the grounds of cost effectiveness is damaging the human rights of older people.

The report, by the International Longevity Centre, warns that the National Institute for Health and Clinical Excellence has created “new ethical dilemmas about allocation of scarce resources” for older people.

It says, “Prescribing drugs according to cost effectiveness may be the opposite of the rights based approach: decisions can condemn patients to deteriorate before the drug will be prescribed, as is the case with Aricept [donepezil] for dementia patients.”

The report also called for the government to act on a recent House of Lords test case, which ruled that private care homes fall outside the scope of the Human Rights Act (Y L v Birmingham City Council and others, HL 20 June 2007). The act covers only homes run by local authorities.

The report says, “The legal chess game may in the long term be the best way of producing a durable result, but in the meantime vulnerable people are left in a wholly unacceptable limbo.”

Frances Butler, vice president of the British Institute of Human Rights, urged the government to plug the loophole.

“This is a serious defect because nine out of 10 care home places are provided by private companies or by charities. Residents cannot properly obtain protection of their human rights when the people who are breaching them do not have human rights responsibilities.”

Ms Butler also urged the government to legislate to protect older people from unfair age discrimination in the health service and other public bodies.

“Women are not invited to breast cancer screening when they reach their 70th birthday. This must make women over 70 think that the government does not consider their health to be worth anything anymore. Age should not be used as a blanket proxy for risk. A more sophisticated scheme is needed,” she said.


COX 2 inhibitor rejected in North America but retained in Europe despite liver risk concerns

Bob Burton CANBERRA

Doctors in the UK will continue to be able to prescribe the cyclo-oxygenase-2 (COX 2) inhibitor Prexige (lumiracoxib) the UK Medicines and Healthcare Products Regulatory Agency has decided, even though it has been withdrawn from the market in Canada and the US Food and Drug Administration (FDA) has refused to approve it.

Health Canada, the government agency with responsibility for drug registration, reviewed the safety of the 100 mg dose of lumiracoxib, following the withdrawal of the 100, 200, and 400 mg doses of lumiracoxib from the Australian market in August (BMJ 2007;335:363). On 4 October the department announced that it had withdrawn marketing approval of the 100 mg dose of lumiracoxib due to the risk of “serious liver-related adverse events” including hepatitis. The department, which had first approved the drug in November 2006, reported that two cases of liver damage in Canada had been associated with the drug and concluded that “the risk of serious liver-related adverse events with Prexige cannot be safely and effectively managed at the 100 mg daily dose.”

While both Canada and Australia have withdrawn marketing approval for the drug, the 100 mg daily dose remains available in more than 50 countries, including the UK and other countries in the European Union and Latin America.
Surgeons get new training facility to practise operations

Zosia Kmietowicz LONDON
Trainee surgeons at the Royal College of Surgeons’ new Wolfson Surgical Skills Centre, in London, can now practise the techniques required for aligning joints during orthopaedic operations, such as knee replacements. The opening of the £3m (€4m; $6m) surgical skills facility this week marks the first stage of the college’s project to provide with a world class centre by 2010.

The Human Tissues Act means that for the first time surgeons in training can practise surgical techniques on donated human bodies before they enter an operating theatre. The fact that the centre’s tables are fully interconnected by monitors means that up to 50 surgeons can learn collaboratively at any one time. And trainees no longer need to travel to the capital to brush up on their skills. A high resolution video wall means that education can be run remotely “with courses run from the college and delivered to trainees locally, nationally, and internationally,” said Dick Rainsbury, director of education at the college.

With the development of more than 150 new surgical procedures in the past 10 years, the demand for training has never been higher. “There are expert surgeons based all around the country, now all trainees can have access to that knowledge no matter where they are based,” said Bernard Ribeiro, president of the college.

74 year old Dutch GP refuses to take part in deputising service

Tony Sheldon UTRECHT
A long running dispute over working practices between the Dutch medical authorities and a septuagenarian general practitioner, portrayed as a cause célèbre for older people’s rights, is to go before the Netherlands’ highest administrative court.

The Council of State, which advises on legislation and governance, will judge whether a 74 year old can be stopped from working as a GP if she refuses to join colleagues in out of hours deputising services.

Dr Tonny Bakhoven from Driebergen near Utrecht says she prefers to see her patients herself during nights, weekends, and holidays. But the certification committee of the Dutch Medical Association, which has a role similar to the UK’s General Medical Council, insists the case is not about age but about meeting the quality requirements for recertification.

All GPs, of whatever age, must apply for recertification every five years to ensure they maintain the quality of care and take part in continuing medical education. One of those requirements is participating in mutual deputising, or locum, services. This, argues the Dutch society of general practitioners, ensures they gain sufficient experience of emergency care, especially with patients they are unfamiliar with.

However this has become a problem for Dr Bakhoven, one of an estimated 190 working GPs in the Netherlands aged over 65. Dr Bakhoven, who qualified in 1958 and has run her current practice for 20 years, believes the system of deputising services is wrong. “I come from a generation where a doctor had their own practice and always worked every night and weekend. I still do that. When something serious happens I want to be involved.”

She considers that the greater free time expected by younger doctors and the idea of deputising is her profession’s downfall. Since her five year certification became due for renewal two years ago, the Medical Association has granted her annual certification to allow her time to meet its requirements.

Dr Bakhoven has attempted to join the deputising services but, she admits, is slow at using its computer system. Dr Bakhoven’s certification has now been withdrawn. She has mounted a legal challenge and the case has been appealed by the general practitioners’ association to the Council of State.

Dr Lourens Kooij, secretary of the medical association’s certification committee, said Dr Bakhoven’s activities “do not satisfy the requirements for recertification in that she does not do any deputising services with other GPs . . . that is a demand in the Netherlands.”

He added: “This has nothing to do with age.

“Last year we renewed the certification of a GP who is 83 and who met all the requirements.”

Dr Tonny Bakhoven admits she is slow using the computer system in the deputising service
Surgeon is suspended over “unconventional donations”

Annette Tuffs HEIDELBERG

A prominent transplant and cancer surgeon from Essen, Christoph Broelsch, has been suspended by Essen University following accusations of fraud and blackmailing patients. Professor Broelsch is also being investigated for tax evasion over “unconventional donations” allegedly made by patients.

Essen University Hospital is one of the largest centers for transplant and cancer surgery in Germany, which has been led by Professor Broelsch since 1998. He is known to be one of the pioneers of liver transplantation.

In May 2007 Professor Broelsch was publicly accused by the relative of a patient with liver cancer of having asked for money to bring forward the date of an operation. Police then began an investigation into whether such offers had been made to other patients, including patients from overseas.

After the accusations were made, Professor Broelsch issued a statement denying that he had ever sought financial reward to perform an operation or blackmailed any patient. However, he said that in certain cases when patients without private health insurance asked him to perform an operation he was obliged by German law to tell them that they had to pay extra for his involvement. He said that in some cases he had offered to forgo his extra payment if the patient made a contribution to his research projects.

At first the hospital merely said that “unconventional donations” had been made, but on 9 October it issued a statement that the situation had changed and that suspension was necessary.

Professor Broelsch has not made any comment following his suspension and could not be contacted by the BMJ.

NHS gets 4% increase in funding a year until 2011

Helen Mooney LONDON

The NHS in England and Wales has been given a 4% a year funding rise for the next three years, with an overall budget rise from £90bn this year to £110bn in 2010.

Announcing the comprehensive spending review, the chancellor, Alistair Darling, said that the cash would fund 20 new hospitals, 140 walk-in centres open seven days a week, and 100 new GP practices.

However, the funding is much less than the 7.2% that the NHS has had each year since 2002 and below the 4.4% recommended by Derek Wanless in his review of the future of the NHS published in September (BMJ 2007;335:529).

But it is more than the 3% that many NHS staff had been expecting.

Hamish Meldrum, chairman of the BMA, said it was vital funds were not squandered on “costly and poor value” deals with the private sector, where he said “profits are often rated higher than patients.

“Excessive use of the private sector in providing NHS care will fragment care for patients, could threaten the existence of many district general hospitals, and risks destroying the proven and trusted model of UK general practice,” he said.

The Pre-Budget Report and Comprehensive Spending Review 2007 is available at www.hm-treasury.gov.uk.

US medical schools have financial ties to drug companies

Janice Hopkins Tanne NEW YORK

Most US medical schools and large teaching hospitals have financial ties to drug companies, according to a survey published this week (JAMA 2007;298:1779-86).

Researchers from Massachusetts General Hospital in Boston, the University of Michigan in Ann Arbor, and the Association of American Medical Colleges surveyed all 125 US medical schools and 15 large teaching hospitals, which often do more research than many medical schools.

They found that 60% of departmental heads had a financial relationship with a drug company as a consultant, member of a scientific advisory board, a paid speaker, an officer, a founder, or a member of the board of directors.

Two thirds of departments at medical schools and large teaching hospitals had relationships with industry that involved research equipment, unrestricted funds, support for research seminars, residency and fellowship training, continuing medical education programmes, discretionary funds to buy food and drink, support for professional meetings, subscriptions to professional journals, and intellectual property licensing.

A spokeswoman for the Pharmaceutical Research and Manufacturers of America told the BMJ that the organisation was unable to comment on the study because it did not collect information on the industry-academic ties reported in the article.
German hospitals demand extra cash to alleviate problems

Annette Tuffs HEIDELBERG

The German Hospital Association has demanded immediate financial help from the federal government because the care of patients is suffering as a result of large deficits and lack of staff.

Georg Baum, the association’s chief executive, presented a representative survey of 304 general hospitals that shows that a third are running at a loss and a third struggle to stay in balance. Despite the loss of 150,000 jobs in the past few years most hospitals are considering further reductions in numbers of nurses and doctors.

Mr Baum blamed the crisis in part on higher wages for all hospital staff and higher value added tax and costs of energy. Also, the introduction of a new hospital financing system, with insurance companies charged a fixed sum for each case rather than for actual daily costs, had added to the problem, he said.

“The rising costs were not covered by consequent increases of hospital budgets,” he said.

The German government, however, has denied that the situation is urgent and that more frequent increases of hospital budgets, “are required,” he said.

For each case rather than for actual daily costs, the introduction of a new hospital financing system, with insurance companies charged a fixed sum for each case rather than for actual daily costs, had added to the problem, he said.

Last December the agency altered its advice to warn that “powdered infant and follow-on formulas are not sterile, which means they can contain harmful bacteria. So it is important to take care when preparing and storing formula, to reduce the risk of babies becoming ill. Bacteria will be killed if formula is made up following the advice, which includes using water that is at least 70°C.”

Now campaigners want the government to implement in full the International Code of Marketing of Breastmilk Substitutes, which was passed by the World Health Assembly in 1981, because they claim that the existing regulations are too weak.

Michael O’Donnell, head of hunger reduction at Save the Children, said, “The law that is supposed to stop formula milk companies from promoting their products and to protect babies and parents is not working. Formula companies are finding increasingly devious ways to beat the ban and continue to bombard parents with misleading information about the alleged similarities of breast and bottle.”

Andrew Radford, deputy director, Unicef UK, said, “The government has regularly stated that it wants to adopt World Health Organization recommendations, which would prohibit advertisements for formula, but has continually failed to act on these promises in the UK. We are, therefore, calling on the government to prohibit all formula milk advertising, so that parents can feed their babies using accurate information, free from commercial pressure.”

The Baby Feeding Law Group, a coalition of organisations that support UK health professionals and mothers, is calling for restrictions in any new legislation, including a ban on the promotion of all breast milk substitutes (including follow-on formula, specialised formulas, and other bottle fed products).


Government should restrict advertising of baby milk products

Peter Moszynski LONDON

Protestors are urging the UK government to tighten safeguards regarding the sale and marketing of infant formula, as the Food Standards Agency prepares to publish proposals for modifying the regulations that control these products.

Last December the agency altered its advice to warn that “powdered infant and follow-on formulas are not sterile, which means they can contain harmful bacteria. So it is important to take care when preparing and storing formula, to reduce the risk of babies becoming ill. Bacteria will be killed if formula is made up following the advice, which includes using water that is at least 70°C.”

Now campaigners want the government to implement in full the International Code of Marketing of Breastmilk Substitutes, which was passed by the World Health Assembly in 1981, because they claim that the existing regulations are too weak.

Michael O’Donnell, head of hunger reduction at Save the Children, said, “The law that is supposed to stop formula milk companies from promoting their products and to protect babies and parents is not working. Formula companies are finding increasingly devious ways to beat the ban and continue to bombard parents with misleading information about the alleged similarities of breast and bottle.”

Andrew Radford, deputy director, Unicef UK, said, “The government has regularly stated that it wants to adopt World Health Organization recommendations, which would prohibit advertisements for formula, but has continually failed to act on these promises in the UK. We are, therefore, calling on the government to prohibit all formula milk advertising, so that parents can feed their babies using accurate information, free from commercial pressure.”

The Baby Feeding Law Group, a coalition of organisations that support UK health professionals and mothers, is calling for restrictions in any new legislation, including a ban on the promotion of all breast milk substitutes (including follow-on formula, specialised formulas, and other bottle fed products).


Backpedalling from Blair’s privatisation agenda

Nicholas Timmins FINANCIAL TIMES

In the four months since Gordon Brown became prime minister, what has happened to the Blairite agenda of choice, competition, and—most particularly—the use of the private sector as key drivers of NHS reform?

Since 27 June, Alan Johnson, the new health secretary, has announced the cancellation of a third round of independent sector treatment centres—when no third round was in fact ever planned. A whole clutch of contracts for second round independent sector treatment centres and for diagnostics—originally due to be signed off late last year and early this year—remain under permanent review. Only a minority are now expected to go ahead.

Members of a high powered board, which advised the health department’s commercial directorate on these and other central contracts, have resigned together. Board members included a Rothschild banker and senior legal, investment, and private finance specialist. They were, they declared, “wasting their time” because, as one member put it, “no one is listening” since Tony Blair went.

The terms “competition” and “use of the private sector” as means to reform the NHS have been entirely absent from the speeches of Gordon Brown and Alan Johnson. Neither do they appear in the chapter on how public services are to be “transformed” in the recent comprehensive spending review—although “choice” does receive a brief, passing reference. The Confederation of...
Royal colleges call for more staff on maternity wards

Andrew Cole LONDON

The four royal colleges that represent UK doctors and nurses involved in childbirth are calling for urgent increases in the number of staff to ensure safe levels of care on maternity wards.

In their report Safer Childbirth the royal colleges for obstetricians and gynaecologists, anaesthetists, midwives, and paediatrics and child health list the minimum standards required for safe childbirth. These include one to one midwifery care for women in established labour and the presence of a consultant obstetrician on the ward at least 40 hours a week.

At the moment, says the report, only 27% of maternity units in England and Wales have enough midwives to guarantee one to one care, and “barely half” say they have an obstetrician available 40 hours a week.

The Royal College of Obstetricians and Gynaecologists’ own survey also shows that fewer than a third of units claiming to have 40 hours of cover by a consultant obstetrician actually have one present on the labour ward for this length of time.

The college says that the number of obstetricians in the United Kingdom would have to increase from 1600 to 2100—and “ideally” to 2500—to meet all the report’s recommendations. The Royal College of Midwives says that 5000 extra midwives would be needed in England alone.

“If the changes suggested in the document are not met we believe that maternity services will be operating at a suboptimal level,” warned the Royal College of Obstetricians and Gynaecologists.

The report’s authors recognise that their recommendations have “considerable financial implications.” But they also note the huge costs of errors in obstetrics. The cost of litigation payments tripled between 1997 and 2002 and reached £269m (€390m; $550m) in 2001-2.

The report also recommends that:

- An anaesthetist should be available 24 hours a day.
- An appropriately qualified healthcare professional should be present at all births in any setting.
- Obstetricians should conduct ward rounds at least twice a day on Saturdays, Sundays, and bank holidays and do a ward round as late as possible every evening (as well as the commitment for there to be at least 40 hours’ cover).

Safer Childbirth: Minimum Standards for the Organisation and Delivery of Care in Labour is available at www.rcog.org.uk

British Industry, the employers’ organisation, fears that reform on these lines is “on the wane.”

Against all that, the department has—11 months late—finally announced the framework contract that will make it easier for primary care trusts to buy in help with commissioning from 14 private firms—four of them big US health insurers and care managers. The department has also published a new advertisement to seek new suppliers of operations and diagnostics ahead of the introduction from next April of “free choice” for patients of any hospital prepared to treat at NHS prices.

The interim report from Ara Darzi on the “next stage” NHS review called for patient choice to be “embedded within the full spectrum of NHS funded care” (BMJ 2007;335:739, 13 Oct). It also said that new providers should be brought in to help extend GPs’ opening hours. Asked a direct question—if he was committed to those proposals—Mr Johnson said, “Yes, there is a role for the private sector here, and there is a role for it in the rest of the NHS.”

In private, ministerial advisers in both Downing Street and Richmond House insist nothing has changed.

In practice, however, it has. The full Blairite agenda is indeed on the wane—at least for now. And some important distinctions are emerging between how what remains of it will be used in primary and secondary care.

The Blair agenda included not just bringing in the private sector to provide additional capacity to the NHS, but its use as a challenge to existing NHS institutions. In the hospital sector, that is evaporating. The original contracts for independent sector treatment centres were meant to be worth £1bn (€1.4bn; $2bn) a year. But the first wave amounted to only £370m. And of the £550m intended to be in the second wave, only about £150m worth of business, if that, now looks likely to emerge. Less than half of the additional diagnostic contracts are now likely to operate. With one important caveat: far from the independent sector performing “up

“The full Blairite agenda is indeed on the wane—at least for now”

...to 15%” of elective procedures, as John Reid suggested in 2004 when health secretary, the proportion next year looks likely to be only 6-7%, according to the healthcare analysts Laing and Buisson.

For commissioning, the framework contract has been approved.

But in primary care, the government may indeed invite new providers in to challenge GPs over opening hours—although that will happen locally, not through any central contract. Ministers seem more comfortable about using the private sector in areas that have fewer doctors than needed than they do in the hospital sector. This is, perhaps, because the risk of large scale disruption is smaller in primary care than in the secondary sector.

The big unknown in this analysis, however, is patient choice.

Four months in, it is clear that the Brown government is much less ideologically committed than the Blair one to the use of the private sector as a challenge to the NHS as opposed to using it, probably temporarily, to provide additional capacity. The days of big, central, ministerially driven, private sector contracts are dead.

But there remains an outside chance that patient choice could yet make Blair’s vision a reality.
Stress at work is bad for your heart

For the purposes of research, stress at work has two components—high workload and low autonomy. Jobs with both characteristics are unhealthy, especially for people with coronary heart disease. In one Canadian study, chronic job strain doubled the risk of recurrent heart disease in 972 adults (mostly men) returning to work after a heart attack (hazard ratio 2.00, 95% CI 1.08 to 3.72).

The researchers measured participant’s job strain soon after their return to work and again two years later. People who scored highly on both occasions were significantly more likely to die of coronary heart disease, have another non-fatal heart attack, or develop unstable angina than those with less stressful jobs, over a follow-up of six years. The association persisted after adjusting for more than two dozen potentially confounding factors. Job strain seems to be an independent predictor of outcome for middle aged men with a history of heart attack.

While these findings are biologically plausible and consistent with research on other psychosocial stresses, the picture may be different for women, says a linked editorial (p 1549). Contrary to the prevailing view, quality of outpatient care for children in the US looks dismal across the board, and must be having an adverse effect on children’s health.

Medical records can be inaccurate, but these authors analysed only those items most likely to be recorded. So it’s hard to blame poor note keeping for the results. *N Engl J Med* 2007;357:1515-23

Childhood obesity is a chronic disease

Children who lose weight generally put it back on again once their weight loss programme has finished. So researchers from the US developed two weight maintenance programmes, each lasting four months, to try and prolong the effects of dieting. The first taught children and parents how to alter their eating and exercise behaviour for good. The second programme on children’s weight

Originally, all 150 children in the trial were 20-100% overweight for their age and sex. All had at least one overweight parent and had completed five months of family orientated treatment, including more exercise and a calorie controlled diet.

An editorial is pragmatic about the results (p 1695), which were modest. Childhood obesity is a chronic and complex disease, says the author. The solutions are likely to be equally complex. *JAMA* 2007;298:1661-73

Bacterial colonisation of the airway linked to wheeze in young children

We know very little about the root causes of wheeze in young children. But there’s some evidence that bacterial infection could be a trigger. When Danish researchers cultured hypopharyngeal aspirates from a cohort of small babies, they found a clear association between bacterial colonisation and development of wheeze up to the age of 5. The one in five infants colonised with *Streptococcus pneumoniae, Haemophilus influenzae*, or *Moraxella catarrhalis* at 1 month of age were more likely to develop persistent wheeze than uncolonised infants (hazard ratio 2.40, 95% CI 1.45 to 3.99), more likely to be admitted to hospital for wheeze (3.85, 1.90 to 7.79), and more likely to have at least one acute severe exacerbation of wheeze

Children in the US receive inadequate outpatient care

We already know that US adults receive relatively poor quality health care. The nation’s children do no better, according to quality analysis of outpatient data. Over two years, a random sample of 1536 children received less than half the care they needed (46.5%, 95% CI 44.5% to 48.4%). Deficits were most obvious for preventive services (40.7%, 38.1% to 43.4%), screening (37.8%, 34.6% to 41.0%), and diagnostic tests (36.3%, 29.8% to 42.7%). But the researchers found a mismatch between recommended and received care in most areas including acute illnesses such as diarrhoea, chronic illnesses such as asthma, and immunisations.

The researchers used medical records to link what children needed with the care they received for 175 quality indicators covering common childhood illnesses and services. Their findings are shocking, says an editorial (p 1695). Contrary to the prevailing view, quality of outpatient care for children in the US looks dismal across the board, and must be having an adverse effect on children’s health.

Medical records can be inaccurate, but these authors analysed only those items most likely to be recorded. So it’s hard to blame poor note keeping for the results. *JAMA* 2007;298:1661-73
Topiramate helps alcohol dependent adults drink less

Topiramate could be a promising new treatment for alcohol dependence, say the manufacturers. In their latest clinical trial, the drug helped alcohol dependent adults cut down on heavy drinking and increase their chances of staying dry for at least a month. Topiramate worked significantly better than placebo in three different analyses, and its effects were big enough to be clinically relevant, says an independent editorial (p 1691).

The 371 participants took up to 300 mg/day of topiramate or placebo for 14 weeks. They also had a brief weekly behavioural intervention designed to help them stick with their assigned treatment. All were alcohol dependent and wanted to cut down or stop. But they had no other mental illnesses, including other substance misuse disorders. The authors selected their participants carefully, and they admit that their results may not translate well to more typical populations of alcohol dependent adults.

Like all drug treatments, topiramate has side effects. In this trial, people taking the drug had more paraesthesiae, taste perversion, anorexia, and poor concentration than controls. Using a smaller dose may help, say the authors.

N Engl J Med 2007;357:1487-95

A fifth of pregnancies end in abortion

In 2003, an estimated fifth of pregnancies worldwide ended in abortion. In Europe, this proportion was nearer a third, but the overall figure is skewed by the countries of the former Soviet Union (Eastern Europe), where an estimated 45% of pregnancies in 2003 were aborted.

These figures come from a systematic analysis of national and international data on abortion. The last such analysis was done in 1995, and crude comparisons suggest there has been a fall in abortion rates and absolute numbers of abortions since then, particularly in developed countries and in China. Abortion rates have fallen more slowly in the developing world.

In 2003, 48% of all abortions worldwide were unsafe, up from 44% in 1995. Almost all unsafe abortions are carried out in developing countries, particularly Africa, Latin America, and the Caribbean. The resulting toll in deaths, hospital admissions, and infertility is a human rights atrocity, says a linked comment (p 1295). National governments in Africa are making some effort to protect women, but they are hamstrung by the restrictive policies of some donors. For example, aid for family planning from the US comes with a gagging clause that denies women access to information about the option of legal safe abortion.

Lancet 2007;370:1338-45

Maternal mortality is falling too slowly to meet international targets

Women in sub-Saharan Africa are 100 times more likely to die during pregnancy and in childbirth than women in developed countries, according to a report from a joint working group of global agencies, including the World Health Organization, Unicef, and the World Bank. The group estimates that maternal mortality for all developing nations taken together was 9/100,000 live births in 2005, compared with 905/100,000 live births in sub-Saharan Africa. In an analysis covering more than 120 countries, Ireland had the lowest maternal mortality in 2005 (one death per 100,000 live births) and Sierra Leone the highest (2100 deaths per 100,000 live births). The bleak outlook for many African women has remained almost unchanged since 1990, says the report, despite a global drop in maternal mortality rates of about 2.5% a year.

The report’s authors had to use statistical techniques to try and compensate for the usual gaps in their mortality data, and they emphasise that all figures are best guesses. Even so, it seems clear that maternal mortality is not falling fast enough to achieve Millennium Development Goal 5—a 75% reduction between 1990 and 2015.

Lancet 2007;370:1311-9

Pravastatin has enduring effects on risk of heart disease in men

In 1995, researchers from Scotland reported that a daily dose of pravastatin could help prevent coronary heart disease (CHD) in middle aged men with hypercholesterolaemia. The original trial lasted five years, but the researchers continued to track the men for another 10. The benefits persisted. In the 10 years since the trial ended, men originally assigned to pravastatin were still significantly less likely to die from coronary heart disease or have a non-fatal heart attack than men originally assigned to a placebo (8.6% vs 10.3%; hazard ratio 0.82, 95% CI 0.69 to 0.96).

Significantly more men in the pravastatin group took a statin in the period after the trial (38.7% vs 35.2% at 5 years, P<0.001), but this small difference is unlikely to explain the final result, say the authors.

In an analysis covering the 15 years from the start of the original trial, men who took pravastatin for five years were less likely to die from cardiovascular disease (7.6% vs 9.0%; 0.81, 0.68 to 0.96) or anything else (18.7% vs 20.5%; 0.88, 0.79 to 0.99) than controls. The researchers found no evidence of a link between statins and cancer. The suggestion of a link with prostate cancer could have been a chance finding— it disappeared when the researchers corrected their analysis for multiple statistical tests.

WHO CARES FOR BABY?

Divisions between maternal and child health services in the developing world, first reported 20 years ago, left services for newborns overlooked. Today, neonatal mortality is still unacceptably high. Hannah Brown reports on initiatives trying to bridge the gap.

There are few better examples of the wasteful and damaging results of uncoordinated health efforts than in Cambodia. Hundreds of well meaning non-governmental organisations have taken up residence in various parts of the country’s 20 provinces since the brutal civil war of the 1970s and 1980s gave way to peace. But although this massive influx of resources is welcome, the confusing array of organisations, programmes, and priorities has proved unable to meet Cambodia’s needs—particularly when it comes to improving maternal and child health.

Many actors mean many approaches. This has created an incomprehensible patchwork of programmes that have a bewildering range of effects on the population’s health. Unequipped for such a complex managerial challenge, the resource poor government has little capacity to get the non-governmental organisations to coordinate or put in place national programmes that can capitalise on the influx of resources. Instead, disparate independent aid offerings are often the only services many of the country’s 14 million citizens see. So coverage is low for even essential interventions and services are heterogeneous and patchy.

Nutrition is one example. Good nutrition is an essential component of programmes to improve both child and maternal health. But a 2002 nutrition sector review in Cambodia done by Helen Keller International—with the cooperation of 58 United Nations agencies and non-governmental organisations working in the country—revealed a jumbled mess of programmes, approaches, interventions, and coverage rates. Some areas received vitamin A supplements and others iron tablets. Where breastfeeding information was available, it did not necessarily conform to international guidelines and was independent of complementary feeding advice. Some programmes targeted infants whereas others were restricted to women or school age girls. What is more, obvious synergies, such as distribution of bed nets alongside vitamin A capsules to children under 5—both of which need to be distributed every six months—were being ignored, wasting resources and causing unnecessary health problems.

Wider problem

This situation, in which organisations set their own priorities in project areas and shout for their own causes, is symptomatic of a problem that affects many efforts to improve child and maternal health, says Flavia Bustreo, deputy director of the Partnership for Maternal, Newborn, and Child Health—a WHO hosted alliance of over 180 organisations, governments, and universities. The partnership aims to put an end to longstanding divisions that have held back progress in reducing mortality for these groups. Her organisation’s first policy document, *Opportunities for Africa’s Newborns*, published in November last year, emphasised the 4 million newborn deaths that occur every year—more deaths than caused by AIDS and malaria combined. But reducing this figure requires an integration of child and maternal health efforts, a process that means confronting a decades old divide.

Although it is counterintuitive for outsiders,
Health programmes have treated maternal mortality and health of newborn babies separately, but the way donors organise money can ruin that united front,” she says.

Although the main thrust of the conflict is about competition for advocacy and resources, there are ideological aspects too. Some champions on the child side perceive the maternal side as a competition, and some on the maternal side have overly focused on obstetric care when their community approaches could include newborns and children too. It is less about individuals than about ideologies that have become entrenched.

Why did it happen?
This problem is not a new one. It was first articulated over two decades ago when Allan Rosenfield and Deborah Maine posed the question: “Where’s the M in MCH?” in a comment article published in the Lancet. They pointed out that most programmes that fall under the banner of “maternal and child health” did little to reduce maternal mortality and chided obstetricians for subspecialising in areas that prioritise high technology rather than basic public health.

“Thirty years ago there were a large number of child deaths and inspirational people like [former Unicef head] Jim Grant really moved forward for child survival,” explains Dr Lawn. The first primary healthcare revolution was also focused on the child, but success in that area brought the realisation that the mother had been forgotten. “That led maternal health people to start to think of the child as competition,” says Dr Lawn.

But despite reams of articles that have since been written about the problem, numerous examples exist of the divide still being perpetuated. Dr Lawn explains that Unicef’s main focus on the child means that many interventions for newborn babies, which require doing things with women, are not linked into their programming. And United Nations Population Fund programmes are linked to women so do not involve the child. Even the countdown for the child survival conference in 2005—part of a new effort to bring accountability—drew criticism for not including the mother, although it is meant to be adding maternal aspects for its second meeting in 2008.

The result of this division is that the increasing global interest in maternal and child health—sparked by the inclusion of specific targets in the millennium development goals, and maintained by a multitude of new initiatives and high profile

Mother and babies queue outside Kantha Bopha Hospital, Phnom Penh

Immunisation clinic for babies at Chumpou Voan Kampot Province, Cambodia
“Instead of sitting in the corner and shouting for your own item, you could shout together”

campaigns by leading medical journals—is not being efficiently translated into better services for women and children on the ground. For example, explains Dr Bustreo, a lot of resources are going into HIV and AIDS and prevention of mother to child transmission, but these finances are not being used to fill the gaps in basic services for mothers at time of delivery or for newborn babies. “The way that advocacy translates into things on the ground is inefficient. The inefficiency of having separate drives is a key argument for looking at integrated service delivery—what we call the continuum of care,” says Dr Bustreo.

Wider publicity for the problem has resulted in some progress in tackling its roots. The World Health Organization has been working on forging stronger links between its division of reproductive health and department of child and adolescent health. According to Dr Lawn, the situation has improved since it was highlighted in the 2005 world health report, Making Every Mother and Child Count.1 “But”, she notes, “it would be fair to say that the newborn side has suffered because it has dropped between the two departments rather than being part of a strategic look at what would make the biggest difference.”

As Dr Bustreo emphasises, national and international policy and programmes are finally shifting towards a continuum of care that integrates maternal, newborn, and child health. But the progress has been shockingly slow, Dr Lawn explains. WHO’s integrated management of childhood illness strategy—looking at child health from the point of view of the individual rather than specific diseases—had nothing about neonates until last year. And its strategy on making pregnancy safer had nothing on the management of newborn illness.

**Signs of hope**

Two fundamental prerequisites for progress are still largely missing: a scientific evidence base for the right kind of integration and the money to make it happen. That is where Dr Bustreo’s partnership comes in. Bringing together the numerous diverse actors in child and maternal health, the Partnership for Maternal, Newborn, and Child Health pledges to intensify political advocacy and mobilisation of resources in addition to advancing scientific analysis of integration. They have a few good examples to work from. “There have been countries that have already come up with national plans for an integrated approach to maternal, newborn, and child health,” she says. “For example, Pakistan launched one last year, Nigeria did it this year, and Tanzania is moving towards that direction. Next year we hope to produce a series of analysis articles asking whether these are the right way of doing integration in terms of improving coverage of services,” she explains.

Despite the uncertainty over technical details, there is a consensus that the benefits of a joint approach are potentially great. “Instead of sitting in the corner and shouting for your own item, you could shout together,” says Dr Lawn. “I think the amount of investment would go up. And by looking at total deaths in all three groups together you increase the public health burden, meaning you are more likely to have consistent investment. If you change policy every couple of years you won’t make any progress,” she adds.

A test of the new spirit of unity will come in this week’s Women Deliver conference in London, whose delegates include UN leaders and government officials. It will be a litmus test of whether newborn health issues have finally made it into mainstream thinking about maternal health, not least because previous gatherings have roundly failed to consider the issue.4 “The last time safe motherhood had a big public meeting in Sri Lanka they had 10 key messages and not one of the messages mentioned a baby. Not even by saying the mother and the baby would benefit,” recalls Dr Lawn.

She believes that the Women Deliver conference is a unique moment for maternal health advocates to make goals that are not in competition with child health and to give newborn health more than a token mention.

So with all this new momentum towards an integrated approach to maternal, child, and newborn health, is there any more hope that it will happen? Dr Bustreo thinks so, not least because there is no other way to reduce mortality now that the easy fixes have been made. Several years ago it was possible to have a large effect on mortality by boosting immunisation rates, for example. But there remains a bulk of deaths that cannot be reduced without interventions at the time of delivery.

“I am quite optimistic,” she says. “I think the issues are converging, so there is no option but for the different camps to work together.” Even if you look at reducing child mortality alone, most deaths are occurring in young babies and neonates. So even partners that are interested only in children will have to accept that they can’t look just at post-neonatal mortality, she explains. “Biologically and psychologically they are together, and there is no way we can look separately.”

One unforeseen benefit of the long running conflict is that maternal and child health is emerging from the invisibility it has previously endured. But, says Dr Bustreo, to capitalise on this new high profile, the next important step must be clear and strategic mobilisation of resources to help fill the gaps. Maternal, newborn, and child health combined currently make up just 2% of development aid from countries in the Organisation for Economic Co-operation and Development. “We need to move to 4%,” asserts Dr Bustreo.

At the end of last month, Norway’s prime minister, Jens Stoltenberg, pledged $1bn (£500m; €700m) for this cause and his Dutch counterpart, Jan Balkenende, added $125m. But Dr Bustreo sees the contributions as just the initial steps in an ongoing increase in funding, through which she hopes to spur new initiatives that document good practice and measure changes in mortality. “To me, the real bottom line is the services that are provided to mothers and children and coverage of these services. So that is where we have to have our eyes,” she says. “Through the Countdown to 2015 process we have that. There is already some sign of progress but significant gaps remain. But if we keep our eyes on those issues, the differences should be minimised.”

**Competing interests:** None declared.


Hannah Brown is a freelance journalist in Cambridge

Hannah@two-cultures.com
Beware of mentioning psychosocial factors

How doctors describe the many interactions between a person, their illness, and society has little purchase in the crudely dualistic world of popular culture.

Although we are constantly told to “engage with the public,” many doctors and academics avoid the media like the plague. This month, like many doctors before me, I walked into a broadcast studio a man and came out an ass.

The story was acupuncture. A major study had been published showing that acupuncture is more effective for back pain than conventional medical treatment, and as I sometimes write about complementary medicine and research methodology I was invited to discuss the study on BBC Radio 4, where nobody can tell that I look about 14.

The very interesting paper (Archives of Internal Medicine 2007;167:1892-8) had three arms. Results from the “sham” and “real” acupuncture arms were indistinguishable—make of that what you will—but both outperformed conventional medical treatment. The patients in the study, I should mention, were people who had already been failed by conventional medical treatment for an average of eight years.

If you’re a doctor, you can probably imagine what I said. The important background information missing from the news reports didn’t concern the study’s methodological details or anything to do with acupuncture: what was missing was a wider understanding of back pain. Back pain isn’t like tuberculosis or a fracture; it’s one of the leading causes of sick leave and misery, but the simple fact that no clear cause is found itself exacerbates distress and causes conflict.

And just as many of the big risk factors for a niggle turning into chronic, longstanding back pain are personal, psychological, and social—things such as working conditions, depression, job dissatisfaction, unavailability of light duty on return to work, and so on—so are many of the interventions.

Anti-inflammatory drugs are undeniably better than placebo, at the cost of possible side effects, but if you were going to look at the evidence beyond pills, then resting in bed is actively harmful (specific exercises can be too), and trial data show that simple educational interventions such as giving advice to “stay active” can speed recovery, reduce chronic disability, and reduce time off work. In fact, in Australia they even put money into that notion, and a simple public information campaign (“Back pain: don’t take it lying down”) was shown to reduce back pain significantly in the whole target population (Medical Journal of Australia 2001;175:456-7).

I talked about this kind of stuff, although I suspect that the Radio 4 people may have been hoping for some old duffer to say that acupuncture is “poppycock.” The acupuncture study raised important issues, I agreed. It’s very important to think about whether and when doctors should go beyond merely prescribing pills, and we often do. But if we accept, in the case of back pain, that acupuncture may in part be a surrogate psychosocial intervention or theatrical placebo, then it’s a pretty expensive one. Maybe we could consider pushing for other options, less expensive and less fashionable, such as brief education interventions, public health information programmes, perhaps cognitive behavioural therapy, and so on.

This all felt pretty sensible. Much of it was lifted from a bog standard review on back pain in the BMJ (2006;332:1430-4). Working doctors are accustomed to thinking beyond the prescription pad, after all, and it’s a peculiar side effect of the branding of alternative therapists that medicine is portrayed in mainstream media as crudely biomechanical.

So I was a bit surprised three days later to hear this read out, in a very angry voice, in the letters slot of BBC Radio 4’s afternoon news programme PM, to a million people: “I would take issue with your speaker Ben Goldacre, who, if I recall correctly, said that 90% of back problems are psychosomatic disorders. What planet is he on? Whilst I would agree that there are a lot of schmucks out there that want to sit around and skive off work every day (and thereby make the problem even worse) . . . never tell me my backache of 20 years is imaginary. OKAY?”

Now I’m quite prepared to accept that I may not have expressed myself very clearly—obviously I don’t think back pain is “psychosomatic,” and I accuse nobody of malingering—but something more interesting is happening here. The finer distinctions between concepts that doctors use to describe the multiple interactions between a person, their illness, and society have little purchase in the crudely dualistic world of popular culture, and sometimes it can seem that there is a hypersensitivity to anyone even mentioning psychosocial risk factors or interventions.

Perhaps it is a matter of who is permitted to discuss them: patients shop for advice, after all, and you don’t go to a crystal therapist for steroids, any more than you go to a urologist for marital guidance. Perhaps even a biomedical doctor merely raising the question is seen as questioning the legitimacy of symptoms and suffering. Perhaps the problem is magnified because we live in a country with millions of people receiving sickness benefit, where many people perceive at least some claimants simply as rebranded “unemployed” people.

But this is a dark corner, framed only by the crude marketing claims of quacks and their notions of “holism.” In a culture where “psychosocial risk factors” can be heard as “psychosomatic illnesses,” and with a popular media where “psychosomatic” simply means “imaginary” and “malingering” (Psychosomatics 2004;4:5:287-90), these negotiations will never be easy.

Ben Goldacre is a doctor and writer, London ben@badscience.net
RESPONSE Fiona Patterson replies to Parashkev Nachev

Although there is over a century’s literature on selection methodologies, rarely has any process provoked such fierce debate as MTAS. Many facts surrounding specialty selection, and our role within it, are not well understood. This is illustrated by Dr Nachev’s personal view (BMJ 2007;335:615), which contains two fundamental factual inaccuracies. The problems surrounding selection methodology are multi-faceted and multi-dimensional, well beyond disregarding CVs, relying on “white space” questions and poor IT delivery.

For the distress caused, I share the anger of the profession. The inquiry led by Sir John Tooke details the facts about our role in the process [www.mmcinquiry.org.uk]. Here, I highlight critical issues to encourage close scrutiny of facts surrounding principles, methodology, and context underlying MTAS, so that there is learning for the future.

Clearly, in the past, selection practices in medicine have been effective. Although few would deny there is scope for improvement, the CV and interview process has generally worked well. So why change? MTAS was devised alongside the MMC (Modemising Medical Careers) programme introduced by the Department of Health, where the fundamental principles underpinning gateways to progression were changed overnight. Consequently this changed the rules governing selection. Traditionally, medicine has relied on robust indicators of attainment on the CV, such as work experience and Royal College examinations for selection decisions. The MMC pathway relied on the belief that doctors could be selected to specialties without previous experience in that specialty. The selection methodology in MTAS was designed for ST1 (the first level of specialty training), not for thousands of doctors already working in specialties. Under MMC principles and in compliance with Postgraduate Medical Education and Training Board (PMETB) [www.pmetb.org.uk/index.php?id=45], we were advised that work experience and exams could not be scored or used to rank applicants. The introduction of run-through ST1 posts was new.

I have worked on selection methodology in medicine for over 12 years. This work informed the development of selection centres and the introduction of new shortlisting tests in general practice. In 2007, all deaneries worked together through a GP national office, where thousands of doctors are successfully appointed using this process. Since 2002, in partnership with doctors we have developed selection methodology for many secondary care specialties and for graduate entry into medical school. Although commonality exists across all specialties and levels, selection criteria for each are distinct, with evidence supporting different priorities between specialties.

Having completed this work, in 2004, I was invited to meet the MMC team to advise on selection methodology into specialty training. I recommended developing a national test for shortlisting (supported by early evidence from general practice) and validated selection centres with full involvement of the royal colleges and with large scale consultation. Following this meeting, I received no further correspondence from the MMC team and no pilots were put in place. In May 2006 we won an open competition tender organised by the Department of Health. Our work included advising on selection methodology for foundation programmes and the general practice selection process. For specialty selection, the scope of work states; “The number of applicants expected to apply for entry into Specialty Training is approximately 6000 and that applications will be via a single electronic national portal entry system (separate project) the working assumption for the closing date will be 5th January 2007.” At the outset we were asked to advise on selection methodology for ST1. We were not asked to deliver selection methodology for doctors in “transition” via ST2, ST3, ST4, and FISTA (fixed term specialist training appointment) posts, nor academic posts. We believed these arrangements would be delivered via local processes.

The rules and conditions governing MTAS were defined by MMC, based on PMETB principles, and, via the Conference of Postgraduate Medical Deans steering group, they represented all stakeholders. Given the time scale (less than 16 weeks) there was no option but to use materials from existing application forms used (over several years) for entry into specialist training. By contrast, in collaboration with the general practice national office, my team designed the shortlisting test with general practitioners, which has shown to work well. For the future, the general practice model has been identified as best practice.

However, this model cannot be transferred into all specialties. Medicine is a broad discipline and secondary care is significantly different from general practice, requiring bespoke selection methodologies. Different selection ratios for specialties and for locations add complexity; one size cannot fit all. Some believe selection practices in other professions can be readily transferred, but medicine in the UK is unique. Few of those deciding policy understand what a clinician does on a daily basis. An important challenge is to translate the needs of the profession to policy makers.

In the past 12 years, advising on selection methodology in all sectors, I have learnt more from collaborating with the medical profession than from any other. Unlike selection approaches used by some organisations, I applaud the focus on psychometric scrutiny, the need for validatory evidence, and the demand to treat human beings with respect and dignity in the process. The fact is, MTAS was not designed by psychologists. Without a full understanding of the issues, we cannot hope to navigate the future, which looks yet more challenging.

Fiona Patterson is professor of organisational psychology, City University, London, and partner, Work Psychology Partnership
How can we regulate medicines better?

Current European licensing regulations give precedence to the interests of drug companies. Silvio Garattini and Vittorio Bertele’ suggest changes to ensure they meet the needs of patients and doctors.

Despite the undoubted advantages of the establishment of the European Medicines Agency (EMEA), criticisms have been made, mostly about its independence and transparency and the evaluation criteria. The 2004 European Commission law expanded the agency’s remit but did not significantly change the methods of regulation. We offer a few proposals aimed at bettering the work of the agency. These may require important changes in current law regulating the pharmaceutical system.

**Agency’s role**

The agency gives opinions on the quality, safety, and efficacy of new drugs that manufacturers want to market in the European Union. The European Commission uses the agency’s opinions to decide whether to grant a licence. Unlike the US Food and Drug Administration, the European agency is not autonomous but an expression of the national agencies in the European Union, which approve most new products. Evaluation through the centralised procedure has become mandatory for biotechnology products and drugs for HIV, cancer, neurovegetative disorders, diabetes, rare diseases, the autoimmune system, and viral diseases. The European agency is important because central authorisation is increasing and because national agencies cannot justify adopting methods and criteria inconsistent with those of the central agency. When national agencies come to differing decisions, the European agency makes a final decision that is binding in all member states.

**Added value of new drugs**

Too many drugs are approved on the basis of surrogate end points that are not validated predictors of therapeutic end points. Numerous cancer treatments have been authorised on the basis of tumour response that is not correlated with an improvement in quality or duration of life. Manufacturers often object that it would take too long to prove a therapeutic benefit and that patients should not be denied potentially useful drugs, particularly for severe diseases. Assuming this is a valid objection, it is still unclear why drugs in the same class have sometimes been authorised before the therapeutic efficacy of the first drug approved has been shown, as was the case with drugs for pulmonary hypertension and HIV and glitazones for diabetes.

Drug companies sometimes do long term post-marketing trials with hard end points in order to define the new drug’s efficacy in relation to available alternatives. However, doctors and patients need to know relative efficacy at the time of approval if they are to make objective decisions about drugs. Moreover, commitments to new studies are seldom fulfilled, the excuse being that since the product has been approved no doctor wants to randomise and no patient wants to be randomised to placebo or an active comparator. Marketing approval is perceived as the recognition of an added value, even though this is often lacking. New drugs have only to show they are of good quality, effective, and safe, independently of any reference or comparison to drugs already on the market. This results in overuse of trials against placebo. Even when new drugs are compared with existing treatments, the trials often seek to show equivalence or non-inferiority rather than superiority to those already available. Such trials could allow drugs into the market that are less active or safe than those in current clinical use. This is because the non-inferiority limit includes a higher incidence of adverse events. The wider the limits the smaller the sample needed and consequently the higher the chance of missing a difference and concluding for non-inferiority. Sometimes limits are so wide that what is considered non-inferior statistically may be worse clinically.

Non-inferiority trials clearly aim at overlooking differences that might preclude entry to the market rather than highlighting them. This tendency occurred even before non-inferiority trials came into use. In a survey of 383 clinical trials testing superiority in major clinical journals, only 16% of the 70 trials with negative results had sufficient power (80%) to detect a 25% difference and only 36% had power to detect a 50% difference.
Only 3% of trials in schizophrenia had enough power to show a 20% improvement in mental state. It is easier to get to the market by proving a new product is similar to standard available treatments than by failing to show it is superior. It is unethical to experiment on patients with the sole aim of obtaining a marketing authorisation. New drugs should be required to have some added value (greater efficacy or less toxicity) to current treatments or be cheaper. The FDA is apparently changing its mind on the suitability of non-inferiority trials, and we hope the EMEA will follow the same path.

Independent research and development
At present, manufacturers prepare the reports seeking approval for a new drug or a new indication. Companies will clearly tend to maximise the benefits and minimise the risk. The expert opinion that accompanies the research reflects this. Industry sponsored research is more favourable to new products than research done by non-profit organisations, as shown by trials on multiple myeloma, schizophrenia, erythropoietin, and mycophenolate mofetil. Authors’ conclusions in randomised clinical trials were significantly more likely to favour experimental interventions if they declared financial competing interests. Study protocols are developed to favour new drugs: industry prefers to use placebo or no therapy as comparator more often than non-commercial sponsors.

One way to overcome this problem would be to introduce some element of independent research by a non-profit organisation. For example, the regulatory agency could require one phase III trial (usually two pivotal trials are needed) to be planned and carried out by an independent organisation credited by the agency, particularly for drugs that are going to be reimbursed by national health services. Independent research should also aim at developing new drugs for conditions that might not be attractive to commercial companies, such as treatments for rare diseases. In such cases the EMEA could commission a study by academic health structures, which would be audited by independent research organisations.

At present independent research occurs only after approval. The Italian regulatory agency, for example, requires drug companies to contribute 5% of their promotional expenses each year, which it uses to fund research. Funds are allocated according to a priority score established by discussion groups including international reviewers, and the projects range from trials to optimise the use of orphan drugs to comparison of drugs licensed for the same indications and observational studies dealing with pharmacovigilance. This initiative could be extended to the premarketing phase and coordinated at the European level.

Disclosure of drug information
Another concern with the European agency is transparency. Unlike the FDA, the EMEA keeps almost all its information secret. Although disclosure of documentation concerning production and drug technology could help competitors, there is no reason to hide data on toxicology and clinical evaluation. This information is essential to understand why a new drug has been approved or a new indication granted. Although the agency releases a European public assessment report, this is a generic document written under the supervision of the company concerned.

Other information that should be made transparent includes the size of the majority that approved a given drug, the reasons of the minority for opposing approval, conflicts of interest, and post-marketing commitments and their fulfilment. When drugs are approved “under exceptional circumstances” on evidence from surrogate end points, the manufacturers usually have to commit to do further research. However, in many cases they ask for extra time or do not meet the commitment at all. A recent FDA survey indicated that only 926 (34%) of the 2701 post-marketing commitments were honoured. No equivalent information is available from the EMEA.

The documents made available by the EMEA do not disclose the issues raised in the evaluation of marketing application, do not reflect the debate in the Committee for Human Medicinal Products (the body that gives the European Commission opinions on new drugs), and therefore do not allow an independent evaluation by people outside the agency.

Implementing active pharmacovigilance
One important weakness of the European drug system is pharmacovigilance. It relies on national activities, which in several cases are limited to spontaneous reports from patients and doctors. Clearly, collection of data about drug toxicity could be strengthened by establishing a network covering all 25 European countries, coordinated by the EMEA. The establishment of EudraVigilance in 2001 was intended to facilitate this, but the system is still far from being fully implemented. EudraVigilance is a data processing network aimed at facilitating the electronic exchange of suspected adverse reaction reports between the European agency, national authorities, marketing authorisation holders, and sponsors of clinical trials in Europe.

Proposed models for future pharmacovigilance include actively looking for toxicity rather than relying on spontaneous reporting. The proactive approach requires research projects to investigate severe adverse reactions such as gastrointestinal bleeding, prolonged Q-T interval, rhabdomyolysis, hepatitis, renal insufficiency, and dependence. Programmes should focus on the signs of toxicity for specific drugs or classes of drugs. In addition, companies should be obliged to present meta-analyses of both beneficial and adverse events in the regular safety update reports they submit to the EMEA.

The EMEA should establish a new pharmacovigilance committee, separate from the Committee for Human Medicinal Products. Decisions to restrict the use of a drug or to withdraw it from the market must be taken by an independent group devoted to pharmacovigilance. The medicinal products committee already has a heavy workload and may be unconsciously resistant to withdrawing a drug that it has approved.
Removal of bias
The current power of drug companies conflicts with the right of patients to receive drugs that have been studied enough to guarantee that they are more effective and safer than current treatments. Major international clinical journals have helped to make the influence of drug companies clearer by requiring that authors declare they know the content of the articles, avoiding ghost writing by people paid by drug companies, state conflicts of interest, and register protocols of clinical trials to highlight changes made after the trial has begun.

The regulatory system is also subject to bias and suspicion. For instance, it is anomalous that the EMEA is part of the EU general directorate of enterprise and industry rather than the general directorate of health and consumers. It is also strange that the general directorate of enterprise regulates compliance to good clinical practice by non-profit organisations.

It is equally surprising that about 70% of the agency’s budget comes from fees paid by the applicants.

Some of our suggestions will make the approval of new drugs and new indications more difficult and prolong the time needed for their introduction into the market. We may therefore need to be more flexible to encourage industrial research. One possibility would be to prolong product patents in exchange for better, safer, more trustworthy, and more affordable innovation. We believe the changes will not only be important for patients but will help stimulate innovative research by drug companies.

Competing interests: SG and VB are director and researcher at the Mario Negri Institute for Pharmacological Research, an independent non-profit research institute devoted to patients’ and national health services’ interests. The Institute could take advantage of the proposal to increase the role of independent research in setting up the documentation supporting marketing authorization applications for new drugs or new clinical indications.

SUMMARY POINTS

- Licensing of new drugs in Europe is increasingly controlled by the European Medicines Agency.
- Current regulations need changing to ensure all new drugs add value.
- Assessment should include a phase III study by an independent organisation.
- Toxicological and clinical information should be made public.
- The European pharmacovigilance system should be implemented.

The Future of the Medical Profession
The current power of drug companies conflicts with the right of patients to receive drugs that have been studied enough to guarantee that they are more effective and safer than current treatments. Major international clinical journals have helped to make the influence of drug companies clearer by requiring that authors declare they know the content of the articles, avoiding ghost writing by people paid by drug companies, state conflicts of interest, and register protocols of clinical trials to highlight changes made after the trial has begun.

The regulatory system is also subject to bias and suspicion. For instance, it is anomalous that the EMEA is part of the EU general directorate of enterprise and industry rather than the general directorate of health and consumers. It is also strange that the general directorate of enterprise regulates compliance to good clinical practice by non-profit organisations.

It is equally surprising that about 70% of the agency’s budget comes from fees paid by the applicants.

Some of our suggestions will make the approval of new drugs and new indications more difficult and prolong the time needed for their introduction into the market. We may therefore need to be more flexible to encourage industrial research. One possibility would be to prolong product patents in exchange for better, safer, more trustworthy, and more affordable innovation. We believe the changes will not only be important for patients but will help stimulate innovative research by drug companies.

Competing interests: SG and VB are director and researcher at the Mario Negri Institute for Pharmacological Research, an independent non-profit research institute devoted to patients’ and national health services’ interests. The Institute could take advantage of the proposal to increase the role of independent research in setting up the documentation supporting marketing authorization applications for new drugs or new clinical indications.

SUMMARY POINTS

- Licensing of new drugs in Europe is increasingly controlled by the European Medicines Agency.
- Current regulations need changing to ensure all new drugs add value.
- Assessment should include a phase III study by an independent organisation.
- Toxicological and clinical information should be made public.
- The European pharmacovigilance system should be implemented.

CORRECTIONS AND CLARIFICATIONS

The future of the medical profession
The printers’ devils were mixing up commas and apostrophes in this editorial by Roger Jones (BMJ 2007;335:53, 14 Jul). The final sentence of the third paragraph should read: “The role of doctors, careers in medicine, and the future of a publicly funded health service...”.

Cardioversion of atrial fibrillation
An error has belatedly emerged in relation to this article by Gregory Y H Lip and colleagues in the ABC of Atrial Fibrillation series (BMJ 1996;312:112-5). In the graph with the caption that reads “Patients in sinus rhythm at time of cardioversion and at one month and six months of follow up according to duration of atrial fibrillation” the labels for the blue and red bars are the wrong way round. The blue bars represent “less than 3 months” of atrial fibrillation not “more than 12 months” and the red bars represent “more than 12 months” of atrial fibrillation not “less than 3 months”.

Minerva
Readers may have been seriously confused when Minerva wrote in one of her items that “exclusive breastfeeding is the best way for HIV positive mothers to feed their babies” (BMJ 2007;335:216, 28 Jul). It depends, of course, on the circumstances. Current international consensus guidance from the World Health Organization states that exclusive breastfeeding is recommended for HIV infected women for the first six months of life “unless replacement feeding is acceptable, feasible, affordable, sustainable and safe for them and their infants before that time,” in which case, avoidance of all breastfeeding is recommended (www.who.int/child-adolescent-health/New_Publications/NUTRITION/consensus_statement.pdf).

Management of asthma in children
In this Clinical Review by J Townsend and colleagues (BMJ 2007;335:253-7, 4 Aug), a mix-up in terminology during editing was not picked up at proof stage, resulting in some confusion for readers. In the section headed “Long acting 2 agonists” five of the six references to “short acting 2 agonists” (all but in the final sentence) should have read “long acting 2 agonists.” In addition, the second and third sentences in the Summary Points box should refer to “long acting 2 agonists not antagonists.”

Diagnostic accuracy and clinical utility of a simplified low cost method of counting CD4 cells with flow cytometry in Malawi: diagnostic accuracy study
In this paper by Calman A MacLennan and colleagues, we gave the wrong unit for the CD4 cell count (BMJ 2007;335:190-4, 28 Jul). Throughout the article we gave the unit as x10^9/l, whereas the correct unit is cells/μl (cells per microlitre). This error has been corrected in the online pdf and html versions of the article.
Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies

Many questions in medical research are investigated in observational studies.1 Much of the research into the cause of diseases relies on cohort, case-control, or cross-sectional studies. Observational studies also have a role in research into the benefits and harms of medical interventions.2 Randomised trials cannot answer all important questions about a given intervention. For example, observational studies are more suitable to detect rare or late adverse effects of treatments, and are more likely to provide an indication of what is achieved in daily medical practice.3

Research should be reported transparently so that readers can follow what was planned, what was done, what was found, and what conclusions were drawn. The credibility of research depends on a critical assessment by others of the strengths and weaknesses in study design, conduct, and analysis. Transparent reporting is also needed to judge whether and how results can be included in systematic reviews.4 5 However, in published observational research important information is often missing or unclear. An analysis of epidemiological studies published in general medical and specialist journals found that the rationale behind the choice of potential confounding variables was often not reported.6 Only few reports of case-control studies in psychiatry explained the methods used to identify cases and controls.7 In a survey of longitudinal studies in stroke research, 17 of 49 articles (35%) did not specify the eligibility criteria.8 Others have argued that without sufficient clarity of reporting, the benefits of research might be achieved more slowly,9 and that there is a need for guidance on reporting observational studies.10 11

Recommendations on the reporting of research can improve reporting quality. The consolidated standards of reporting trials (CONSORT) statement was developed in 1996 and revised five years later.12 Many medical journals supported this initiative,13 which has helped to improve the quality of reports of randomised trials.14 15 Similar initiatives have followed for other research areas—for example, for the reporting of meta-analyses of randomised trials16 or diagnostic studies.17

We established a network of methodologists, researchers, and journal editors to develop recommendations for reporting observational research: the strengthening the reporting of observational studies in epidemiology (STROBE) statement.

Aims and use of STROBE statement

The STROBE statement is a checklist of items that should be addressed in articles reporting on the three main study designs of analytical epidemiology: cohort, case-control, and cross-sectional studies. The intention is solely to provide guidance on how to report observational research well: these recommendations are not prescriptions for designing or conducting studies. Also, while clarity of reporting is a prerequisite to evaluation, the checklist is not an instrument to evaluate the quality of observational research.

Here we present the STROBE statement and explain how it was developed. In a detailed companion article, the explanation and elaboration article,18-20 we justify the inclusion of the different checklist items and give methodological background and published examples of what we consider transparent reporting. We strongly recommend using the STROBE checklist in conjunction with the explanatory article, which is available freely on the websites of the publishing journals.18-20

Development of STROBE statement

We established the STROBE Initiative in 2004, obtained funding for a workshop and set up a website (www.strobe-statement.org). We searched textbooks, bibliographic databases, reference lists, and personal files for relevant material, including previous recommendations, empirical studies of reporting, and articles describing relevant methodological research. Because observational research makes use of many different study designs, we felt that the scope of STROBE had to be clearly defined early on. We decided to focus on the three study designs that are used most widely in analytical observational research: cohort, case-control, and cross-sectional studies.

We organised a two-day workshop in Bristol in September 2004. Twenty-three people attended this meeting, including editorial staff from Annals of Internal Medicine, BMJ, Bulletin of the World Health Organization, International Journal of Epidemiology, JAMA, Preventive Medicine, and the Lancet as well as epidemiologists, methodologists, statisticians, and practitioners from Europe and North America. Written contributions were sought from 10 other people who declared an interest in contributing to STROBE but could not attend.

Three working groups identified items that were
deemed to be important to include in checklists for each type of study. A provisional list of items prepared in advance (available from our website) was used to facilitate discussions. The three draft checklists were then discussed by all participants and, where possible, items were revised to make them applicable to all three study designs. In a final plenary session, the group decided on the strategy for finalising and disseminating the STROBE statement.

After the workshop we drafted a combined checklist including all three designs and made it available on our website. We invited participants and additional scientists

---

**STROBE statement—checklist of items that should be included in reports of observational studies**

<table>
<thead>
<tr>
<th>Item No</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td></td>
</tr>
</tbody>
</table>
| 1 | (a) Indicate the study’s design with a commonly used term in the title or the abstract  
(b) Provide in the abstract an informative and balanced summary of what was done and what was found |
| **Introduction** | |
| **Background/rationale** | |
| 2 | Explain the scientific background and rationale for the investigation being reported |
| **Objectives** | |
| 3 | State specific objectives, including any prespecified hypotheses |
| **Methods** | |
| **Study design** | |
| 4 | Present key elements of study design early in the paper |
| **Setting** | |
| 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection |
| **Participants** | |
| 6 | (a) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  
*Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  
*Cross sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants  
(b) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  
*Case-control study*—For matched studies, give matching criteria and the number of controls per case |
| **Variables** | |
| 7 | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group |
| **Data sources/ measurement** | |
| 8* | Describe any efforts to address potential sources of bias |
| **Bias** | |
| 9 | Explain how the study size was arrived at |
| **Study size** | |
| 10 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why |
| **Quantitative variables** | |
| 11 | Describe all statistical methods, including those used to control for confounding |
| **Statistical methods** | |
| 12 | (a) Describe each statistical method used to examine subgroups and interactions  
(b) Explain how missing data were addressed  
(c) Consider use of a flow diagram |
| **Results** | |
| **Participants** | |
| 13* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed  
(b) Give reasons for non-participation at each stage  
(c) Consider use of a flow diagram |
| **Descriptive data** | |
| 14* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders  
(b) Indicate number of participants with missing data for each variable of interest  
(c) Consider use of a flow diagram |
| **Outcome data** | |
| 15* | *Cohort study*—Report numbers of outcome events or summary measures over time  
*Case-control study*—Report numbers in each exposure category, or summary measures of exposure  
*Cross sectional study*—Report numbers of outcome events or summary measures |
| **Main results** | |
| 16 | (a) Report the numbers of individuals at each stage of the study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed  
(b) Give reasons for non-participation at each stage  
(c) Consider use of a flow diagram |
| **Other analyses** | |
| 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses |
| **Discussion** | |
| **Key results** | |
| 18 | Summarise key results with reference to study objectives |
| **Limitations** | |
| 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias |
| **Interpretation** | |
| 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence |
| **Generalisability** | |
| 21 | Discuss the generalisability (external validity) of the study results |
| **Other information** | |
| **Funding** | |
| 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based |

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross sectional studies.

The STROBE checklist is best used in conjunction with the explanation and elaboration article. This article and separate versions of the checklist for cohort, case-control, and cross sectional studies are available at www.strobe-statement.org.
and editors to comment on this draft checklist. We subsequently published three revisions.

**STROBE components**
The STROBE statement is a checklist of 22 items that we consider essential for good reporting of observational studies (table). These items relate to the article's title and abstract (item 1), the introduction (items 2 and 3), methods (items 4-12), results (items 13-17), discussion sections (items 18-21), and other information (item 22 on funding). Eighteen items are common to all three designs, while four (items 6, 12, 14, and 15) are design specific, with different versions for all or part of the item. For some items (indicated by asterisks), information should be given separately for cases and controls in case-control studies, or exposed and unexposed groups in cohort and cross sectional studies. Although the table is a single checklist, the STROBE website provides separate checklists for each of the three study designs.

**Implications and limitations**
The STROBE statement was developed to assist authors when writing up analytical observational studies, to support editors and reviewers when considering such articles for publication, and to help readers when critically appraising published articles.

Observational studies serve a wide range of purposes, on a continuum from the discovery of new findings to the confirmation or refutation of previous findings. Some studies are essentially exploratory and raise interesting hypotheses. Others pursue clearly defined hypotheses in available data. In yet another type of studies, the collection of new data is planned carefully on the basis of an existing hypothesis. We believe the present checklist can be useful for all these studies, since the readers always need to know what was planned (and what was not), what was done, what was found, and what the results mean.

We acknowledge that STROBE is currently limited to three main observational study designs. We would welcome extensions that adapt the checklist to other designs—for example, case crossover studies or ecological studies—and also to specific topics. Four extensions are now available for the CONSORT statement. A first extension to STROBE is under way for gene-disease association studies: the STROKE extension to Genetic Association studies (STREGA) initiative. We ask those who aim to develop extensions of the STROBE statement to contact the coordinating group first to avoid duplication of effort.

The STROBE statement should not be interpreted as an attempt to prescribe the reporting of observational research in a rigid format. The checklist items should be addressed in sufficient detail and with clarity somewhere in an article, but the order and format for presenting information depends on author preferences, journal style, and the traditions of the research field. For instance, we discuss the reporting of results under a number of separate items, while recognising that authors might address several items within a single section of text or in a table. Also, item 22, on the source of funding and the role of funders, could be addressed in an appendix or in the methods section of the article. We do not aim at standardising reporting. Authors of randomised clinical trials were asked by an editor of a specialist medical journal to “CONSORT” their manuscripts on submission. We believe that manuscripts should not be “STROBEd,” in the sense of regulating style or terminology. We encourage authors to use narrative elements, including the description of illustrative cases, to complement the essential information about their study, and to make their articles an interesting read.

We emphasise that the STROBE statement was not developed as a tool for assessing the quality of published observational research. Such instruments have been developed by other groups and were the subject of a recent systematic review. In the explanatory article we used several examples of good reporting from studies whose results were not confirmed in further research; the important feature was the good reporting, not the quality of the research. However, if authors and journals adopt the STROBE statement, issues such as confounding, bias, and generalisability could become more transparent, which might help temper the over-enthusiastic reporting of new findings in the scientific community and popular media, and improve the methods of studies in the long term. Better reporting may also help to have more informed decisions about when new studies are needed, and what they should address.

We did not undertake a comprehensive systematic review for each of the checklist items and sub-items, or do our own research to fill gaps in the evidence base. Furthermore, although no one was excluded from the process, the composition of the group of contributors was influenced by existing networks and was not representative in terms of geography and probably was not representative in terms of research interests and disciplines. We stress that STROBE and other recommendations on the reporting of research should be seen as evolving documents that require continual assessment, refinement, and, if necessary, change. We welcome suggestions for the further dissemination of STROBE—for example, by republishing this article in specialist journals and in journals published in other languages. Groups or individuals who intend to translate the checklist to other languages should consult the coordinating group beforehand. We will revise the checklist in the future, taking into account comments, criticism, new evidence, and experience from its use. We invite readers to submit their comments through the STROBE website.
Amateur boxing and risk of chronic traumatic brain injury: systematic review of observational studies

Mike Loosemore, lead sports physician,1 Charles H Knowles, clinical senior lecturer and honorary consultant surgeon,2 Greg P Whyte, professor of sport and exercise science3

ABSTRACT
Objective To evaluate the risk of chronic traumatic brain injury from amateur boxing.
Setting Secondary research performed by combination of sport physicians and clinical academics.
Design, data sources, and methods Systematic review of observational studies in which chronic traumatic brain injury was defined as any abnormality on clinical neurological examination, psychometric testing, neuroimaging studies, and electroencephalography. Studies were identified through database (1950 to date) and bibliographic searches without language restrictions. Two reviewers extracted study characteristics, quality, and data, with adherence to a protocol developed from a widely recommended method for systematic review of observational studies (MOOSE).
Results 36 papers had relevant extractable data (from a detailed evaluation of 93 studies of 943 identified from the initial search). Quality of evidence was generally poor. The best quality studies were those with a cohort design and those that used psychometric tests. These yielded the most negative results: only four of 17 (24%) better quality studies found any indication of chronic traumatic brain injury in a minority of boxers studied.
Conclusion There is no strong evidence to associate chronic traumatic brain injury with amateur boxing.

INTRODUCTION
In light of evidence of acute and chronic injuries associated with boxing, the British Medical Association (BMA) has passed a series of resolutions at its annual representative meetings calling for boxing to be made illegal.12 The latest report from the BMA Board of Science Working Party on Boxing (now disbanded), published as a briefing paper, continues to campaign for a complete ban on boxing (amateur and professional), mainly because of the purported risk of cumulative brain injury (chronic traumatic brain injury).2 Severe acute injuries in boxing (including those resulting in fatality), however, are relatively rare compared with other sports, even when professional and amateur boxing are grouped together.245

A series of important changes in rules and equipment aimed at improving the safety of boxing have been gradually introduced by boxing authorities since the early 20th century.67 Whether such changes have improved safety remains contentious. The box shows changes relevant in amateur boxing. On the basis of published data available at the time, the BMA 2001 report acknowledged that the evidence for chronic traumatic brain injury in amateur boxing was “far less clear cut” than in professional boxing.1 There have been several publications since this BMA report that have continued to examine the link between boxing, including the amateur sport, and chronic traumatic brain injury.12

We carried out a systematic review to determine whether amateur boxing leads to chronic traumatic brain injury. We did not consider professional boxing, the incidence of acute injuries, or the moral or legal arguments regarding the sport. A problem with

Changes in rules and equipment in amateur boxing

1906— Requirement for a medical examination before the contest
1947— Referees allowed in the ring
1950— Boxing medicals and medical cards introduced with imposition of mandatory suspensions for certain injuries. Doctor must be present at ring side
1962— Establishment of the Medical Commission of the Amateur Boxing Association (ABA)
1964— Introduction of the “standing 8” count
1972— First publication of Medical Aspects of Amateur Boxing
1984— Head guards introduced for the Los Angeles Olympics
1992— Computerised scoring system introduced at the Barcelona Olympics
1996— Structure of the bouts changed from three rounds of three minutes to four rounds of two minutes
2000— Introduction of the “outclassed” rule (Sydney Olympics): bout stopped automatically if one boxer leads the other by 20 points in any but the final round
2002— Ringside physician regulations changed to include suitably trained paramedics present when the doctor is not trained or equipped for resuscitation (Medical Aspects of Boxing, 2002)
MeSH terms were still manageable number of retrieved titles. In addition, based on cases described to him by promoters.8 This was the lowest thresholds for surrogate markers of chronic traumatic brain injury. Historically, the first description of a link between boxing and cerebral dysfunction was that of Martland, who described “punch drunk” (although this was actually based on cases described to him by promoters).8 This extreme form of injury, perhaps partially encompassed by the current term “chronic traumatic brain encephalopathy”90 is rare. We clearly needed to consider much more subtle indicators that may be surrogate markers of chronic traumatic brain injury. In the absence of any ideal standards for this, we took the lowest thresholds—that is, any consistent change in the results of neurological examination, brain imaging, psychometric testing, electroencephalography, including a few other relevant studies for completion. We included studies of amateur boxers (including military and police), with the intervention (exposure) being participation in the sport and from which we could extract data.

METHODS

Though the quality and heterogeneity of available data meant that we could not undertake a meta-analysis, we have adhered as far as possible to the QUOROM statement for systematic reviews.10 Because all included studies were observational in design, we also adhered to a protocol developed from a widely recommended method for systematic review/meta-analysis of observational studies (MOOSE).11

Search strategy

Two authors (ML (initial search) and CK (final arbitrator in selection)) carried out a comprehensive search of the literature using Medline and Premedline 1950 to December 2006, Embase, Evidence Based Medicine (EBM) reviews (including the Cochrane database of systematic reviews and the Cochrane central register of controlled trials), and the SPORTDiscus database. The only search term used was “boxing” because of the still manageable number of retrieved titles. In addition, MeSH terms were “(“Boxing/adverse effects” OR “Boxing/injuries” OR “Boxing/mortality” OR “Boxing/physiology” OR “Boxing/psychology” OR “Boxing/psychology” OR “Boxing/mortality”)” — that is, any consistent change in the results of neurological examination, brain imaging, psychometric testing, electroencephalography, including a few other relevant studies for completion. We included studies of amateur boxers (including military and police), with the intervention (exposure) being participation in the sport and from which we could extract data.

Selection

The figure outlines the study selection process. We included all studies from which we could extract data. In addition, we also adhered to a protocol developed from a widely recommended method for systematic review/meta-analysis of observational studies (MOOSE).11

Data extraction and synthesis

ML extracted data, which were checked by CK. As far as possible, we obtained numerical data, though outcome measures were largely categorical in case series (proportions of participants with positive findings) or expressed as group differences in controlled studies. No quantitative data synthesis was performed. Exposure times were collected and expressed as median or mean number of bouts. Accepting that exposure to injury relates to quality, quantity, and length of bouts, we also included the type (level) of boxing where this was recorded. In cohort studies, exposure was presented as length of follow-up in years as well as number of bouts during the study period.
RESULTS

Literature identification, study design, and quality

We identified 943 citations on the basis of initial search terms, of which we selected 36 articles for the systematic review (from 93 retrieved for detailed evaluation, see figure). Most exclusions were because there were no original data (n=26) or data were on acute injuries only (n=12) or were duplicate data (n=10). We excluded four studies because we could not separate data from amateur and professional boxers, including one often cited paper. Five foreign case series from 1959-68 were irretrievable (although these would not necessarily have been included). Of the 36 selected, 16 evaluated findings from psychometric tests, 11 from brain imaging, 14 from electroencephalography, and 12 from neurological examination, with several including more than one outcome measure (63 methods in all). We included four cohort studies, four controlled before and after studies, and 11 case-control studies, with the remainder (n=17) being case series (six of which were prospective—that is, before and after studies in which the cases acted as their own controls).

Overall quality was poor (median score 2/6, range 0-6) (table 1). Table 2 shows the characteristics of studies and table 3 the main results. All are tabulated by quality followed by time since publication—that is, most recent first (in some instances, some studies had different designs or quality and numbers of participants for different outcome measures so we then included the best quality in table 1). We have summarised results for the main outcome measures below in order of general quality.

### Table 1 | Quality of included studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Prospective</th>
<th>Groups comparable on confounding factors</th>
<th>Blinded outcome</th>
<th>Long enough follow-up</th>
<th>Exposure response measured</th>
<th>Appropriate statistics</th>
<th>Overall quality (max 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porter w1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>Porter w2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>Zetterberg w3</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Moriarty w4</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Stewart w5</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Butler w6</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Haglund w7</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Brayne w8</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Master w9</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Heilbronner w10</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Kemp w11</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Murelius w12</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Haglund w13</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Haglund w14</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Levin w15</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Brooks w16</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Thomassen w17</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Holzgrafe w18</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Jedlinski w19</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Rodriguez w20</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>McLatchie w21</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Ross w22</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>1</td>
</tr>
<tr>
<td>Jordan w23</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Legwold w24</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Jordan w25</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Casson w26</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Kastew 27</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Corselli 28</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Beaussam 29</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Szymusi 30</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Moriyasu 31</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Nesarajah 32</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Beaussam 33</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Blonstein 34</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Pampus 35</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Blonstein 36</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
</tbody>
</table>

NA=not applicable.
Table 2 | Characteristics of included studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study design</th>
<th>Type of boxing, duration of follow-up, No of bouts (mean unless stated)</th>
<th>Outcome measures used</th>
<th>No of cases</th>
<th>No of controls</th>
<th>Selection of cases and controls</th>
<th>Methods used to control for confounding (when applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porter[1]</td>
<td>Cohort</td>
<td>Club, 9 years, 80</td>
<td>Psychometric</td>
<td>20</td>
<td>20</td>
<td>Random</td>
<td>Age, geographical, sex, socioeconomic status</td>
</tr>
<tr>
<td>Porter[2]</td>
<td>Cohort</td>
<td>Club, 2 years, 50</td>
<td>Psychometric</td>
<td>20</td>
<td>20</td>
<td>Random</td>
<td>Age, geographical, sex, socioeconomic status</td>
</tr>
<tr>
<td>Zetterberg[3]</td>
<td>Controlled before-after</td>
<td>NS, 3 months, 1</td>
<td>Cerebrospinal fluid biochemistry</td>
<td>14</td>
<td>10</td>
<td>NS</td>
<td>Age, sex</td>
</tr>
<tr>
<td>Moriarty[4]</td>
<td>Controlled before-after</td>
<td>Club, 7 days, 1</td>
<td>Psychometric</td>
<td>85</td>
<td>30</td>
<td>1 tournament</td>
<td>Age, sex, education</td>
</tr>
<tr>
<td>Stewart[5]</td>
<td>Case series before-after</td>
<td>Club, 2 years, 0-&gt;11</td>
<td>Psychometric, EEG, Brain evoked potentials</td>
<td>369</td>
<td>0</td>
<td>Invitation</td>
<td>Before and after (that is, act as own controls)</td>
</tr>
<tr>
<td>Butler[6]</td>
<td>Cohort</td>
<td>Club, 2 years, 4</td>
<td>Psychometric</td>
<td>86</td>
<td>78</td>
<td>Invitation</td>
<td>Age, sex</td>
</tr>
<tr>
<td>Haglund[7]</td>
<td>Case-control</td>
<td>Club, NA, 28 (estimate)</td>
<td>CT, MRI</td>
<td>47</td>
<td>50</td>
<td>Random</td>
<td>Age</td>
</tr>
<tr>
<td>Brayne[8]</td>
<td>Controlled before-after</td>
<td>NS, 2 hours, 1</td>
<td>Creatine kinase BB</td>
<td>16</td>
<td>16</td>
<td>Invitation</td>
<td>Age, sex</td>
</tr>
<tr>
<td>Master[9]</td>
<td>Controlled before-after</td>
<td>Club, 1 bout, 1</td>
<td>Psychometric</td>
<td>38</td>
<td>28</td>
<td>Random</td>
<td>Age, sex, education, weight</td>
</tr>
<tr>
<td>Heilbronner[10]</td>
<td>Case series before-after</td>
<td>NS, 5 min, 1 bout</td>
<td>Psychometric</td>
<td>28</td>
<td>0</td>
<td>Invitation</td>
<td>Before and after (that is, act as own controls)</td>
</tr>
<tr>
<td>Kemp[11]</td>
<td>Case-control</td>
<td>Military, NA, 40 (median)</td>
<td>Psychometric, SPECT</td>
<td>34</td>
<td>34</td>
<td>Invitation</td>
<td>Age, sex</td>
</tr>
<tr>
<td>Murelius[12]</td>
<td>Case-control</td>
<td>Club, NA, 28 (estimate)</td>
<td>Psychometric</td>
<td>50</td>
<td>50</td>
<td>Random</td>
<td>Age, education, sex</td>
</tr>
<tr>
<td>Haglund[13]</td>
<td>Case-control</td>
<td>Club, NA, 28 (estimate)</td>
<td>Neurological, Psychometric, Pt MAO</td>
<td>50</td>
<td>50</td>
<td>Random</td>
<td>Age, education, sex</td>
</tr>
<tr>
<td>Haglund[14]</td>
<td>Case-control</td>
<td>Club, NA, 28 (estimate)</td>
<td>EEG, brain evoked potentials</td>
<td>50</td>
<td>50</td>
<td>Random</td>
<td>Age, education, sex</td>
</tr>
<tr>
<td>Levin[15]</td>
<td>Cohort</td>
<td>Club, 0.5 years, 2-7</td>
<td>Psychometric, MRI</td>
<td>2</td>
<td>13</td>
<td>Invitation</td>
<td>Age, sex, education, socioeconomic status</td>
</tr>
<tr>
<td>Brooks[16]</td>
<td>Case-control</td>
<td>Club, NA, 26</td>
<td>Psychometric</td>
<td>29</td>
<td>19</td>
<td>Invitation</td>
<td>Age, sex, education, ethnicity</td>
</tr>
<tr>
<td>Thomasse[17]</td>
<td>Case-control</td>
<td>Champions, NA, 76 (median)</td>
<td>Psychometric, EEG, neurological</td>
<td>53</td>
<td>53</td>
<td>Invitation</td>
<td>Age, sex, education, occupation,</td>
</tr>
<tr>
<td>Holzgraf[18]</td>
<td>Case-series before-after</td>
<td>Top class, 2 months, 1</td>
<td>MRI</td>
<td>13</td>
<td>0</td>
<td>Invitation</td>
<td>Before and after (that is, act as own controls)</td>
</tr>
<tr>
<td>Jedlinski[19]</td>
<td>Case-control</td>
<td>&gt;100 bouts, NA, 153</td>
<td>Psychometric, EEG, neurological</td>
<td>60</td>
<td>30</td>
<td>Invitation</td>
<td>Reported as &quot;characterologic&quot;</td>
</tr>
<tr>
<td>Rodriguez[20]</td>
<td>Case series</td>
<td>NS, NA, NS</td>
<td>Regional cerebral blood flow</td>
<td>7</td>
<td>0</td>
<td>invitation</td>
<td>NA</td>
</tr>
<tr>
<td>McLatchie[21]</td>
<td>Case series</td>
<td>Club, NA, 4-200</td>
<td>Psychometric, CT, EEG, neurological</td>
<td>20</td>
<td>0</td>
<td>Invitation</td>
<td>NA</td>
</tr>
<tr>
<td>Ross[22]</td>
<td>Case series</td>
<td>Club, NA, 13-150</td>
<td>CT, EEG, neurological</td>
<td>13</td>
<td>0</td>
<td>Invitation</td>
<td>NA</td>
</tr>
<tr>
<td>Jordan[23]</td>
<td>Case series</td>
<td>NS, NA, NS</td>
<td>MRI, CT</td>
<td>4</td>
<td>0</td>
<td>Selected*</td>
<td>NA</td>
</tr>
<tr>
<td>Legwol[24]</td>
<td>Case series</td>
<td>Military, NA, 4</td>
<td>Neurological</td>
<td>7000</td>
<td>0</td>
<td>Compulsory</td>
<td>NA</td>
</tr>
<tr>
<td>Jordan[25]</td>
<td>Case series</td>
<td>Knockouts, NS, 13</td>
<td>MRI, neurological</td>
<td>9</td>
<td>0</td>
<td>Selected</td>
<td>NA</td>
</tr>
<tr>
<td>Casson[26]</td>
<td>Case series</td>
<td>Club, NA, 0-80</td>
<td>Psychometric, CT, EEG, neurological</td>
<td>5</td>
<td>0</td>
<td>Invitation</td>
<td>NA</td>
</tr>
<tr>
<td>Kaste[27]</td>
<td>Case series</td>
<td>National, NA, 129</td>
<td>Psychometric, CT, EEG, neurological</td>
<td>8</td>
<td>0</td>
<td>Invitation</td>
<td>NA</td>
</tr>
<tr>
<td>Corsell[28]</td>
<td>Case series</td>
<td>NS (post-mortem study), NA, NS</td>
<td>Histological</td>
<td>3</td>
<td>0</td>
<td>Selected†</td>
<td>NA</td>
</tr>
<tr>
<td>Beauissant[29]</td>
<td>Case series before-after</td>
<td>NS, 1 bout, 1</td>
<td>EEG, neurological</td>
<td>123</td>
<td>0</td>
<td>Invitation</td>
<td>Before and after (that is, act as own controls)</td>
</tr>
<tr>
<td>Szymusik[30]</td>
<td>Case series</td>
<td>Long career, NA, &gt;100</td>
<td>EEG, neurological</td>
<td>60</td>
<td>0</td>
<td>Invitation</td>
<td>NA</td>
</tr>
<tr>
<td>Moriyas[31]</td>
<td>Case-control</td>
<td>NS, 1 bout, NS</td>
<td>EEG</td>
<td>10</td>
<td>300</td>
<td>NS</td>
<td>Non-boxing head injury patients</td>
</tr>
<tr>
<td>Nesarajah[32]</td>
<td>Case-control</td>
<td>Mixed, NA, NS</td>
<td>EEG</td>
<td>50</td>
<td>75</td>
<td>Invitation</td>
<td>NS</td>
</tr>
<tr>
<td>Beauissant[33]</td>
<td>Case series before-after</td>
<td>Club, 1 bout, 1</td>
<td>EEG</td>
<td>52</td>
<td>0</td>
<td>NS</td>
<td>Before and after (that is, act as own controls)</td>
</tr>
<tr>
<td>Blonstein[34]</td>
<td>Case series</td>
<td>Knockouts, NA, NS</td>
<td>EEG</td>
<td>29</td>
<td>NA</td>
<td>Selected</td>
<td>NA</td>
</tr>
<tr>
<td>Pampus[35]</td>
<td>Case series before-after</td>
<td>NS, NA, 1-4 bouts</td>
<td>EEG, neurological</td>
<td>207</td>
<td>0</td>
<td>Selected</td>
<td>Before and after (that is, act as own controls)</td>
</tr>
<tr>
<td>Blonstein[36]</td>
<td>Case-control</td>
<td>Knockouts, 1 bout</td>
<td>EEG</td>
<td>24</td>
<td>24</td>
<td>Selected</td>
<td>Non-knockout boxers</td>
</tr>
</tbody>
</table>

EEG=electroencephalography; MRI=magnetic resonance imaging; CT=computed tomography; SPECT=single photon emission computed tomography; Pt MAO=platelet monoamine oxidase inhibitor; NA=not applicable; NS=not stated.

*Selected on basis of referral to clinic with neurological problems.
†Selected on basis of dying in psychiatric hospital and having reportedly boxed at some time in life.
Psychometric testing

Direct comparison between studies was confounded by the use of more than 20 different psychometric tests (up to 12 tests in a single study); however these tests give the highest quality evidence and were used in the four cohort studies and in two controlled before and after studies. The longer duration cohort studies found that, though there were differences from controls in baseline measurements in some psychometric tests (reflecting educational background), there was no longitudinal effect of boxing on psychometric testing, even at nine years. Indeed in three studies, boxers out-performed controls on some tests.

Controlled before and after studies observed the acute effects of a boxing bout on performance, but the durations of altered results on psychometric assessment were not reported as boxers were not followed up long term. The positive findings in the smaller study were not replicated in the later, larger study, which had a longer exposure and found no overall differences, though a degree of association with exposure was present on subgroup analysis. One cohort study and two large well controlled series found an isolated abnormality of finger tapping in the non-dominant hand. A further study with several positive findings used multivariate models to explore the effect of boxing on results of psychometric testing and introduced additional covariates to control for confounding factors. Although the case series had heterogeneous findings, a large well conducted, albeit uncontrolled, prospective study found no changes in results of psychometric tests from baseline over a two year period.

Neuroimaging

Imaging studies using contemporary techniques—computed tomography, magnetic resonance imaging, SPECT (technetium-99m hexamethylpropyleneamineoxime single photon emission computed tomography), and isotope studies of cerebral blood flow (Xe133 CBF)—have built on earlier work using pneumoencephalography in professional boxers. Only one pneumoencephalographic study included an amateur boxer, who had the only normal results on encephalography. Though these studies provided the next best evidence after psychometric testing, the overall quality was poor with no cohort studies except that of Butler et al, who reported in the methods that 67% of the boxers underwent computed tomography but commented no further on this in the results or discussion.

Most of the other studies had small numbers of participants (often reflecting a selected subgroup of the whole study group). Most found no consistent abnormalities, and results correlated poorly with findings from other tests when used. For example, Kemp et al, despite identifying abnormalities both on psychometric testing and SPECT, showed no correlation between these findings. The highest positive yield from a case series was in seven of 13 boxers who underwent computed tomography, though the exact abnormalities were not described. Other series found abnormalities in individual boxers who had competed in only seven and 14 bouts and were, at the time of study, aged 55 and 37, respectively.

Electroencephalography and brain evoked potentials

In the 1940s to 1960s researchers extensively explored the potential of electroencephalography to indicate acute injury or chronic traumatic brain injury in amateur and professional boxers with variable results. Numerous early studies (case series and before and after studies) showed changes in the electroencephalogram in about half of boxers studied, with more findings immediately after bouts, although these findings were not followed up longitudinally. Two recent case series found abnormal results on electroencephalography in about half of amateur boxers studied, although results were inversely correlated with advancing age and experience (more findings in younger subjects and with fewer bouts). A third found some abnormalities in three of 10 amateur boxers, though two were aged 14 and 16 and the other was aged 53 and all had normal results on psychometric testing, neurological examination, and computed tomography. While case-control studies from the 1960s observed more findings in boxers than controls, these findings were in stark contrast to more recent case-control series and one prospective series that found no changes compared with controls or from baseline function.

Clinical neurological examination

There are several reports of clinical neurological abnormalities in small numbers of amateur boxers selected on the basis of evident symptoms or recorded acute neurological injury. Nine case series that did not select, however, showed a wide variation in prevalence and severity of findings (from none to 60%) (although we consider that the statistical method used to detect this, Mann-Whitney U test, was not appropriate). Other studies found no correlation with exposure or other methods of testing when used. Of the three case-control series, two found that non-specific findings such as tremor, nystagmus, slurred speech, and fine movement abnormalities were similarly present in controls. One large Polish study, however, found significant differences in the incidence of organic neurological dysfunction between high exposure group and controls or lower exposure groups.

Other outcomes

In 1973 Corsellis presented evidence of histological changes in the brains of 15 boxers, of whom three were
Table 3 | Results of included studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Outcome measure</th>
<th>Measure of effect</th>
<th>Results in exposed group v controls (where applicable)</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porter™</td>
<td>Psychometric Group comparison*</td>
<td>Better scores in 2/12 tests v controls</td>
<td>P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Porter™</td>
<td>Psychometric Group comparison*</td>
<td>Deterioration in finger tapping dominant hand (1/12 tests)</td>
<td>P&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Zetterberg***</td>
<td>CSF biomarkers Group comparison*</td>
<td>Increase in 3/6 CSF biomarkers v controls and baseline (at rest)</td>
<td>P=0.04/0.001</td>
<td></td>
</tr>
<tr>
<td>Moriarty***</td>
<td>Psychometric Group comparison†</td>
<td>Improvement in 1/5 tests v controls and baseline</td>
<td>P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Stewart***</td>
<td>Psychometric Odds ratio, contingency analysis</td>
<td>No differences in contingency or odds of abnormality with exposure</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Stewart***</td>
<td>EEG Odds ratio, contingency analysis</td>
<td>No differences in contingency or odds of abnormality with exposure</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Stewart***</td>
<td>Brain evoked potentials Odds ratio, contingency analysis</td>
<td>No differences in contingency or odds of abnormality with exposure</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Butler™</td>
<td>Psychometric Group comparison*</td>
<td>Improvements in 10/12 tests from baseline</td>
<td>P&lt;0.01-0.001</td>
<td></td>
</tr>
<tr>
<td>Haglund™</td>
<td>MRI Group comparison, contingency analysis</td>
<td>Small numbers of abnormalities in boxers and controls</td>
<td>Not stated</td>
<td></td>
</tr>
<tr>
<td>Haglund™</td>
<td>CT Group comparison, contingency analysis</td>
<td>CSP in 2/47 boxers v 4/50 controls, no group differences</td>
<td>Not stated</td>
<td></td>
</tr>
<tr>
<td>Brayne™</td>
<td>Creatine kinase BB Group comparison, correlation</td>
<td>Significantly increased levels v controls and with exposure</td>
<td>P&lt;0.01, P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Master™</td>
<td>Psychometric Group comparison*</td>
<td>Significant changes v controls in 5/8 tests used</td>
<td>P&lt;0.001-0.047</td>
<td></td>
</tr>
<tr>
<td>Heilbronner***</td>
<td>Psychometric Multivariate analysis</td>
<td>Significant changes, both positive (motor) and negative (memory)</td>
<td>P&lt;0.0001-0.004</td>
<td></td>
</tr>
<tr>
<td>Kemp***</td>
<td>Psychometric Group comparison†</td>
<td>Significant difference v controls in 4/5 tests</td>
<td>P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Kemp***</td>
<td>SPECT Contingency analysis</td>
<td>14/34 boxers v 5/34 controls abnormal</td>
<td>P&lt;0.02</td>
<td></td>
</tr>
<tr>
<td>Murelius™</td>
<td>Psychometric Group comparison†</td>
<td>No significant difference except finger tapping (dominant hand)</td>
<td>P&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Haglund™</td>
<td>Platelet monoamine oxidase Group comparison*</td>
<td>No difference between exposed group and controls</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Haglund™</td>
<td>Neurological Contingency analysis</td>
<td>1/47 boxers v 3/50 controls abnormal</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Haglund™</td>
<td>Brain evoked potentials Group comparison†</td>
<td>No significant difference between exposed group and controls</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Levin™</td>
<td>Psychometric Multivariate analysis</td>
<td>Exposed group moved closer to control group results</td>
<td>P=0.10-0.89</td>
<td></td>
</tr>
<tr>
<td>Levin™</td>
<td>MRI NA</td>
<td>No abnormalities found in boxers (MRI not performed in controls)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Brooks***</td>
<td>Psychometric Group comparison*</td>
<td>No evidence of neuropsychological abnormalities in boxers</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Thomassen™</td>
<td>Psychometric Multivariate analysis</td>
<td>No significant differences except finger tapping (dominant hand)</td>
<td>P&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Thomassen™</td>
<td>EEG Multivariate analysis</td>
<td>No significant difference between exposed and control group</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Thomassen™</td>
<td>Neurological Contingency analysis</td>
<td>“Sparse discrete” findings in boxers and controls</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Holzgrafe***</td>
<td>MRI Contingency analysis</td>
<td>0/13 abnormal before and after exposure</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Jellinski™</td>
<td>Psychometric Contingency analysis, correlation</td>
<td>11/60 abnormal v 0/30 controls, correlation with increasing bouts</td>
<td>Not stated, r=0.50</td>
<td></td>
</tr>
<tr>
<td>Jellinski™</td>
<td>EEG Contingency analysis, correlation</td>
<td>24/60 abnormal v 2/30 controls, correlation with increasing bouts</td>
<td>Not stated, r=0.50</td>
<td></td>
</tr>
<tr>
<td>Jellinski™</td>
<td>Neurological Contingency analysis, correlation</td>
<td>33/60 abnormal v 3/30 controls, correlation with increasing bouts</td>
<td>Not stated</td>
<td></td>
</tr>
<tr>
<td>Rodriguez™</td>
<td>Regional cerebral blood flow NA</td>
<td>No abnormalities</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>McLatchie™</td>
<td>Psychometric Group comparison*</td>
<td>9/16 abnormal, significant differences v controls in 3/10 tests</td>
<td>P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>McLatchie™</td>
<td>CT Correlation</td>
<td>1/20 abnormal in exposed group (dilated ventricles)</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>McLatchie™</td>
<td>EEG Correlation</td>
<td>8/20 abnormal (various) correlating with increasing number of fights</td>
<td>P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>McLatchie™</td>
<td>Neurological Correlation</td>
<td>7/20 abnormal correlating with increasing number of fights</td>
<td>P&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Ross™</td>
<td>CT NA</td>
<td>7/13 abnormal (not specified)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Ross™</td>
<td>EEG NA</td>
<td>4/8 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Ross™</td>
<td>Neurological NA</td>
<td>1/8 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Jordan™</td>
<td>MRI NA</td>
<td>1/4 abnormal (congenital or post-traumatic cyst of hippocampus)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Jordan™</td>
<td>CT NA</td>
<td>1/4 abnormal (congenital or post-traumatic cyst of hippocampus)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Legwold™</td>
<td>Neurological NA</td>
<td>68 concussions/7000 bouts, none resulting in neurological dysfunction</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Jordan™</td>
<td>MRI NA</td>
<td>0/9 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Jordan™</td>
<td>Neurological NA</td>
<td>0/9 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Casson™</td>
<td>Psychometric NA</td>
<td>2/5 abnormal as defined by “impairment index”</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Casson™</td>
<td>CT NA</td>
<td>1/5 abnormal (generalised cerebral atrophy)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Casson™</td>
<td>EEG NA</td>
<td>1/4 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Casson™</td>
<td>Neurological NA</td>
<td>1/5 abnormal (mild “organic mental syndrome,” right Babinsky)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Kaste™</td>
<td>Psychometric NA</td>
<td>0/8 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Kaste™</td>
<td>CT NA</td>
<td>1/8 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Kaste™</td>
<td>EEG NA</td>
<td>4/7 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Kaste™</td>
<td>Brain evoked potentials NA</td>
<td>1/7 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Kaste™</td>
<td>Neurological NA</td>
<td>0/8 abnormal</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Corsellis™</td>
<td>Histological NA</td>
<td>0/3 abnormal</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>
amateur. Aside from possible flaws of attribution in this study (such as alcohol, syphilis, and head injuries from other causes) in association with positive findings in professional boxers, the authors concluded that no changes specific to boxing were present in the amateurs. Two studies examined neurochemical changes in the blood and cerebrospinal fluid of boxers after competition compared with athletic or non-athletic controls, respectively, and found significantly higher concentrations in boxers. In the latter studies each boxer underwent two lumbar punctures. These increases were said to indicate disruption of the blood-brain barrier or acute neuronal and astroglial injury.

**DISCUSSION**

In this systematic review we found no evidence for a strong association between amateur boxing and chronic traumatic brain injury. In boxing the head might get hit repeatedly with resultant concussion, though less than in several more popular sports—such as rugby union and equestrian activities—which may harm cerebral function. Whether clinically measurable long term brain injury occurs is a different and more important question. There is reasonable clinical, radiological, and histopathological evidence that this is the case in a proportion of professional boxers (10-20% in most studies), although most studies were performed at a time when safety standards were far less stringent than they are today.

Amateur boxing is a different sport from professional boxing, including in its motivation to participate, rules, and equipment, but, most importantly, there is considerably greater exposure to injury in professionals (increased frequency and force of punches over a greater duration of career). We looked at the data for chronic traumatic brain injury in amateur boxing alone. Although no formal synthesis was performed, the data can be described in summary. Overall, 15 of 36 studies (42%) included in the systematic review concluded that relevant abnormalities were present, at least in a proportion of boxers studied. When we expressed this as a function of all methods tested (see table 3) we attained a similar figure (28/63, 44%).

**Limitations**

It would clearly be impossible to perform a double blind randomised controlled trial for amateur boxing, though in general study design and conduct could have been greatly improved. Few studies were of sufficient quality to conclude anything other than a weak association when positive findings were reported, and none was sufficiently powered (no sample size calculations performed) to exclude a type II error when results were negative. Only two studies supplied confidence intervals for the main results. There was a definite tendency towards positive findings in studies of poorer quality and design. For instance, none of the four cohort studies had positive results, with three actually showing improvements over the study period. This contrasts with the finding of abnormalities in over 50% of case series. With a cut-off of quality ≥3/6, only four of 17 (24%) studies and five of 20 (20%) methods of testing yielded abnormal results. The latter is in contrast to studies that scored ≤2 on quality, in which 62% (23/37) yielded positive results. Although perhaps not of importance, only two of 14 studies performed from 1990 onwards concluded that any measurable abnormality was present (and one of these was in a single boxer). The importance of using controls was illustrated by several case-control studies that showed that potentially severe abnormalities on clinical neurological examination were present equally in the control group.

**Bias**

Methods of selection were rarely adequately explained and occasionally performed on the basis of prior abnormal clinical or investigative findings. In terms of design, when controls were used these were poorly selected in terms of possible confounding factors. For instance, in one study that used psychometric tests, the controls (rugby and water polo players) were drawn from an undergraduate population, whereas many of the boxers had not completed their full time education. As no data on IQ had been gathered this factor could not be assessed, and it is acknowledged that education and vocabulary have a large weighting on results of neuropsychometric
WHAT IS ALREADY KNOWN ON THIS TOPIC

The safety of amateur boxing in terms of risk of chronic traumatic brain injury continues to be questioned. No recent or systematic review has been performed to assess the evidence for this.

WHAT THIS STUDY ADDS

A systematic review of observational studies indicates that, although the quality of evidence supporting or refuting the hypothesis was poor, the association between amateur boxing and chronic traumatic brain injury is not strong.

testing. In respect of performance, remarkably in only two studies were observers blinded. Finally, though we did not carry out a formal analysis of publication bias, studies showing adverse effects might have been more likely to get published.

In studies that sought an association between exposure and outcome, few found an effect, raising the question of false attribution. Indeed, questions of specificity exist for almost all methods used. In psychometric testing, three well conducted studies found a single significant difference in the finger tapping test of the non-dominant hand. This finding has been replicated in other studies and seems to be worse with increased exposure to boxing. The association between finger tapping response and brain damage is difficult to assess in boxers, however, given the chronic damage to the fingers directly associated with repeated punches. The observation in electroencephalography studies of an inverse association between exposure or age and positive findings might similarly be explained by the recognised false positive rate of electroencephalography, particularly in the young. The problems implicit in analysis are also illustrated by the revision of findings (from highly significant to zero) by a single group in two consecutive publications. Imaging methods have similar inherent difficulties of interpretation with the relevance of some much championed findings, such as cavum septum pellucidum and ventricular abnormalities, questioned by others. This issue is well reviewed elsewhere, and the relevance of many of these abnormalities remains dubious. In respect of blood concentrations of creatine kinase BB, increased levels have also been observed in oarsmen and marathon runners. When, as in some series, participants in their 50s were described with clinical neurological abnormalities after a limited exposure (as few as seven bouts) some 30 years previously, the sole attribution of these findings to boxing must surely be questioned.

Conversely, the current range of tests might lack sensitivity to detect subtle changes in neural structure or function. All these tests must be regarded as surrogate markers for the notional concept of chronic traumatic brain injury and clearly no conclusion can be reached on this without an ideal test for comparison or indeed a clear definition of what might constitute clinically relevant injury. Nevertheless, tests regarded as sensitive in general neurological practice have all been used. In particular, psychometric testing, regarded by some as the most sensitive, provided the most conclusive negative results. Similarly, it is generally accepted that magnetic resonance imaging is the best method of determining subtle parenchymal damage and degenerative change. In the six studies that used this, only one case series of four boxers concluded that relevant abnormality was present. This was a cyst in a single boxer, which was possibly congenital. No abnormalities were found in the single cohort study that used magnetic resonance imaging.

Finally, because of the short duration or “snapshot” design of nearly all studies (except that of longer follow-up cohort studies in which no detrimental effects were found), it is impossible to conclude whether or not longer exposure would have eventually led to chronic injury or whether such changes might present in much later life when further neuronal loss occurs with ageing. Implicit within this latter argument is the possibility that subclinical, sub-psychometric, and sub-radiological brain damage incurred as an amateur may contribute to that which becomes clinically evident in those who subsequently have a long professional boxing career. This was not, however, indicated by findings at nine year follow-up.

Conclusions

Amateur boxing is becoming an increasingly popular participation sport, especially within universities and for both sexes. The safety of boxing is an issue that stimulates emotive responses on both sides of the debate, and calls to ban the sport continue. This review neither seeks to endorse nor oppose the sport of amateur boxing. It is perhaps a question of personal philosophy whether it is incumbent on boxing to prove that it is safe, or on those who oppose it to prove that it is deleterious (although it might be argued that those wanting to alter the status quo have the responsibility to prove this). Nevertheless, on the basis of this systematic review, we conclude that the current evidence, such as it exists, for chronic traumatic brain injury as a consequence of amateur boxing is not strong.

Contributors: ML and CK made substantial contributions to conception and design of the study, acquisition, analysis, and interpretation of data, and drafting and revising the article. GW critically revised the article for important intellectual content. All authors gave final approval of the version to be published. ML is guarantor.

Funding: None.

Competing interests: ML is lead sports physician (London Region) at the English Institute of Sport, Olympic Medical Institute, and physician to the British Amateur Boxing Team. ML and CK are commissioners on the ABA of England Medical Commission.

Ethical approval: Not required.

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


16 Critchley M. Medical aspects of boxing, particularly from neurological standpoint. *BMJ* 1957;i:357-62.


Accepted: 7 August 2007
**Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials**

Catherine J Minns Lowe, research fellow,1 Karen L Barker, director,2 Michael Dewey, special lecturer,3 Catherine M Sackley, professor of physiotherapy research1

**ABSTRACT**

**Objective** To evaluate the effectiveness of physiotherapy exercise after elective primary total knee arthroplasty in patients with osteoarthritis.

**Design** Systematic review.


**Review methods** Randomised controlled trials were reviewed if they included a physiotherapy exercise intervention compared with usual or standard physiotherapy care, or compared two types of exercise physiotherapy interventions meeting the review criteria, after discharge from hospital after elective primary total knee arthroplasty for osteoarthritis.

**Outcome measures** Functional activities of daily living, walking, quality of life, muscle strength, and range of motion in the knee joint. Trial quality was extensively evaluated. Narrative synthesis plus meta-analyses with fixed effect models, weighted mean differences, standardised effect sizes, and tests for heterogeneity.

**Results** Six trials were identified, five of which were suitable for inclusion in meta-analyses. There was a small to moderate standardised effect size (0.33, 95% confidence interval 0.07 to 0.58) in favour of functional exercise for function three to four months postoperatively. There were also small to moderate weighted mean differences of 2.9 (0.61 to 5.2) for range of joint motion and 1.66 (−1 to 4.3) for quality of life in favour of functional exercise three to four months postoperatively. Benefits of treatment were no longer evident at one year.

**Conclusions** Interventions including physiotherapy functional exercises after discharge result in short term benefit after elective primary total knee arthroplasty. Effect sizes are small to moderate, with no long term benefit.

**INTRODUCTION**

Osteoarthritis is the commonest cause of disability in older people,1 with painful knee osteoarthritis affecting 10% of people aged over 55 in the United Kingdom.2 Over 80% of patients experience limitations in performing activities of daily living, such as mobility outside the home, household chores, and work duties.3 In 2005, patients with osteoarthritis accounted for at least 55 495 primary knee joint arthroplasties in England and Wales.4 As the length of hospital stay after joint arthroplasty surgery has markedly and rapidly decreased,5 and given that patients who undergo knee arthroplasty may still experience considerable functional impairment postoperatively,6 the effectiveness of physiotherapy after discharge is a valid question. The present uncertainty regarding effectiveness makes it difficult for commissioning organisations, healthcare practitioners, and patients to make decisions regarding such physiotherapy. We systematically reviewed randomised controlled trials to determine the effectiveness of physiotherapy exercise after discharge in terms of improving function, quality of life, walking, range of motion in the knee joint, and muscle strength for patients with osteoarthritis after elective primary unilateral total knee arthroplasty.

**METHODS**

**Searching**

In March 2005 and in April 2007 we identified randomised controlled trials by simultaneously searching AMED (from 1985), CINAHL (from 1982), Embase (from 1974), Kings Fund database (from 1979), and Medline (from 1966) via Knowledge Access 24/7 (KA24). We also searched the Cochrane library, PEDro physiotherapy evidence database, and the Department of Health national research register. In July 2005 and April 2007 we handsearched Physiotherapy (1985- March 2007 inclusive) and Physical Therapy (1985-April 2007 inclusive) to double check for trials. The conference proceedings in the Journal of Bone and Joint Surgery (Britain) (1985-2006 inclusive) were also handsearched, as were the reference lists of included trials.

As it is difficult to locate physiotherapy trials, we considered that using multiple general searches was the optimum method. This review is part of a series that included both knee and hip search terms. Table 1
summaries the searches. No language restrictions were applied.

Selection
We sought randomised controlled trials of patients undergoing elective total knee arthroplasty for osteoarthritis who received an intervention of physiotherapy exercise after discharge from hospital. We used broad definitions of “physiotherapy” and “exercise” to include any exercises or exercise programme advised or provided by physiotherapists or physical therapists during the rehabilitative period after discharge from hospital after surgery in the outpatient, community, or home setting. We excluded trials in which the intervention consisted of an electrical adjunct to physiotherapy, such as use of continuous passive motion. Physiotherapy exercise interventions included outpatient physiotherapy sessions and functional physiotherapy programmes, in which exercises are based on functional activities. Trials were included if they investigated a physiotherapy intervention compared with usual or standard care or compared two different types of relevant physiotherapy intervention. Usual or standard care refers to the continuation of home exercise programmes provided to patients during a stay in hospital. These programmes usually consist of isometric or simple strengthening exercises, exercises to regain range of movement, and stretches. Effectiveness outcomes were measures of functional activities of daily living, walking, self reported measures of quality of life, muscle strength, and range of motion in the knee joint. As most trials use functional measures rather than specific pain outcomes, we did not include pain as an effectiveness outcome. Two

Table 1 | Search strategy for systematic review

<table>
<thead>
<tr>
<th>Sources, searches, and search terms</th>
<th>Mar-Jul 2005 hits*</th>
<th>2005-April 2007 hits*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “rehabilitation” AND “trials” (whole document)</td>
<td>587 (25)</td>
<td>180 (0)</td>
</tr>
<tr>
<td>2. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “rehabilitation” AND “trials” (title)</td>
<td>118 (11)</td>
<td>1 (0)</td>
</tr>
<tr>
<td>3. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “physiotherapy” AND “trials” (title)</td>
<td>2 (0)</td>
<td>4 (0)</td>
</tr>
<tr>
<td>4. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “physiotherapy” (title)</td>
<td>39 (0)</td>
<td>14 (0)</td>
</tr>
<tr>
<td>5. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “physical therapy” (title)</td>
<td>43 (8)</td>
<td>15 (0)</td>
</tr>
<tr>
<td>6. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “home programme” (title)</td>
<td>2 (0)</td>
<td>1 (0)</td>
</tr>
<tr>
<td>7. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “home programme” (whole document)</td>
<td>22 (2)</td>
<td>27 (0)</td>
</tr>
<tr>
<td>8. “hip” OR “knee” (whole document) AND “replacement” OR “arthroplast$” (whole document) AND “occupational therapy” (whole document)</td>
<td>35 (0)</td>
<td>3 (0)</td>
</tr>
<tr>
<td>9. “hip” OR “knee” (whole document) AND “occupational therapist$” (title)</td>
<td>0 (0)</td>
<td>3 (0)</td>
</tr>
</tbody>
</table>

Cochrane library (Cochrane reviews, CCRT, DARE):
1. Browsed by topic—musculoskeletal, search narrowed—osteoarthritis, search narrowed—rehabilitation | 9 | 11 |
2. General search term “joint replacement” | 80 | 18 |

PEDro physiotherapy evidence database:
1. “joint replacement AND rehabilitation” | 1 (0) | 17 (0) |
2. “joint replacement” | 5 (0) | 45 (0) |

Department of Health national research register:
1. “joint replacement AND rehabilitation” | 0 | 19 (1) |
2. “joint replacement AND physiotherapy” | 7 | 9 (0) |
3. “joint replacement AND exercise” | 2 (0) | 6 (0) |
4. “joint replacement AND physical therapy” | 3 (0) | 5 (0) |
5. “joint arthroplasty AND physiotherapy” | 0 | 0 |
6. “joint arthroplasty AND rehabilitation” | 2 (1) | 2 (0) |
7. “joint replacement AND occupational therapy” | 5 (0) | 5 (0) |

Physiotherapy key journal—hand search of contents pages | None new | None new |
Physical Therapy key journal—hand search of contents pages | None new | None new |
JBIS [Br]—hand search of all conference proceedings | 2 new | 1 |
Reference lists—hand searching of papers included in review | 1 new | None new |
Totals | 965 (50) | 386 (2) |

*Numbers after removal of duplicates commands when available.
Reviewers (CML and CS) assessed and agreed on study eligibility.

Validity assessment, data abstraction, and quality assessment

We developed and piloted a data extraction form using quality indicators from the CONSORT statement \(^7\) and the CASP guidelines \(^8\) (Table 2). Similar analysis of individual quality components has previously been used in reviews of physiotherapy \(^9\) and is advocated to avoid known problems associated with existing composite scores \(^10\). Items could be marked as yes, no, unclear, or partial. Items were marked as yes only if they fully and explicitly met the detailed criteria laid out in the CONSORT standards. \(^7\) Two reviewers (CML and KB) independently extracted the data. KB was masked to the key details of each paper and the extent to which masking was successful was assessed. The masking rates were 80% for authors, 20% for journals, 80% for author affiliations, and 80% for funding sources, all of which except journal of publication were considered successful. The level of agreement between reviewers was 69.09% (κ = 0.524, intraclass correlation coefficient (2,1) = 0.49, 95% confidence interval 0.30 to 0.63).

We resolved initial disagreements regarding study quality by discussion until consensus was reached. Major disagreement was rare; usually disagreement was the more minor “yes” to “partial/unclear” or

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Quality component checklist and quality evaluation of six trials included in review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contained in study</td>
<td>Codine et al, 2004(^w1)</td>
</tr>
<tr>
<td>Rationale for study</td>
<td>Yes</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>Yes</td>
</tr>
<tr>
<td>Recruitment method</td>
<td>No</td>
</tr>
<tr>
<td>Settings and location of study</td>
<td>No</td>
</tr>
<tr>
<td>Intervention</td>
<td>Partial</td>
</tr>
<tr>
<td>Objectives/hypotheses</td>
<td>Partial</td>
</tr>
<tr>
<td>Defined outcome measures</td>
<td>Yes</td>
</tr>
<tr>
<td>Quality enhancers (such as multiple observations)</td>
<td>No</td>
</tr>
<tr>
<td>Sample size determination</td>
<td>No</td>
</tr>
<tr>
<td>Randomisation</td>
<td>§</td>
</tr>
<tr>
<td>Randomisation sequence generation</td>
<td>No</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>No</td>
</tr>
<tr>
<td>Randomisation implementation methods</td>
<td>No</td>
</tr>
<tr>
<td>Blinding of participants</td>
<td>Inappropriate</td>
</tr>
<tr>
<td>Blinding of those administering intervention</td>
<td>Inappropriate</td>
</tr>
<tr>
<td>Blinding of outcome/assessments</td>
<td>Yes</td>
</tr>
<tr>
<td>Statistical methods</td>
<td>¶</td>
</tr>
<tr>
<td>Flow of participants through each stage</td>
<td>No</td>
</tr>
<tr>
<td>Recruitment and follow-up dates</td>
<td>No</td>
</tr>
<tr>
<td>Baseline demographics</td>
<td>Partial</td>
</tr>
<tr>
<td>Numbers analysed (and ITT)</td>
<td>No</td>
</tr>
<tr>
<td>Summary of results</td>
<td>No</td>
</tr>
<tr>
<td>Estimated effect sizes</td>
<td>No</td>
</tr>
<tr>
<td>Precision</td>
<td>No</td>
</tr>
<tr>
<td>Results for each outcome</td>
<td>Yes</td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>No</td>
</tr>
<tr>
<td>Adverse events</td>
<td>No</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Partial</td>
</tr>
<tr>
<td>Generalisability</td>
<td>Partial</td>
</tr>
<tr>
<td>Results placed into context</td>
<td>Partial</td>
</tr>
<tr>
<td>Quality (good enough to include in meta-analyses)</td>
<td>No</td>
</tr>
</tbody>
</table>

*Published abstract and information from authors only, therefore reporting is incomplete.
†Intervention described but little description of home exercise programme described to both groups.
‡Feasibility trial that provided sample size calculations as part of results section.
§First participant was drawn randomly then alternatively assigned (this process was witnessed).
¶Additional information from author stated that sample size was determined by statistician who calculated n=30 in each arm to be sufficient.
**Intention to treat analysis intended and performed but per protocol analysis presented because loss to follow-up in control group favoured intervention group.
"no" to "partial/unclear" and 100% agreement was obtained. A third reviewer (CS) was available in the event of consensus not being reached, but this was not required. Where key study details were absent or unclear we contacted authors for further information.

We considered studies to be of good quality if they were good enough to include in meta-analyses. Table 2 presents quality assessment findings for each study. We excluded one study\(^6\) from the meta-analysis because participants were allocated by alternation. All trial outcomes were measured by assessors masked to allocation. As table 2 shows, most studies clearly reported the flow of participants through the trial, justified the sample size, and included intention to treat analyses. Several quality indicators were not fully discussed in all papers, such as allocation concealment and details regarding the implementation of randomisation methods.

### Table 3 | Studies excluded from the review

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not a randomised clinical trial</td>
<td>Tum-Sugden 1976,(^{67}) Ulreich et al 1997,(^{68}) Benedetti et al 2003,(^{69}) Kitter et al 1989,(^{70}) Waters 1974(^{71})</td>
</tr>
<tr>
<td>Osteopathic manipulation intervention</td>
<td>Licciardone et al 2004(^{82})</td>
</tr>
<tr>
<td>Neuromuscular stimulation intervention</td>
<td>Stevens 2002(^{83})</td>
</tr>
<tr>
<td>Preoperative intervention</td>
<td>Gursen and Ahrens 2003(^{84})</td>
</tr>
<tr>
<td>Comparison of exercise and continuous passive motion</td>
<td>Worland et al 1998(^{85})</td>
</tr>
<tr>
<td>Study halted early with no results</td>
<td>Stanley 2004(^{86})</td>
</tr>
<tr>
<td>Duplicate trial report</td>
<td>Mockford et al 2006(^{87})</td>
</tr>
</tbody>
</table>

### Table 4 | Heterogeneity $\chi^2$ and $I^2$ test results for range of outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Three to four months after surgery</th>
<th>12 months after surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function</td>
<td>$\chi^2=0.78, df=2, P=0.68$</td>
<td>$\chi^2=2.35, df=3, P=0.50$</td>
</tr>
<tr>
<td>Walking</td>
<td>$\chi^2=0.54, df=1, P=0.46$</td>
<td>$\chi^2=0.41, df=1, P=0.52$</td>
</tr>
<tr>
<td>Joint range of motion</td>
<td>$\chi^2=1.46, df=2, P=0.48$</td>
<td>$\chi^2=2.28, df=3, P=0.52$</td>
</tr>
<tr>
<td>Quality of life</td>
<td>$\chi^2=0.91, df=2, P=0.63$</td>
<td>$\chi^2=0.91, df=2, P=0.63$</td>
</tr>
</tbody>
</table>

\(*I^2=0% in all tests.*

### Results

We identified and screened 27 potentially relevant studies. Of these, six studies\(^{6-11}\) were included in the systematic review and five\(^{12-16}\) in the meta-analysis (fig 1). Table 3 gives details of excluded studies.\(^{67-87}\) Table 4 provides the results of the analysis of heterogeneity.

### Study characteristics

Table 5 summarises the characteristics of the included studies and provides information regarding the participants, interventions, main outcomes, and conclusions reached by authors.

### Summary of the interventions and comparisons

With the exception of one trial,\(^{66}\) in-depth details of the intervention and comparison groups were available from the papers and authors (table 6).

The trial interventions were similar to each other in that they provided additional physiotherapy exercises or treatment after discharge after total knee arthroplasty, often involving programmes of functional

### Quantitative data synthesis

We carried out meta-analyses for knee function, walking, range of joint motion, and quality of life with R2.3.1 and the rmeta package.\(^{11}\) Our outcome was the score at the chosen time point rather than the change in score as this maximised the number of comparable studies. The time points used were three to four months after surgery and 12 months after surgery. If the same measure was reported we used weighted mean differences, otherwise we used standardised effect sizes (small (0.2), medium (0.5), and large (0.8)\(^{12}\)). We used fixed effect models and 95% confidence intervals throughout and performed tests of heterogeneity ($\chi^2$) at a 5% significance level, though we accept these have low power because few studies were available for meta-analyses. We also calculated $I^2$ to give a measurement of the degree of heterogeneity between the trials in the meta-analysis. Random effects models were not considered as there was no compelling evidence of heterogeneity and estimating the variation between studies is difficult with such low numbers. The differences were calculated so that positive differences indicate that the effect favoured treatment and negative differences that the effect favoured control or usual care. We considered it inappropriate to assess publication bias because of the small number of trials.
weight bearing exercise. The study by Rajan et al provided few details regarding the intervention.\textsuperscript{w6} Though most interventions included functional weight bearing exercises, Codine et al investigated the effect of eccentric isokinetic muscle strengthening with a CYBEX dynamometer.\textsuperscript{w1} Interventions usually started within two weeks of discharge. Outpatient programmes generally lasted up to 12 weeks, while home exercise programmes were recommended for up to a year or indefinitely in one case.\textsuperscript{w3}

The comparison groups were mainly control groups in which no additional outpatient physiotherapy was organised. Patients were expected to continue with the traditional home exercise programme—namely, isometric strengthening and range of movement exercises plus gait training or re-education provided to all patients during their stay in hospital.

Quantitative data synthesis

*Measures of function (five trials)*

Five of the studies contained a measure of function.\textsuperscript{w1-w5} The measures used included the 12 item Oxford knee score,\textsuperscript{w2} which measures functional ability, including pain, (scores 12-60, low score indicates high function) (Frost et al used one item of this score\textsuperscript{w2}); the American Knee Society clinical rating score,\textsuperscript{w1} which measures pain, movement, stability, and functional activity (scores 0-100, high score indicates favourable); the 24 item Western Ontario and McMaster Universities osteoarthritis index (WOMAC),\textsuperscript{w3} which has domains for pain, stiffness, and function (scored as a percentage by Moffet et al\textsuperscript{w5} and out of 0-170 for functional ability, including pain, (scores 12-60, low score indicates high function) (Frost et al used one item of this score\textsuperscript{w2}); the Bartlett patellar score,\textsuperscript{w4} which measures anterior knee pain, quadriceps strength, and function (scores 3-30, high scores are favourable).

Within the individual trials, three found no significant differences between groups.\textsuperscript{w1-w3} Frost et al found significant differences within groups for the treatment arm, indicating a benefit of treatment.\textsuperscript{w2} Mockford and Beverland presented no results in their published abstract but supplied summary statistics for their outcomes,\textsuperscript{w4} allowing us to include their study in the meta-analysis. Moffet et al found significant differences between the two groups, in favour of the intervention, at four and six months after arthroplasty but not at 12 months.\textsuperscript{w5}

Figure 2 shows the three studies with data on functioning at three to four months and 12 months after

![Figure 2](https://example.com/figure2.png)
surgery. Where studies included more than one measure of function we decided to use the Oxford knee and the WOMAC scores as these encompassed all component trials. No trial included both these scores. At three to four months the standardised effect size was 0.33 (95% confidence interval 0.07 to 0.58), which is considered small to moderate.12 At 12 months, with one additional study, the effect size was close to zero at −0.07 and the confidence interval (−0.28 to 0.14) included zero.

Walking (three trials)

Three knee arthroplasty trials used some form of outcome measure for walking.13,14,15 The measures reported included walking speed over a 10 metre distance, measured in m/sec,14 and a six minute timed walking test, measured in metres.15,16,17 The study by Moffet et al15 reported on time walking over a 50 metre walkway.

The results from these three trials were mixed. One trial found no significant differences between groups,15

| Table 6 | Summary of trial interventions and comparisons included in the knee replacement trials (n=6) |
|----------------|----------------------------------|------------------|------------------|
| Early programme provided to all trial participants | Intervention details | Comparison group details |
| Codine et al, 20041,2 | Knee mobilisation (continuous passive motion and manual), isometric muscle strengthening for all knee and leg muscle groups, “propropioceptive enhancement,” walking exercises | Submaximal hamstrings muscle eccentric isokinetic strengthening (passive resist mode). Torque produced by half of torque measured during testing. CYBEX dynamometer training speed of 10 degrees/s. Range of motion during flexion was conducted in active assist mode. Programme from day 10-30 post total knee arthroscopy; CYBEX 5 mins/day, 5 days/week for 3 weeks | Nil else added |
| Frost et al, 20042,3 | Gait re-education, mobilising, and strengthening exercises (including active knee flexion, straight leg raises, inner range quadriceps, isometric quadriceps) | Warm up: sitting knee flexions 10 repetitions. Chair rise: baseline number set and increased every alternate day up to 2 mins, then repeat up to 3 times/day; Walk: 1 min normal pace, increment 30 s/day until up to 10 mins, repeat 2-3/day. Leg lifts to step/thick book: Baseline 1 set. Increase by 1 per day to 2 min duration. Then repeat 2-3/day | Long sitting: static quadriceps (5 s hold); straight leg raise (5 s hold), inner range quadriceps (5 s hold). Supine: knee flexion heel slides. Standing: knee flexion (5 s hold). Sitting: long arc quadrats (3 s hold) |
| Kramer et al, 20033,3 | Performed 3 times/day until 12 week check-up, then 2 once a day. Stage I—supine: knee flexion and extension. Long sitting: autoassisted knee flexion, ankle dorsiflexion in knee extension with calf stretch. Supine: isometric knee extension, inner range quadriceps. Supported sitting: hamstrings stretch. Sitting: active knee extension. 10 repetitions. Stretch/holding for 5 s. Stage II—prone lying: quadriceps stretch. Standing: quadriceps stretch, soleus stretch, Achilles stretch, knee flexion (progress to ankle weights). Supine: straight leg raise. Sitting: resisted knee extension, resisted knee flexion, sit-stand-sit. 10 repetitions, stretch/holds for 5 s. Optional exercises: exercise bike; standing wall sits 5-10 repetitions; standing slow short squats 10 repetitions. | Outpatient physiotherapy weeks 2-12 after surgery: up to two sessions/week, each session about 1 hour. Exercises could be added/modified, therapeutic modalities (ice, heat, ultrasound), joint mobilisations, or other measures as appropriate. Patients requested to complete common home exercise only twice on clinic session days | Physiotherapist phoned patient ≤1 during weeks 2-5 and 7-12 after surgery for queries, advice. Patients able to phone with queries |
| Mockford and Beverland, 20044,5 | Supine: active ankle dorsiflexion and plantarflexion (10 repetitions), isometric quadriceps and hamstring (3 repetitions), straight leg raise (repetitions vary), physiotherapist assisted knee flexion (5 repetitions), heels slides (3 lots of 10 repetitions, 2 mins rest between each set), active knee extensions over roll/bar (3 sets of 10 repetitions, as able, 2 mins rest between each set). High sitting: proprioceptive neuromuscular facilitation. Hamstrings pulley with 2 kg weight (3 lots of 10 repetitions, 2 mins rest between each set). Gait re-education with crutches/sticks as appropriate. Stairs practice | 9 outpatient physiotherapy sessions in 6 weeks (2 sessions in weeks 1, 2, and 3, and 1 session in weeks 4-6). Week 1—heel slides, isometric quadriceps and hamstrings, straight leg raise, active knee extension over bar/roll and hamstrings pulley—all as in early programme. Proprioceptive neuromuscular facilitation, physiological mobilisations for flexion and extension (3 sets of 10 repetitions with 2 mins rest between sets). Weeks 2-3—standing: weight shifts (10 repetitions), quarter squats (10 repetitions). Pone lying: autoassisted quadriceps stretch (10 repetitions). Sit-stand-sit (10 repetitions), Gait re-education. Weeks 4-6—proprioceptive work in parallel bars. Gait re-education. Standing: stepping over cones, wobble board, step-ups (10 repetitions). Exercise bike for 5 mins | No outpatient physiotherapy |
| Moffet et al, 20045,5 | “Simple exercises” to regain lower limb strength (quadriceps, hamstrings, hip abductors, and hip extensors) and to increase knee range of motion. Advice about knee positioning, icing application, and gait retraining | 12 sessions of outpatient physiotherapy in 6-8 weeks, 60-90 mins each at clinic visit starting 2 months after surgery. Plus individual home exercise programme. Warm up (5-10 mins); lower limb flexion/extension, alternate ankle dorsiflexion-plantarflexion, hamstrings stretch. Specific strength exercises (15 mins): isometric knee extension in 0° and 60° flexion at visits 1-2; isometric hamstrings 60° flexion at visits 3-6; concentric eccentric hip abductors against gravity at visits 1-4. Functional task oriented exercises (15-20 mins): get up-sit down at visits 1-6; knee extensor strengthening in standing with Theraband at visits 1-6; controlled bilateral knee flexion-extension in standing at visits 1-8; unilateral knee flexion to 90° in standing at visits 7-10; climbing on platform/stairs at visits 3-12; walk backwards on slope and/or laterally while crossing lower limbs at visits 3-12; walk in place, with large amplitude hip and knee flexion and upper limb movements at visits 9-12. Endurance exercises (5-20 mins): walk at visits 3-12, exercise bike at visits 4-12. Cool down (10 mins); slow walk, strength, ice | Usual care. 26% patients had home visit. No attempt made to limit care. Information about usual care obtained by questionnaire and by phone interviews with patients and physiotherapist |
| Rajan et al, 20044,5 | All patients given home exercise programme to follow on discharge | Outpatient physiotherapy, 4-6 sessions | No organised physiotherapist |
another found differences approaching significance, and the third trial found significant differences within intervention groups. Figure 3 shows that the intervention had no overall influence on walking at either three or 12 months.

Range of joint motion (five trials)
Five of the total knee arthroplasty trials used the range of motion in the knee joint as an outcome measure. Although all measurements were provided in degrees, the method of achieving results varied. Codine et al used a goniometer integrated into a dynamometer to measure knee flexion and extension, while Mockford and Beverland used a goniometer to measure active and passive flexion and extension. Frost et al and Kramer et al both measured active flexion only, and Rajan et al measured range of motion in the knee as a single value.

Once again, the results were mixed. Codine et al found a significant difference in knee extension between the two groups at 10 days, though, despite randomisation, extension was different in the two groups at baseline. Another study concluded that there was a significant difference in active knee movement in favour of the intervention group but not in the passive range. In the pilot study by Frost et al there was a trend for less loss of range in the functional group than in the traditional exercise group but the study was small and the difference was not significant. Two other studies also found no significant differences.

All the studies on range of movement in the knee joint used the same measure (degrees); therefore figure 4 shows the weighted mean differences and confidence intervals. The three month summary shows an increase of 2.9° (0.61° to 5.2°), which is considered small to moderate. At 12 months the effect was smaller, about 1°, and the confidence interval (−1.10° to 3.00°) included zero.

Quality of life (three trials)
Three trials included measures of quality of life. The SF-36 health survey provides an eight scale profile of functional health and wellbeing scores with low scores indicating poor health. Kramer et al used the SF-36, and Moffet et al used a French translation of the same score. Moffet et al also provided the physical component and mental component scores of the SF-12, as did Mockford and Beverland.

One trial found no significant differences between the groups. One other trial has not yet presented statistical analyses for this measure. The final trial found small significant differences in favour of the intervention group for the role-physical dimension of the SF-36, and Moffet et al used a French translation of the same score.

Figure 5 presents the studies with data on quality of life. At three to four months the studies used the same measure, the SF-12, and so we have presented weighted mean difference results. At 12 months after surgery, however, not all studies used the same measure and therefore we used standardised effect sizes.

At three to four months after surgery the weighted mean difference was 1.7 (−1.0 to 4.3), indicating a small effect in favour of the intervention. At 12 months the
effect was close to zero with a standardised effect size of 0.03 (−0.20 to 0.25).

**Muscle strength**

None of the trials included in the review directly measured muscle strength.

**DISCUSSION**

This systematic review provides support for the use of physiotherapy exercise interventions with exercises based on functional activities after discharge, rather than traditional home exercise and advice programmes, to obtain short term benefit after elective primary knee arthroplasty. There was a small to moderate standardised effect size in favour of functional exercise for function three to four months postoperatively. Small to moderate weighted mean differences, in favour of functional exercise interventions, were seen for range of joint motion and quality of life three to four months postoperatively. Any benefits seen after treatment did not persist to one year follow-up.

**Strengths and weaknesses of review procedures**

Physiotherapy literature remains a difficult area to search, with numerous bibliographic databases and unindexed journals. While we made every attempt to identify studies in any language, other studies might exist. We believe, however, that this review remains the most comprehensive to date.

Trial quality was good overall. Of the five adequately randomised studies included in the meta-analyses, most were sufficiently powered with adequate strategies to conceal allocation and outcome measurements obtained by assessors blinded to treatment allocation. Yet, like most physiotherapy trials, studies were relatively small, with 554 participants in the five trials included in the meta-analyses and 614 participants included overall in the review.

The most commonly used outcomes were function, predominantly subjective measures of functional ability, and range of joint motion as an objective measure. While range of joint motion is important, its usefulness as an outcome measure of physiotherapy interventions is limited as other factors, such as prosthetic design, preoperative knee motion, and surgical technique, also influence postoperative range of joint motion. None of the trials directly measured muscle strength, although one included leg extensor power instead of muscle exercises and exercises to increase range of motion in the joint.

There were no apparent problems with our data extraction processes. Although many quality checklists and scales exist, there is no accepted ideal score; component approaches are often preferred as the wide variety of scores and weighting systems available mean that the same trial may score as both high quality and low quality depending on which score is used. Additionally, many scoring systems downgrade the quality rating of a trial if it is not double blinded. For many physiotherapy trials, such as those in this review, patients and therapists inevitably know the treatment allocation and this is not an indication of low or high trial quality. For these reasons we used a component approach, although we accept this is controversial.

The $\chi^2$ tests did not indicate major problems with heterogeneity in any of the eight analyses, but these were limited by low power. The $F$ results also indicated no observed heterogeneity. The number of available studies, and their size, does limit this review and prevents its findings from being conclusive. It is perhaps surprising that so few published trials exist for such a common practice. This may be partially attributable to the general lack of research on rehabilitation in orthopaedic surgery patients after discharge, rather than knee arthroplasty patients as such, as we also found few existing trials investigating exercise and rehabilitation after elective hip arthroplasty.

**Clinical implications**

Presently, given the reduction in length of hospital stay, compressed inpatient rehabilitation, and the limitations of the available evidence, it seems reasonable to refer patients for a short course of physiotherapy after discharge to provide short term benefit. While range of motion may be limited as an outcome measure of physiotherapy, the small to moderate standardised effect size obtained for function, which favours the intervention, is considered clinically important. This reflects actual improvements in one or more important aspects of function reported by patients after they received the treatment intervention. The type of physiotherapy provided also needs consideration. In the short term physiotherapy exercise interventions with exercises based on functional activities may be more effective after total knee arthroplasty than traditional exercise programmes, which concentrate on isotonic muscle exercises and exercises to increase range of motion in the joint.

**Future directions**

Although there were few studies and they were not large, they are still likely to have detected most worthwhile effects. These tentative findings suggest that further research would be worthwhile to reduce the current level of uncertainty. There seemed to be no benefits related to treatment at one year, though the evidence is not conclusive. The content of the intervention could be better designed and further tested. Interventions to date have largely consisted of exercise programmes and gait rehabilitation, mainly targeting impairment and helping patients to recover from the effects of surgery rather than specifically targeting limitations in activity or restrictions in participation. From the wider field of rehabilitation as a whole, however, such task training seems highly relevant. A recent systematic review, which assessed physiotherapy on functional outcome after stroke, found that effective studies contained focused exercise programmes within which the relevant functional tasks were directly trained. Research is currently underway to determine whether a brief feasible physiotherapy intervention of this type, supplied after discharge, affects patient’s functional ability one year...
WHAT IS ALREADY KNOWN ON THIS TOPIC

Osteoarthritis is the commonest cause of disability in older people, and total knee joint arthroplasty is a common orthopaedic procedure. Uncertainty exists regarding whether physiotherapy after discharge should be routinely provided to patients after elective primary knee arthroplasty for osteoarthritis.

WHAT THIS STUDY ADDS

Functional physiotherapy exercise soon after discharge results in short term benefit after elective primary knee arthroplasty. No benefit was seen at one year.

after knee arthroplasty. An investigation into the health economics is also included.

We thank Robert Bourne, David Beverland, P Codine, Helen Frost, Patricia Humphreys, John Kramer, and Brian Mockford for providing additional data for the review. Mike Clarke and students on the “Systematic Reviews” Module, May 2005, University of Oxford Department for Continuing Education, commented on the design of the review during its planning.

Contributors: CJML designed the review, undertook the review searches, screened trials for eligibility, assessed the quality of the trials, assisted with data analysis, and drafted the paper. She is guarantor. KLB supervised the review, assessed the quality of trials, and reviewed the draft paper. MD designed and undertook the meta-analyses for the review and reviewed the draft paper. CMS supervised the review, screened trials for eligibility, and cowrote the paper. Funding: CJML is funded by a nursing and allied health professional researcher development award from the Department of Health and NHS research and development. CMS is funded by a primary care career scientist award from the Department of Health and NHS research and development.

Competing interests: None declared.

Ethical approval: Oxford local research ethics committee (AQREC No AO3.018).

Provenance and peer review: Not commissioned; peer reviewed.


Accepted: 8 August 2007
Effect of prolonged and exclusive breast feeding on risk of allergy and asthma: cluster randomised trial

Michael S Kramer, James McGill professor; and scientific director,1 Lidia Matush, chief, 2 Irina Vanilovich, senior scientist,3 Robert Platt, associate professor; and investigator and associate director,4 Natalia Bogdanovich, senior scientist,3 Zinaida Sevkovskaya, senior scientist,1 Irina Dzikovich, senior scientist,5 Gyorgy Shishko, director,5 Bruce Mazer,, associate professor; and head,6 for the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group

ABSTRACT

Objective To assess whether exclusive and prolonged breast feeding reduces the risk of childhood asthma and allergy by age 6.5 years.

Design Cluster randomised trial.

Setting 31 Belarussian maternity hospitals and their affiliated policlincs.

Participants A total of 17 046 mother-infant pairs were enrolled, of whom 13 889 (81.5%) were followed up at age 6.5 years.

Intervention Breastfeeding promotion intervention modelled on the WHO/UNICEF baby friendly hospital initiative.

Main outcome measures International study of asthma and allergies in childhood (ISAAC) questionnaire and skin prick tests of five inhalant antigens.

Results The experimental intervention led to a large increase in exclusive breast feeding at 3 months (44.3% v 6.4%; P<0.001) and a significantly higher prevalence of any breast feeding at all ages up to and including 12 months. The experimental group had no reduction in risk or even an increase in risk with breast feeding.13-20

Conclusions These results do not support a protective effect of prolonged and exclusive breast feeding on asthma or allergy.

Trial registration Current Controlled Trials ISRCTN37687716.

INTRODUCTION

Whether breast feeding protects against the development of allergy and asthma has been frequently studied and hotly debated for more than 70 years.1-8 Research findings indicating a beneficial effect have been most consistent for atopic eczema during infancy, but the evidence on asthma and other atopic outcomes (including hay fever, food allergies, and positive skin tests) has been far more mixed. Some studies have reported greater degrees of protection with more exclusive and prolonged breast feeding,9-12 and several have noted a larger protective effect in children prone to atopy.13-17 Other studies, however, have reported no reduction in risk or even an increase in risk with breast feeding.18-20

A variety of methodological problems are likely to have affected the evidence base on infant feeding and allergic disease. Virtually all of the evidence is based on observational studies. Case-control studies are prone to recall bias, which may be systematic if the people who ascertain the previous feeding history are aware of the case versus control status of the study participants. Cohort studies, on the other hand, are prone to biased assessment of outcomes when observers are not blinded to the previous infant feeding history, are not showing that breast feeding is protective). Finally, publication bias may have affected the evidence base, with selection of studies with positive findings (that is, showing that breast feeding is protective).

A variety of methodological problems are likely to have affected the evidence base on infant feeding and allergic disease. Virtually all of the evidence is based on observational studies. Case-control studies are prone to recall bias, which may be systematic if the people who ascertain the previous feeding history are aware of the case versus control status of the study participants. Cohort studies, on the other hand, are prone to biased assessment of outcomes when observers are not blinded to the previous infant feeding history, are not showing that breast feeding is protective). Finally, publication bias may have affected the evidence base, with selection of studies with positive findings (that is, showing that breast feeding is protective).

One potential solution to these methodological problems is a randomised controlled trial. Although randomising healthy mothers and infants to breast feeding versus formula feeding is indefensible and probably unethical, randomising them to an intervention that
Table 1 | Baseline characteristics of children followed up at age 6.5 years in experimental and control groups. Values are numbers (percentages) unless stated otherwise

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental group (n=7108)</th>
<th>Control group (n=6781)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (years):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>1017 (14.3)</td>
<td>897 (13.2)</td>
</tr>
<tr>
<td>20-34</td>
<td>5783 (81.4)</td>
<td>5598 (82.6)</td>
</tr>
<tr>
<td>≥35</td>
<td>308 (4.3)</td>
<td>286 (4.2)</td>
</tr>
<tr>
<td>Maternal education:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete secondary</td>
<td>310 (4.4)</td>
<td>292 (3.8)</td>
</tr>
<tr>
<td>Complete secondary</td>
<td>2441 (34.3)</td>
<td>2011 (29.7)</td>
</tr>
<tr>
<td>Advanced secondary/partial university</td>
<td>3395 (47.6)</td>
<td>3693 (54.5)</td>
</tr>
<tr>
<td>Complete university</td>
<td>962 (13.5)</td>
<td>875 (12.9)</td>
</tr>
<tr>
<td>Older children living in household:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4176 (58.8)</td>
<td>3695 (54.5)</td>
</tr>
<tr>
<td>1</td>
<td>2368 (33.3)</td>
<td>2450 (36.1)</td>
</tr>
<tr>
<td>≥2</td>
<td>564 (7.9)</td>
<td>636 (9.4)</td>
</tr>
<tr>
<td>Positive family history of atopy</td>
<td>359 (5.1)</td>
<td>226 (3.3)</td>
</tr>
<tr>
<td>Maternal smoking during pregnancy</td>
<td>184 (2.6)</td>
<td>108 (1.6)</td>
</tr>
<tr>
<td>Male child</td>
<td>3653 (51.4)</td>
<td>3528 (52.0)</td>
</tr>
<tr>
<td>Mean (SD) birth weight (g)</td>
<td>3440 (418)</td>
<td>3441 (423)</td>
</tr>
</tbody>
</table>

METHODS

The detailed methods of PROBIT and the results during the first year of follow-up have been previously reported.22 The units (clusters) of randomisation were maternity hospitals and one affiliated polyclinic (outpatient clinic where children are followed for well child and illness care) for each hospital, with double randomisation based on both a random numbers table and a coin flip.22 The experimental intervention was based on the baby friendly hospital initiative, which was developed by the World Health Organization (WHO) and the United Nations Children’s Fund (UNICEF) to promote and support breast feeding, particularly among mothers who have chosen to start breast feeding.23 The control maternity hospitals and polyclinics continued the practices and policies in effect at the time of randomisation. The trial results are based on a total of 17 046 healthy breastfed infants from 31 maternity hospitals/polyclinics; all were born at term in 1996–7, weighed at least 2500 g, and were enrolled during their postpartum stay.22 To our knowledge, PROBIT is the largest randomised trial ever done in the area of human lactation. It conforms to the CONSORT recommendations for the design, analysis, and reporting of cluster randomised trials.24

As previously reported,22 the two randomised groups were similar in baseline sociodemographic and clinical variables, including maternal age, education, number of other children at home, proportion who had breast fed a previous child for at least three months, family history of atopy, caesarean delivery, maternal smoking during pregnancy, birth weight, gestational age, and five minute Apgar score. The experimental intervention led to a substantial difference in the duration of any breast feeding that was maintained throughout the first year of follow-up: 72.7% versus 60.0% were still breast feeding at 3 months, 49.8% versus 36.1% at 6 months, 36.1% versus 24.4% at 9 months, and 19.7% versus 11.4% at 12 months in the experimental and control groups. In addition, the prevalence of exclusive breast feeding was sevenfold higher in the experimental group at 3 months (43.3% vs 6.4%; P<0.001), although low in both groups at 6 months (7.9% vs 0.6%; P=0.01).22

One paediatrician in each of 24 of the 31 polyclinics did follow-up interviews and examinations at age 6.5 years from December 2002 to April 2005; in the remaining seven high volume clinics, follow-up visits were shared by two paediatricians. Allergy symptoms and diagnoses were ascertained with the international study of asthma and allergy in childhood (ISAAC) questionnaire, which had already been translated and validated by the ISAAC investigators.23 In addition, the paediatricians did skin prick tests to five antigens (Allergy Canada): house dust mite, cat, birch pollen, mixed northern grasses, and Alternaria. Saline was included as a negative control and histamine (1 mg/ml) as a positive control. The criteria for a positive result were a mean wheal ≥3 mm or flare ≥10 mm, calculated as the mean of the longest diameter and orthogonal diameter after subtracting the mean of these...
Table 3 | Positive skin prick test wheal remeasurements (n=119). Values are intraclass correlation coefficients (95% confidence intervals)

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Longest diameter</th>
<th>Orthogonal diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>House dust mite</td>
<td>0.96 (0.95 to 0.98)</td>
<td>0.96 (0.94 to 0.97)</td>
</tr>
<tr>
<td>Cat</td>
<td>0.93 (0.90 to 0.95)</td>
<td>0.86 (0.81 to 0.94)</td>
</tr>
<tr>
<td>Birch pollen</td>
<td>0.94 (0.91 to 0.95)</td>
<td>0.97 (0.96 to 0.98)</td>
</tr>
<tr>
<td>Mixed northern grasses</td>
<td>0.95 (0.92 to 0.96)</td>
<td>0.98 (0.98 to 0.99)</td>
</tr>
<tr>
<td>Alternaria</td>
<td>0.94 (0.91 to 0.96)</td>
<td>0.92 (0.89 to 0.95)</td>
</tr>
</tbody>
</table>

Table 4 | ISAAC results. Values are numbers (percentages) positive unless stated otherwise

<table>
<thead>
<tr>
<th>Question</th>
<th>Experimental group</th>
<th>Control group</th>
<th>ICC</th>
<th>Cluster adjusted odds ratio* (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever had wheezing</td>
<td>778 (11.0)</td>
<td>651 (9.6)</td>
<td>0.03</td>
<td>1.1 (0.6 to 1.8)</td>
</tr>
<tr>
<td>Wheezing in past 12 months</td>
<td>238 (3.4)</td>
<td>188 (2.8)</td>
<td>0.01</td>
<td>1.0 (0.7 to 1.6)</td>
</tr>
<tr>
<td>Ever had asthma</td>
<td>97 (1.4)</td>
<td>68 (1.0)</td>
<td>0.00</td>
<td>1.2 (0.7 to 1.9)</td>
</tr>
<tr>
<td>Ever had hay fever symptoms</td>
<td>384 (5.4)</td>
<td>257 (3.8)</td>
<td>0.02</td>
<td>1.1 (0.6 to 1.9)</td>
</tr>
<tr>
<td>Hay fever symptoms in past 12 months</td>
<td>262 (3.7)</td>
<td>192 (2.8)</td>
<td>0.01</td>
<td>1.0 (0.6 to 1.6)</td>
</tr>
<tr>
<td>Recurrent itchy rash</td>
<td>350 (4.9)</td>
<td>241 (3.6)</td>
<td>0.02</td>
<td>1.3 (0.7 to 2.2)</td>
</tr>
<tr>
<td>Ever had eczema</td>
<td>69 (1.0)</td>
<td>72 (1.1)</td>
<td>0.00</td>
<td>1.0 (0.5 to 1.8)</td>
</tr>
</tbody>
</table>

ICC=intraclass correlation coefficient.

*For experimental group versus control group.

RESULTS

A total of 13 889 children were seen in follow-up for PROBIT II, representing 81.5% of the 17 046 originally randomised. Of the 3157 children randomised but not followed up, 88 had died, 2938 were lost to follow-up, and 131 were unable or unwilling to come for their PROBIT II visit. Follow-up rates were similar in the experimental (80.2%) and control (82.9%) polyclinics but varied considerably by polyclinic—from 56.1% at one of the Minsk polyclinics to 94.6% at Klimovichi, a small rural based polyclinic. The mean age at follow-up was 6.6 (SD 0.3) years. As shown in table 1, the children followed up in the experimental and control groups were similar in baseline characteristics, with small differences paralleling those seen (and previously reported) at randomisation.

Table 2 shows the audit results. The data shown are the κ coefficients (and their 95% confidence intervals) between the results at the initial clinic visit and the results at the audit visit. These κ values are high with respect to wheezing and hay fever symptoms and moderate for reported diagnosis of asthma or symptoms and diagnosis of atopic eczema. Concordance was high for the skin prick test results, but only 54 (28%) of the 190 audited children (or their parents) agreed to the repeat skin tests. The results of the random skin prick test wheal remeasurements indicated exceptionally high intraclass correlation coefficients for both the longest diameter and orthogonal diameter for all five antigens tested (table 3).

Table 4 shows the trial results for the ISAAC questionnaire, the number and proportion of those with positive responses in the experimental and control groups, the intraclass correlation coefficients reflecting the degree of within polyclinic clustering, and the cluster adjusted odds ratios and their 95% confidence intervals. The results showed a very low degree of clustering—that is, very little tendency for mothers of children attending the same polyclinic to respond more similarly to each other than those of children attending...
Table 5 | Skin prick test results. Values are numbers (percentages) positive unless stated otherwise

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Experimental group (n=5551)</th>
<th>Control group (n=5595)</th>
<th>ICC</th>
<th>Cluster adjusted odds ratio* (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>House dust mite</td>
<td>805 (14.5)</td>
<td>603 (10.8)</td>
<td>0.14</td>
<td>1.1 (0.5 to 2.4)</td>
</tr>
<tr>
<td>Cat</td>
<td>648 (11.7)</td>
<td>491 (8.8)</td>
<td>0.20</td>
<td>1.2 (0.5 to 2.8)</td>
</tr>
<tr>
<td>Birch pollen</td>
<td>526 (9.5)</td>
<td>393 (7.0)</td>
<td>0.18</td>
<td>1.2 (0.5 to 2.9)</td>
</tr>
<tr>
<td>Mixed northern grasses</td>
<td>712 (12.6)</td>
<td>491 (8.8)</td>
<td>0.17</td>
<td>1.0 (0.5 to 2.3)</td>
</tr>
<tr>
<td>Alternaria</td>
<td>480 (8.6)</td>
<td>340 (6.1)</td>
<td>0.18</td>
<td>1.5 (0.5 to 4.4)</td>
</tr>
<tr>
<td>≥1 positive</td>
<td>1496 (27.0)</td>
<td>1013 (18.1)</td>
<td>0.19</td>
<td>1.2 (0.5 to 2.6)</td>
</tr>
</tbody>
</table>

ICC=intraclass correlation coefficient. *For experimental group versus control group.

Table 6 | Results of sensitivity analysis for skin prick test results. Values are numbers (percentages) positive unless stated otherwise

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Experimental group (n=4100)</th>
<th>Control group (n=4906)</th>
<th>ICC</th>
<th>Cluster adjusted odds ratio* (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>House dust mite</td>
<td>504 (12.3)</td>
<td>299 (6.1)</td>
<td>0.04</td>
<td>2.0 (1.2 to 3.4)</td>
</tr>
<tr>
<td>Cat</td>
<td>347 (8.5)</td>
<td>182 (3.7)</td>
<td>0.05</td>
<td>2.1 (1.1 to 3.9)</td>
</tr>
<tr>
<td>Birch pollen</td>
<td>273 (6.7)</td>
<td>125 (2.5)</td>
<td>0.03</td>
<td>2.3 (1.3 to 4.1)</td>
</tr>
<tr>
<td>Mixed northern grasses</td>
<td>369 (9.0)</td>
<td>209 (4.3)</td>
<td>0.06</td>
<td>1.5 (0.8 to 2.8)</td>
</tr>
<tr>
<td>Alternaria</td>
<td>258 (6.3)</td>
<td>77 (1.6)</td>
<td>0.05</td>
<td>3.5 (1.6 to 7.7)</td>
</tr>
<tr>
<td>≥1 positive</td>
<td>929 (22.7)</td>
<td>579 (11.8)</td>
<td>0.07</td>
<td>2.0 (1.1 to 3.4)</td>
</tr>
</tbody>
</table>

ICC=intraclass correlation coefficient. *For experimental group versus control group.

The results from this large cluster randomised trial indicate that the experimental intervention to promote breast feeding did not reduce the risk of asthma, hay fever, or eczema at age 6.5 years despite large increases in the duration and exclusivity of breast feeding; nor did the intervention succeed in reducing the prevalence of positive skin prick tests. We observed high inter-paediatrician variability (and consequently wide confidence intervals around the adjusted odds ratios) in results of skin prick tests. After exclusion of six polyclinics with suspiciously high rates of positive skin prick test results, the relative odds of positive skin prick test, nor were there any problems detected during monitoring visits carried out soon after data collection had begun. Redefining positive skin prick tests on the sole basis of the size of the wheal did not reduce the high positivity rates in the six discrepant clinics, nor did it uncover any protective effects of the experimental intervention (data available on request).

We did a sensitivity analysis (n=9006) after excluding the six polyclinic sites with high rates of positive skin prick test results (three of these polyclinics were in the experimental group and three were in the control group). Table 6 shows the results of the sensitivity analysis. Intraclass clustering of skin prick test results was reduced substantially. The proportions of positive test results were considerably lower in both the experimental and control groups, yet the differences between the two groups increased, with significantly elevated odds ratios in the 2-3 range for all but one antigen (mixed northern grasses).

Statistical models with interaction terms showed no evidence that the overall effects of the intervention on allergic symptoms and diagnoses were modified by family history of atopy, except for a history of ever wheezing, for which the experimental intervention increased the risk to a greater extent among those with a positive family history (P<0.005). For the skin prick tests, we found a significant interaction in the same direction (greater risk increase in those with a positive family history) only for mixed northern grasses (P=0.02).

Finally, we also examined the observational associations between the duration and exclusivity of breast feeding and the prevalence of allergic symptoms and diagnoses and skin prick test results. For allergic symptoms and diagnoses, we found borderline significant reductions in history of eczema both with more prolonged any breast feeding and with more prolonged exclusive breast feeding (P=0.08 for both associations, based on χ² tests for trend). Associations were stronger and in the opposite direction for skin prick test results, especially after elimination of the six sites with high rates of positive tests; we found highly significant increases in positive skin prick test results with exclusive breast feeding for 3 to <6 months and ≥6 months versus <3 months for house dust mite, cat, birth pollen, mixed northern grasses, and Alternaria (P<0.001 for all five antigens, based on χ² tests for trend).

**DISCUSSION**

The results from this large cluster randomised trial indicate that the experimental intervention to promote breast feeding did not reduce the risk of asthma, hay fever, or eczema at age 6.5 years despite large increases in the duration and exclusivity of breast feeding; nor did the intervention succeed in reducing the prevalence of positive skin prick tests. We observed high inter-paediatrician variability (and consequently wide confidence intervals around the adjusted odds ratios) in results of skin prick tests. After exclusion of six polyclinics with suspiciously high rates of positive skin prick test results, the relative odds of positive skin prick test, nor were any problems detected during monitoring visits carried out soon after data collection had begun. Redefining positive skin prick tests on the sole basis of the size of the wheal did not reduce the high positivity rates in the six discrepant clinics, nor did it uncover any protective effects of the experimental intervention (data available on request).

We did a sensitivity analysis (n=9006) after excluding the six polyclinic sites with high rates of positive skin prick test results (three of these polyclinics were in the experimental group and three were in the control group). Table 6 shows the results of the sensitivity analysis. Intraclass clustering of skin prick test results was reduced substantially. The proportions of positive test results were considerably lower in both the experimental and control groups, yet the differences between the two groups increased, with significantly elevated odds ratios in the 2-3 range for all but one antigen (mixed northern grasses).

Statistical models with interaction terms showed no evidence that the overall effects of the intervention on allergic symptoms and diagnoses were modified by family history of atopy, except for a history of ever wheezing, for which the experimental intervention increased the risk to a greater extent among those with a positive family history (P<0.005). For the skin prick tests, we found a significant interaction in the same direction (greater risk increase in those with a positive family history) only for mixed northern grasses (P=0.02).

Finally, we also examined the observational associations between the duration and exclusivity of breast feeding and the prevalence of allergic symptoms and diagnoses and skin prick test results. For allergic symptoms and diagnoses, we found borderline significant reductions in history of eczema both with more prolonged any breast feeding and with more prolonged exclusive breast feeding (P=0.08 for both associations, based on χ² tests for trend). Associations were stronger and in the opposite direction for skin prick test results, especially after elimination of the six sites with high rates of positive tests; we found highly significant increases in positive skin prick test results with exclusive breast feeding for 3 to <6 months and ≥6 months versus <3 months for house dust mite, cat, birth pollen, mixed northern grasses, and Alternaria (P<0.001 for all five antigens, based on χ² tests for trend).
Evidence is conflicting as to whether prolonged and exclusive breast feeding increases, decreases, or has no effect on the risks of asthma and allergy. All of the available evidence is based on observational studies.

**WHAT THIS STUDY ADDS**

Prolonged and exclusive breast feeding had no protective effect on allergic symptoms and diagnoses or on positive skin prick tests.

The device used for administration, the potency of the extracts, the spacing between the antigens, the position of the arm used for placement, and even the time of day, affect the results. All testers were trained by the same investigator (BM), all centres were provided with the same testing equipment, and the extracts were prepared from the same lots for all participants. Other possible explanations include failure to wipe the testing needle after histamine or antigens, contamination of the extract preparation with histamine, trauma with the skin prick test needle, and placement of the test antigens too closely together. The small number (n=54) of children with repeat skin tests at the audit, and the auditing paediatricians’ use of the polyclinics’ antigen solutions, prevented use of the audit data to confirm or refute these possibilities.

Given that we found significantly increased risks of positive skin prick tests in the experimental group only after excluding the six suspect polyclinics, we cannot be confident that the experimental intervention actually caused the increased risks. We feel on safer ground in inferring no reduction in risk. Given these results based on a large randomised trial and the inconsistent benefits reported in previous studies, public health measures to increase the initiation, duration, and exclusivity of breast feeding seem unlikely to have a major impact on reducing the incidence of atopic diseases. The fact that most atopic outcomes have increased in incidence over the past several decades, simultaneous with the renaissance in breast feeding, strongly suggests that breast feeding does not have a potent protective effect at the population level. Thus, our results underline the importance of seeking other explanations for the recent epidemic of allergy and asthma and of investigating other potential causative factors to develop and test new preventive interventions.

**Contributors:** MSK, RP, and BM contributed to obtaining funding for this project and to the design, analysis, interpretation, writing, and revision of the manuscript. LM, IV, NB, ZS, ID, and GS contributed both to the design of the study and to the planning, implementation, and monitoring of the field work in Belarus. MSK is the guarantor.

**Funding:** Grant (MOP No 53155) from the Canadian Institutes of Health Research. The funder had no role in study design, data collection, data analysis, data interpretation, or writing of this manuscript.

**Competing interests:** None declared.

**Ethical approval:** The research ethics board of the Montreal Children’s Hospital of the McGill University Health Centre approved this project (including the 6.5-year follow-up).

**Provenance and peer review:** Non-commissioned; externally peer reviewed.

4. Gdalevich M, Mimouni D, David M, Mimouni M. Breast-feeding and the onset of atopic dermatitis in childhood: a systematic review and...


Accepted: 27 July 2007
Uterine prolapse is the herniation of the uterus into or beyond the vagina as a result of failure of the ligamentous and fascial supports. It often coexists with prolapse of the vaginal walls, involving the bladder or rectum. In the United Kingdom, the disorder accounts for 20% of women waiting for major gynaecological surgery.1

Why should I read this article?
Most women with symptoms of prolapse will present to primary care, and initial assessment and treatment occurs here.2 An understanding of the pathophysiology, assessment, and management of prolapse is essential for the primary care team to streamline appropriate referrals to hospital. This article aims to cover these topics and to provide an overview of current management of secondary care.

How common is uterine prolapse?
The exact prevalence is unknown. Forty per cent of participants in the women’s health initiative (WHI) trial in the United States had some degree of prolapse. Uterine prolapse was found in 14% of the 27,342 women enrolled in the study.3 Another US study of 149,554 women found an 11% lifetime risk of surgery for prolapse or incontinence in the United States.4 The Oxford Family Planning Association study in the United Kingdom followed more than 17,000 women aged 25-39.5 The annual incidence of hospital admission with prolapse was 20.4/10,000, and the annual incidence of surgery for prolapse was 16.2/10,000. Many studies do not distinguish between prolapse of all pelvic organs and prolapse of the uterus alone, which makes it difficult to determine the true incidence.

Four hundred and twelve women originally enrolled in the WHI study were followed up to assess progression of prolapse. Spontaneous regression was common, especially for grade 1 prolapse— the progression rate was 1.9/100 women years and the regression rate was 48/100 women years.6 Thus, prolapse is not always progressive.

Why does prolapse occur?
Anatomy
Some knowledge of normal vaginal support is needed to understand the pathophysiology of pelvic organ prolapse. Delancey’s three levels of support (figure) are now accepted worldwide.7 Level 1: The cardinal-uterosacral ligament complex provides apical attachment of the uterus and vaginal vault to the bony sacrum. Uterine prolapse occurs when the cardinal-uterosacral ligament complex breaks or is attenuated. Level 2: The arcus tendineous fascia pelvis and the fascia overlying the levator ani muscles provide support to the middle part of the vagina. Level 3: The urogenital diaphragm and the perineal body provide support to the lower part of the vagina.

Risk factors
The aetiology of pelvic organ prolapse is multifactorial (box 1). The pelvic organ support study found age to be a risk factor for pelvic organ prolapse—risk doubled with each decade of life.9 Increasing parity was also associated with increasing severity of prolapse. Of the 17,000 women in the Oxford family planning study, those with a history of two vaginal deliveries were 8.4 times more likely to have surgery for prolapse than those with no such history.5

Although vaginal delivery is clearly associated, specific obstetric risk factors remain controversial. Macrosomia, prolonged second stage of labour, episiotomy, anal sphincter injury, epidural analgesia, and the use of forceps and oxytocin have
all been proposed as risk factors but have not been proved.

Women who are overweight (body mass index 25-30; odds ratio 2.51, 95% confidence interval 1.18 to 5.35) or obese (>30; 2.56, 1.23 to 5.35) are at high risk of developing prolapse.\(^\text{10}\) Heritable or genetic factors might play a part. In a case control study of 108 women with and without prolapse, a higher risk of prolapse was seen in women with a mother (3.2, 1.1 to 7.6) or sister (2.4, 1.0 to 5.6) reporting prolapse.\(^\text{11}\)

Although menopause is often cited as a risk factor for pelvic organ prolapse, a study of 270 women from the WHI trial who had undergone hysterectomy found no association between oestrogen status (use of hormone replacement therapy) and prolapse.\(^\text{12}\)

**What are the symptoms?**

Many symptoms have been attributed to prolapse (box 2), although none of them are specific, except for seeing or feeling a vaginal bulge. A study of 497 women in the US under annual review for prolapse\(^\text{10}\) showed that the number of symptoms and the problems these caused increased as the stage of prolapse increased from stage 0 (no prolapse) to stage III (prolapse outside the vagina) (box 3).

**Examining the patient**

A pelvic examination should be done (using a Sim’s single bladed speculum) to define the extent of the prolapse and establish the compartments of the vagina affected (anterior, posterior, or apical). The patient should be at rest and straining during a Valsalva manoeuvre. The oestrogen status of the tissues (signs of vaginal atrophy) and the size and mobility of the uterus and adnexae should be assessed.
Several prolapse grading systems exist, but the only system that has been robustly tested for both inter-observer and intraobserver reliability is the pelvic organ prolapse quantification system. This system defines the extent of prolapse by measuring the descent of anterior, posterior, and apical segments of the vaginal wall relative to the vaginal hymen. A full description is beyond the scope of our review, but the prolapse is minimal (stage I). Some women may prefer observation for advanced prolapse—they should be examined periodically to look for development of new symptoms or disorders (such as obstructed urination or defecation, vaginal erosion).

### Management

#### Observation

The extent of the prolapse does not correlate well with the symptoms. Watchful waiting is most appropriate if the prolapse is minimal (stage I). Some women may prefer observation for advanced prolapse—they should be examined periodically to look for development of new symptoms or disorders (such as obstructed urination or defecation, vaginal erosion).

#### Conservative treatment

**Pelvic floor muscle training**

Pelvic floor muscle training is an effective treatment for urinary incontinence, but its role in managing prolapse is unclear. A Cochrane review of conservative management of uterine prolapse published in 2006 concluded that there was no evidence from randomised trials and that further trials were needed. A feasibility study for the pelvic organ prolapse physiotherapy study—which will evaluate the effectiveness of pelvic floor muscle training in treating pelvic prolapse—has since been completed, and a fully powered randomised follow on trial in 16 global centres is due to start in 2007.

**Pessaries**

Vaginal pessaries are the only currently available non-surgical intervention for managing women with prolapse (box 4).

Although evidence to support the use of pessaries is not robust, they are used by 86% of gynaecologists and 98% of urogynaecologists. In a prospective study of 100 consecutive women with symptomatic pelvic organ prolapse fitted with a pessary, 73 women retained the pessary two weeks later. After two months, 92% of these women were satisfied with the pessary; virtually all symptoms of prolapse and 50% of urinary symptoms had resolved, although occult stress incontinence was unmasked in 21% of the women.

We found just one small prospective study that looked at whether pessaries can alter the natural history of prolapse. Fifty six women were prospectively evaluated using the pelvic organ prolapse quantification system. All women had a pessary fitted for at least one year. Of the 19 women who continued to use pessaries, the stage of the prolapse improved in four.

In 2004, a Cochrane review found no randomised trials of pessary use in women with prolapse. A feasibility study for a randomised trial of the use of vaginal pessaries combined with pelvic floor exercises is due to start in 2007.

### Surgical treatment

In England and Wales in 2005–2006, 22 274 operations were performed for “vaginal prolapse.” The literature reports outcomes from surgery for uterine prolapse alone, and in conjunction with vaginal prolapse repairs at the same time. Many recent papers are confounded by heterogeneity of the patients studied, and a considerable proportion of patients had continence surgery with suburethral tape procedures at the same time. Hysterectomy for uterine prolapse...
prolapse can be performed via the abdominal route or vaginal route, although vaginal hysterectomy is preferred in the UK. In a study of all hysterectomies performed in the UK during 1993 and 1994, one third were done vaginally, and 95% of these were for prolapse.24

The greatest challenge in surgery for uterine prolapse is to prevent subsequent prolapse of either the vault or anterior or posterior walls of the vagina. Hysterectomy alone fails to correct the loss of integrity of the cardinal-uterosacral ligament complex and weakening of the pelvic diaphragm. A variety of procedures are available to support the vaginal vault at the time of hysterectomy. These include the vaginal procedures McCall culdoplasty; plication of the uterosacral ligament; sacrospinous or prespinous fixation for vaginal vault prolapse; and sacrocolpopexy (performed via an open procedure or laparoscopically). A retrospective case control study compared 62 women having sacrospinous fixation with 62 women having McCall culdoplasty at the time of vaginal hysterectomy. It found that women who had McCall culdoplasty had fewer recurrences (15% vs 27%).25

We found only one published Cochrane review that looked at surgery for all types of prolapse. The meta-analysis of trials of vault suspension procedures showed that abdominal sacrocolpopexy was associated with a lower recurrence of vault prolapse and less dyspareunia than vaginal sacrospinous colpopexy. However, too few data on subjective success rate, patient satisfaction, and effect on quality of life were available to make reliable conclusions.26

In the past five years, several randomised and non-randomised studies have assessed the efficacy of various procedures with and without artificial meshes for preventing vault prolapse after hysterectomy. At present, however, even after meta-analysis the data have limited ability to inform decision making, particularly with respect to long term efficacy and the effect on sexual function. Large randomised trials to look at long term cure and complications after surgery for uterine prolapse are urgently needed.

What do we still need to know?

Many aspects of the cause, incidence, prevalence, and natural history of pelvic organ prolapse are unclear. No consensus or evidence exist about whether and how to treat women with prolapse. We do not know whether we should intervene in the absence of symptoms, despite anatomical changes. We also do not know the optimal timing for intervention, or whether early intervention reduces the incidence of recurrence.

We do not know whether conservative management can prevent or delay the need for surgery, and the best surgical approach to achieve anatomical cure, resolution of symptoms, and low rates of recurrence in not known. Large prospective randomised studies with long term follow-up are needed to answer all these questions.

Contributors: AD searched the literature, obtained the primary papers, and drafted the paper. RECT and CJM reviewed and revised the draft manuscript and approved the final version. DGT discussed the search strategy, reviewed the papers, contributed to and revised the draft manuscript, and approved the final version. DGT is guarantor.

Competing interests: DGT currently sits on the advisory board of clinical studies funded by Eli Lilly and Company and Johnson and Johnson Medical. Consultancy payments for these studies are managed by the University of Leicester research and business office and are used to support his research. DGT is also principal investigator on three investigator initiated studies funded by grants from Johnson and Johnson, Astellas Pharma, and UCB Pharma. In 2006, he received grants towards attending international scientific meetings from American Medical Systems, Astellas, Pfizer, UCB Pharma, and Janssen Cilag. CJM has provided surgical training in the use of the TVT and TVT-O.

Box 4 | Key points for fitting pessaries and their subsequent management

Fitting
• Ensure that the patient’s bladder and bowel are empty
• The pessary fits well if a finger can be swept between the pessary and the walls of the vagina
• The goal is to fit the largest pessary that does not cause discomfort
• Ask the patient to walk around, bend, and micturate to ensure that the pessary is retained

Management
• No consensus exists on how frequently to see a patient after a pessary is fitted successfully
• At each follow-up visit ask the patient about any new symptoms and inspect for erosions, ulcers, or discharge
• Fitting a pessary can unmask symptoms of urinary incontinence

Additional educational resources

Resources for health professionals
Conservative management of pelvic organ prolapse in women. www.cochrane.org/reviews/en/ab003882.html
Mechanical devices for pelvic organ prolapse in women. www.cochrane.org/reviews/en/ab004010.html
Surgical management of pelvic organ prolapse in women. www.cochrane.org/reviews/en/ab004014.html

Resources for patients
NHS Direct http://cks.library.nhs.uk/patient_information_leaflet/prolapse_of_the_uterus—Information on causes, symptoms, and treatment of prolapse of the uterus
Proplase of the Pelvic Organ. www.womenshealthlondon.org.uk/leaflets/prolapse/prolapse.html—Detailed information on uterine prolapse
JAMA http://jama.ama-assn.org/cgi/reprint/293/16/2054.pdf—Patient page on uterine prolapse
SUMMARY POINTS

Uterine prolapse can occur at the same time as prolapse of the anterior or posterior vaginal compartments.

Little is known about the prevalence and natural progression of prolapse.

Initially, patients should be assessed and managed conservatively in primary care.

Conservative management is advised for patients who are not fit for surgery or do not want surgery.

Surgical treatment for uterine prolapse should incorporate procedures to prevent recurrence.

Reliable evidence for both conservative and surgical treatment options is lacking, but randomised trials are under way.

Body heat

Last year, I accompanied a group to Everest Base Camp as the team doctor. My job was to tackle both common medical ailments and any problems related to altitude sickness. We all arrived safely, and we were making our way back from Pheriche (4100 metres) altitude when big black clouds formed out of nowhere and it started to snow heavily. With the prospect of hot chocolate and a warm fire at the end of the day and our down jackets keeping us warm, we fought on against the chilling winds and the blinding snow.

Finally, we reached the place where we were staying for the night and lost no time in huddling around a blazing fire.

But then one of our group, a 26 year woman, started shivering severely. Her hands and feet were extremely cold and numb. We brought her near to the flames, piled blankets on her, and vigorously rubbed her hands and feet. These slowly turned pink, but she felt no better.

We added more blankets, threw more wood on the fire, and gave her a hot water bag to hold close against her body. However, despite all our endeavours, nothing was working.

As our companion continued to shiver uncontrollably, we added more blankets, threw more wood on the fire, and gave her a hot water bag to hold close against her body. However, despite all our endeavours, nothing was working.

As our companion continued to shiver uncontrollably, we added more blankets, threw more wood on the fire, and gave her a hot water bag to hold close against her body. However, despite all our endeavours, nothing was working.
Acromegaly

Jon Danzig

Acromegaly may be a rare condition and difficult to diagnose. This patient’s story highlights the importance of inquisitive questioning in cases with no obvious diagnosis.

Something strange started happening to me around 10 years ago. My shoe size grew from 9 to 11, my income permanently dried up, and the girlfriend I was going to marry left after we stopped having sex. It wasn’t just the sex, or lack of it. I also dramatically changed, both personally and physically. My nose grew bigger, my forehead enlarged, and my face and hands became puffy and coarse. But these were all gradual changes that I only realised in retrospect.

From 1997 to the summer of 2001, I saw all manner of medical experts. My general practitioner collected almost 100 pages of notes about me. The doctors found nothing fundamentally wrong. I came to think it was me. I must have just “lost it.” My general practitioner sent me to the practice counsellor. I told him I was worried that I had stopped being creative. He replied that I’d got to an age when I couldn’t expect to be creative any more. I was only 42.

An opportunity to discover what was wrong came and passed. A small lump rapidly grew in my right breast—gynaecomastia. “Stop taking vitamin pills,” said one consultant. “Have surgery,” said another. Neither doctor, though, arranged hormone tests to determine the cause. A fortnight before surgery, the lump disappeared. Later it reappeared in my left breast. Ignore it, advised a general practitioner.

My dermatologist missed the diagnosis by not discovering the true cause of chronic mountainous acne on my back that was extremely painful. My orthopaedic doctor lost an opportunity by not finding the reason for joint pains. My dentist missed a chance by not recognising the widening gaps in my lower front teeth and problems with my lower jaw overbite. The doctors in my general practice failed by not understanding the combination of all my complaints, including depression, apathy, and snoring. When my testosterone was discovered to be substantially below normal, one general practitioner told me to come back in 12 months.

So I struggled for years, undiagnosed and with symptoms that stopped me from earning a living. Even when at last I was referred to an endocrinologist, he was also unable to make a timely diagnosis. It was only by becoming a more informed patient that I eventually added up the sum of all my symptoms and spelt out acromegaly. My suspicion sent me into a deep shock.

My endocrinologist failed to confirm a diagnosis of acromegaly in the six months I was his patient. Instead, he proposed a “non-functioning tumour.” Only belatedly was I sent for an oral glucose tolerance test, the standard marker for acromegaly. As the endocrinologist then went on holiday, I asked my general practitioner’s receptionist to obtain the results for me but was met with reluctance. I became exasperated and said, “This could be a life threatening illness.”

A general practitioner rang back: “Your growth hormone levels are completely normal, and you do not have any insufficiency. This is hardly a life threatening illness.” I remember feeling disappointed. Maybe my suspicion was completely wrong. “I thought my growth hormone level might have been too high,” I said. I was sweating and my heart was beating fast. “Well,” said the doctor, “as part of the test, you had taken a growth hormone accelerator, so I expect the
The diagnosis at last provided an explanation for all my strange and incapacitating symptoms, which was a huge relief. I also felt let down. Surely any of my doctors should have been able to recognise what was wrong with me much earlier? Apparently, the typical delay in diagnosis for acromegaly is 10 years or more. Isn’t this a failure of medicine? The challenge for doctors, and patients, is to catch the illness early, years before the extensive damage and deformities occur.

Naturally, I wanted a cure. Yet I knew that with a large tumour my chances were low. The position of the tumour also caused concern as it was dangerously close to the carotid artery. This made me determined to find the most experienced neurosurgeon. That led me to Professor Rudolf Fahlbusch, the then director of neurosurgery at the University of Erlangen-Nuremberg in Germany. He had performed over 4000 pituitary operations, probably more than anyone else in the world. I was right to seek the best. I needed two operations, and afterwards Professor Fahlbusch wrote that the trans-sphenoidal surgery on me was “one of the most difficult and risky of my surgical life.”

After the operations my face started going back to normal and my testosterone levels increased. Later, prolactin and thyroid levels returned to normal. My sleep improved and the out of body experiences stopped. The cystic acne on my back disappeared. Most importantly, the growth hormone levels also went down considerably but, alas, not enough. Blood tests showed I still had acromegaly.

Eventually, I sought the help of a leading acromegaly expert, Professor John Wass at the Churchill Hospital in Oxford. He fought hard for me to be treated on the NHS with octreotide (Sandostatin Lar, Novartis) injections, costing almost £1000 a shot, which I now receive every six weeks. My growth hormone levels are now at last in the normal range.

I feel I’ve made considerable progress. Unfortunately, though, I am not yet back to my work. Although I have my good days, I still get easily damaged, not all of which can be undone. Nevertheless, I’m an optimist and remain hopeful that I may eventually be able to resume my career and life’s plans in general.

My relationship is greatly improved with my general practitioner, Stephen Cohen. We’ve been on a long journey together and learnt considerably from it. From a difficult start, ours is now a true doctor-patient partnership. Dr Cohen has become my strongest ally through the medical maze and has helped to restore my confidence in the profession. Recently he told me, “I’ve come to the conclusion that any doctor who’s not prepared to learn may as well retire.” All the doctors at the practice have met to try and learn from my case. That, I applaud.

**Doctor’s perspective**

Jon Danzig has been one of the most challenging patients in my 32 years of general practice. I was initially frustrated, and probably iritated, with his frequent attendances for what seemed at that time to be a disparate collection of symptoms. I never seemed to be able to find any treatable cause for any of them, neither could any of the various specialists to whom I referred him. I reserved a drawer in my consulting room just for the paperwork that he generated—both from his own researches and the endless stream of hospital letters.

After the penny finally dropped and acromegaly was diagnosed, Jon said: “If only you had asked me if my shoe size had changed.” Well, of course we can all be wiser after the event, and because I saw him so regularly, I failed to notice the asymmetrical enlarging of his jaw—unlike his friend who saw him for the first time in years.

I have come to respect Jon’s quiet persistence in pursuing his quest for a diagnosis. Even since the diagnosis, he has had to challenge the rigidly held views of experts, getting them, as well, to listen to his observations and the results of his own researches. He has taught me the value of listening to my “persistent” patients more carefully, not just of giving a weary sigh on seeing their familiar name on my appointment list with the consequent shutting off of my receptive faculties.

Stephen Cohen, senior partner at The Nap Surgery, Kings Langley WD4 8ET Steve.Cohen@gp-e82129.nhs.uk

levels would be a bit higher.” “No,” I replied, “the glucose I took should have suppressed growth hormone if my pituitary had been functioning normally.”

“Hold on a minute,” said the doctor. “I’ll check the test results again.” She returned to tell me my growth hormone levels. “That’s high above normal,” I exclaimed. “Surely it means I have a condition known as acromegaly?” “I’m afraid you’re now talking beyond my depth,” admitted the doctor. “I think it’s best you wait for your endocrinologist to come back from holiday.”

I didn’t want to wait. Instead, I wrote to an acromegaly expert, Professor John Monson, at St Bartholomew’s Hospital, London. I enclosed my test results and told him I had acromegaly. The next day he phoned to ask, “Who made the diagnosis of acromegaly?” I paused. “I did,” I said, somewhat hesitantly, waiting for his rebuke. “Well,” said the professor, “you’re absolutely right.” When we met a few days later, in August 2001, Professor Monson said he could tell immediately that I had acromegaly just by looking at me. The test results confirmed that my growth hormone levels were around 2500% above normal. There was no doubt I had acromegaly.

The diagnosis at last provided an explanation for all my strange and incapacitating symptoms, which was a huge relief. I also felt let down. Surely any one of my doctors should have been able to recognise what was wrong with me much earlier? Apparently, the typical delay in diagnosis for acromegaly is 10 years or more. Isn’t this a failure of medicine? The challenge for doctors, and patients, is to catch
How to become an MP

PERSONAL VIEW Steven Duncan Ford

Ours is a participatory democracy—startling chiefly for the lack of participation. Membership of the hopelessly debt laden political parties is at an all time low, and the esteem in which parliament is held is likewise low. A variety of means to improve the situation have been mooted, but little seems to change. Could we doctors make a change? It is often noted that the public has more trust in doctors and other healthcare workers than in politicians; this is the right time, a last chance perhaps, to take our approval ratings out for a psephological test drive.

We have witnessed problems arising from the healthcare “reforms” and can predict more of the same. As a group we have been on the receiving end of much government inspired unpleasantness: the precedence of budgets over care of patients, fragmentation of care, Modernising Medical Careers and the medical training application service (MTAS), crassly incompetent management at all levels, contract chaos, hostile press briefings, black propaganda, downright lies, tendentious media comment, privatisation, unfeasibly complex recertification, licensure, and appraisal procedures, and so on. Is alerting the populace to the problems and offering alternatives in some way problematic or unprofessional? Are parliamentarians, their advisers, and the commentators more competent at generating and implementing reform than we are?

The broad sentiment of the profession toward the current reforms is negative, although no one shrinks from the need to improve continuously in personal as well as system performance and standards. If we are sincere in our reservations about the course proposed for health care in the United Kingdom, is it legitimate or excusable to do nothing? If we are not sincere in our reservations, then we should put off Luddism and apply all our strength to the task. Malcontent, whining, or acquiescence is neither a sustainable posture nor a pretty sight. Choices can be placed before the electorate. “Contestability,” in so far as it has a meaning, need not apply only to the supply of health services. If anything, with the dearth of inspiring policy on so many aspects of national life, national government itself stands in greater need of a searching assessment of contestability than the NHS.

With few exceptions, healthcare workers—not just doctors—are hard working, dynamic, committed, educated, adept at problem solving, resilient under pressure, articulate, imaginative, disposed to serve the community, of broadly liberal outlook, and of beneficent intent. These surely are the sort of values that the electorate might care to see more widely represented in parliament. By comparison, parliamentarians are increasingly narrowly confined in their views on policy, preoccupied with party matters to the exclusion of governing effectively.

Although the hope for a 2007 election has come to nothing, we now have time to marshal our forces. Please may we have an independent candidate in every constituency: a healthcare worker who can offer the electorate a better choice of health policies than those on offer from the established parties? We can do this. We can make a change.


The application paperwork, from your local acting returning officer, is available only after an election has been called. The papers must be returned at least two weeks before polling day, along with £500 in cash or a banker’s draft—no personal cheques or credit cards are accepted.

MPs are not better than the rest of us and do not deserve the unopposed scope for harm that the electorate has so far given them. Acquiescence is complicity—examine your conscience before discounting yourself as a candidate in your own constituency. At the least, an all constituency health professional candidacy will push the health issue up the political agenda, and that might promote fresh thinking and debate. It’s time to put up or shut up.

Steven Duncan Ford is a general practitioner, Haydon and Allen Valleys Medical Practice, Hexham, Northumberland. mail@stevenford.co.uk

Competing interest: SDF intends to stand at the next general election.

Please may we have an independent candidate in every constituency: a healthcare worker who can offer the electorate a better choice of health policies than those on offer from the established parties?
An exhibition focusing on representations of the sex act down the ages shows that the way sex is depicted is as unchanging as sex itself, Colin Martin finds.

Entry to the gallery is restricted to over 18s; however, thoughts of sex shops, men in raincoats, or catalogues wrapped in brown paper are inappropriate. Fig leaves have fallen—the first exhibit is a plaster leaf made to cover the genitals of a plaster cast of Michelangelo’s David to spare Queen Victoria’s blushes—but intellectual propriety remains firmly in place throughout the show. Intimate viewing spaces, decorated in 18th century colours, hint at earlier, private closets of erotica but are ideal for close inspection of works on paper.

Depicting Jupiter’s exploits was a standard recourse for Western artists when patrons sought overtly sexual subjects. Rembrandt’s 1659 drypoint etching Jupiter and Antiope depicts the god as a lusty half man, half goat, lasciviously pulling aside bedcovers to reveal Antiope’s discreetly shadowed pudenda. Unlike artists in Europe, Indian, Chinese, and Japanese graphic artists of the 17th to 19th centuries depicted sex more explicitly. Illustrations for Japanese woodblock prints known as “shunga” were intended for use in brothels and private homes, to excite both sexes. A Japanese watercolour painted around 1830 shows a man and woman stimulating each other, their genitalia instructively depicted on a larger scale than their naked bodies. Oysters are scattered in the foreground.

The exhibition considers the roles of Sigmund Freud and Alfred C Kinsey in interpreting and studying sexual practices. Freud’s work revealed that traumatic memories were often related to sexual experiences before puberty. His influence is apparent in contemporary work by Viennese artists, including Egon Schiele. Eros (1911), a young man with a red erect penis, and Woman with Black Stockings (1913), displaying her pudenda, convey feverish, psychologically charged sensuality. Later, the Surrealists were indebted to the psychoanalyst’s notion of unconsciousness as the seat of erotic desire.

Kinsey’s work is represented by a series of projected images of photographs of sexual acts, which constituted the “primary material” he used in his groundbreaking surveys of human sexual behaviour, climaxing in the publication of Sexual Behaviour in the Human Male (1948) and Sexual Behaviour in the Human Female (1953). The Association of Women Students at Indiana University, which in 1938 asked him to teach human sexual biology to married students or those about to marry, can’t have known what they were starting.

Collecting and classifying specimens was second nature for the entomologist Kinsey, who compulsively collected 75,000 erotic photographs and films and 7000 other kinds of visual images of sex. “No other survey with a scientific basis has ever been enriched by a collection of imagery in such a fundamental way,” comments Wallace.

An unknown photographer’s 1943 image, donated to the Kinsey Institute in 1960, classified as Coitus Female Supine Ventral-Ventral Male Prone, captures a naked couple on the front seat of a car. The man’s head turns over his left shoulder towards the photographer with a quizzical expression, in a sensuous tableau that could equally have been carved by a Roman sculptor, showing a startled satyr caught in flagrante with a nymph. Attitudes regarding displaying depictions of coitus might alter over centuries; the depictions, however, haven’t changed.

Colin Martin is an independent consultant in healthcare communication, London.
MTAS in GP land

I once took a six month career break and jetted around the world. Before we knew about climate change, this was guiltless travel, with no token planting of a small tree in the garden as “carbon offsetting.” Huddled in rain battered bus shelters across South East Asia, I joined the ragbag of travelling Brits: lorry drivers, miners, hairdressers, and the ubiquitous Oxbridge gap year kids. We had been released from our respective social shackles, and we found camaraderie and humour. I had space to reflect. At the time I was a career obstetrician, but I knew this was mere denial: I wasn’t emotionally robust enough. So I came home to fulfil my destiny and start GP training.

Professor John Took has been leading a comprehensive review of Modernising Medical Careers, sparked by the debacle over the medical training application service (MTAS) (BMJ 2007;335:737). Its suggestions are to add more flexibility and local accountability and to reinstate traditional assessments of “excellence,” while rationalising the duration of training—all for the good. Training of GPs receives only a couple of footnotes, as it was largely untouched by the MTAS problems—mainly, I suspect, because it remains unpopular with upwardly mobile young medics. I could say that it doesn’t bother me that general practice is seen as a last resort for the medical elite, but this would not be true. “You’re too good to be a GP” was one of the most insensitive remarks anyone ever made about my career choice.

The review recommends extending GP training to five years—this is long overdue. But any extra training must take place in our specialist setting: general practice. Otherwise the suspicion is that our registrars will be used as ever to plug holes in hospital rotas. The review also repeats an aspiration that all specialist registrars should join us briefly in GP land. This may cause much eye rolling from certain hospital colleagues, but the community can teach much about patients’ health seeking behaviour and the dangers of medicalisation.

This review still has a faint whiff of hospital elitism, but we doctors no longer stride the medical plantation in pith helmets. We are huddled together in our corroded profession, battered by the deluge of political interference and unrealistic expectations. This is a time for camaraderie and breaking traditional medical shackles.

Des Spence is a general practitioner, Glasgow destwo@yahoo.co.uk

A man who called hospitals to account

Sir Derek Wanless is not the first banker to call health services to account. Preceding the former NatWest chief executive by more than a century, the unsung pioneer of modern hospital management Sir Henry Burdett (1847-1920) started out as a humble clerk behind a desk at Midland Bank in Birmingham.

A vicar’s son with ambitions to become a doctor, the young Burdett had to abandon his altruistic aims at 17 when his father’s death put paid to his university plans. In the balance sheet of history, his loss was health care’s gain. Burdett’s brief spell with the world’s local bank sowed a passion for figures, which inspired him to modernise the world’s local health services.

Appointed manager of the Queen’s Hospital, Birmingham, four years later, Burdett embarked on a lifelong crusade to improve standards of accounting and management in hospitals worldwide. Although he spent only 12 years as a hospital manager—his second post was at the helm of the Dreadnought Hospital, Greenwich—before turning to business, Burdett launched a revolution in management practices which reverberates today.

Championing ideas ahead of their time, Burdett pioneered hospital league tables, clinical audit, standardised accounting, and general management. He supported lay authority over doctors, introduced the first registration system for nurses, and was the main mover behind the foundation of the King’s Fund, which would commission a study by his banking successor 110 years later.

Tirelessly collecting hospital statistics with the aid of a secretary, a primitive typewriter, and the Victorian postal system, Burdett published figures for everything from mortality rates to the cost of leeches in his annual tome, The Hospitals Year-Book. Scathing in his criticism of inefficiency, he pronounced St Thomas’ “among the least successfully managed hospitals in London”—and Guy’s was “in an even worse plight.”

His views did not endear him to the clinical professions. Florence Nightingale refused to be photographed with him, and his reforms at Addenbrooke’s prompted one doctor to declare: “Everybody was contented until in an evil day they called into their counsel Sir Henry Burdett.”

But for all the snobbery against the erstwhile bank teller, Burdett was impeccably connected. He was knighted by Queen Victoria and counted the Prince of Wales among his closest friends. His advice was sought throughout the British Empire and beyond. As a prototype management consultant, he offered his expertise to hospitals from Chicago to Russia, and his Christmas Day was never complete without a ward visit. Although he was fiercely opposed to state intervention—he described politicians as “the curse of the sick”—his reforms paved the way for the modern NHS.

Wendy Moore is a freelance writer and author, London wendymoore@ntlworld.com

Des Spence
FROM THE FRONTLINE
PAST CARING
wendymoore@ntlworld.com
VIEWS & REVIEWS

BMJ | 20 OCTOBER 2007 | VOLUME 335
**Miscellaneous truths**

The *Harleian Miscellany* is a selection of early pamphlets from the library of the Earl of Oxford, catalogued by Samuel Johnson, and first published in 1744. Whoever opens one of its 10 volumes is almost certain to be detained for hours, for who can resist titles such as “A short, legal, useful, safe, and easy Prescription to recover our Kingdom, Church and Nation, from their present dangerous, destructive, destructive Confusion, and worse than Bedlam Madness,” a title to ensure perennial contemporary relevance, though published in 1659?

Or again, “The She-Wedding: or, a Mad Marriage, between Mary, a Seaman’s Mistress, and Margaret, a Carpenter’s Wife, at Deptford, being a full Relation of a cunning Intrigue, carried on and managed by two Women, to hide the Discovery of a great Belly, and make the Parents of her Sweet-heart provide for the same,” published in 1684, and which begins with an eloquent but no longer admissible animadversion on the nature of women:

“It hath been the policy of the prince of darkness in all ages, when any work of his was to be carried on, which required a more than ordinary cunning, to employ a female craft therein: nor indeed from his first attempt in that kind, in the betraying our mother Eve, did he ever find reason to blame his discretion in the said method, since he scarce ever failed thereby of his ends.”

There is much of medical interest in the miscellany—for example, “A Discourse, setting forth the unhappy Condition of the Practice of Physick, and offering some Means to put it into a better; for the Interests of Patients, no less, or rather much more, than of Physicians,” by Jonathan Goddard, professor at the Gresham College, published in 1670.

“No less, or rather much more” seems to betray a certain unease or anxiety: methinks the doctor doth protest too much. Yet which of us in present day Britain could strongly disagree with Dr Goddard in his defence of the profession?

“The art of physick hath had, in common with other arts and professions, the infelicity to be abused by the professors thereof; who, either out of insatiable avarice to make the utmost gain to themselves thereby, or out of pride and state, or humour, have given just occasion to the world to judge, that they had not that care and consideration of the lives and healths of persons with whom they had to do, as, in humanity, reason and conscience, they ought to have had.”

No doubt this is still so; and newspapers call me several times a year to ask me to reveal what a bad lot doctors are. But we ought to get things into perspective, as Dr Goddard did a third of a millennium ago: “Admitting this to be inexcusable, as to the persons guilty of it; yet it may be said, as to the present professors thereof, that there was never in any age, less grievance or cause for complaint upon any such account.”

Bravo! Of no age is this more true than our own. Nevertheless, Dr Goddard goes on to diagnose the reason for hostility to the profession: “That distinction between the vices of persons, and of arts and professions, is so clear and obvious, that whosoever transfers those of one upon the other, must needs appear deficient in the use of his reason, or else partial and injurious.”

But then, governments are both deficient in their use of reason and partial and injurious.

Theodore Dalrymple is a writer and retired doctor

---

**MEDICAL CLASSICS**

**The Cunning Man**

By Robertson Davies

First published by Viking in 1994; reprinted by Penguin, 1996

“Should I have taken the false teeth?” Thus, the curt opening of *The Cunning Man* by Robertson Davies, typical of Canada’s foremost man of letters, professor of English, and author of three brilliant trilogies telling stories about the odd ways in which people behave. The speaker is Dr Jonathan Hullah, Toronto physician and sometime police surgeon, whose reminiscences are prompted by the sudden death of his old friend, the saintly Father Ninian Hobbes, while celebrating Good Friday mass at St Aidan’s church. The question will be answered only after 400-odd pages of what could be called a detective story, but a seasoned reader will know that there will be diversions into any topic that takes the fancy, as well as interesting if not eccentric characters and many quotable one-liners.

When young Hullah nearly died from scarlet fever, he reckoned his life was saved by Elsie Smoke, a native American who set up her tepee in his garden and whose secrets he tried to fathom. He was a precocious lad and a Freudian fanatic as a medical student until a teacher said the theories were only “a flash in the bedpan.” His favourite bedside reading was Thomas Browne’s *Religio Medici*, which, he said, “brought a sweet humanism to the gross materialism of much medical information.” In Europe as a medical officer in the 1939-45 war, he was put in charge of a ward in an Oxford hospital containing 26 shell shocked victims of friendly fire: “They talked, I listened for an hour three times a week.” In the bookshop Blackwell’s he came across an anthology of poems called *The Reader’s Companion* and used it to take their minds off their misfortunes and restore their optimism in the future. He reckoned this therapy was a great success, but the top brass were less enthusiastic.

Back in Toronto, Hullah converted a barn near St Aidan’s into a consulting room and embarked on what he called a type of psychosomatic medicine where mind and body mingle and “untangling the relationship is the Devil’s own work.” When you carry out a comprehensive examination, he says, you must be alert to the demeanour, tone of voice, and even the smell of the patient: you will not really know them until you have seen them in their own home. Obvious physical disease does not exempt you from a duty to relieve the grievances and raw deals that dog people’s lives. Outcomes, successful or otherwise, are illustrated by a series of case histories, but because of his diagnostic skills he came to be known as the Cunning Man, after the wise man that “used to be found in every English village.”

Alex Paton, retired consultant physician, Oxfordshire alexpaton@doctors.org.uk

---

**BETWEEN THE LINES**

Theodore Dalrymple

Newspapers call me several times a year to ask me to reveal what a bad lot doctors are

---

**Robertson Davies:**

brilliant storyteller

---

The speaker is Dr Jonathan Hullah, Toronto physician and sometime police surgeon, whose reminiscences are prompted by the sudden death of his old friend, the saintly Father Ninian Hobbes, while celebrating Good Friday mass at St Aidan’s church. The question will be answered only after 400-odd pages of what could be called a detective story, but a seasoned reader will know that there will be diversions into any topic that takes the fancy, as well as interesting if not eccentric characters and many quotable one-liners.

When young Hullah nearly died from scarlet fever, he reckoned his life was saved by Elsie Smoke, a native American who set up her tepee in his garden and whose secrets he tried to fathom. He was a precocious lad and a Freudian fanatic as a medical student until a teacher said the theories were only “a flash in the bedpan.” His favourite bedside reading was Thomas Browne’s *Religio Medici*, which, he said, “brought a sweet humanism to the gross materialism of much medical information.” In Europe as a medical officer in the 1939-45 war, he was put in charge of a ward in an Oxford hospital containing 26 shell shocked victims of friendly fire: “They talked, I listened for an hour three times a week.” In the bookshop Blackwell’s he came across an anthology of poems called *The Reader’s Companion* and used it to take their minds off their misfortunes and restore their optimism in the future. He reckoned this therapy was a great success, but the top brass were less enthusiastic.

Back in Toronto, Hullah converted a barn near St Aidan’s into a consulting room and embarked on what he called a type of psychosomatic medicine where mind and body mingle and “untangling the relationship is the Devil’s own work.” When you carry out a comprehensive examination, he says, you must be alert to the demeanour, tone of voice, and even the smell of the patient: you will not really know them until you have seen them in their own home. Obvious physical disease does not exempt you from a duty to relieve the grievances and raw deals that dog people’s lives. Outcomes, successful or otherwise, are illustrated by a series of case histories, but because of his diagnostic skills he came to be known as the Cunning Man, after the wise man that “used to be found in every English village.”

Alex Paton, retired consultant physician, Oxfordshire alexpaton@doctors.org.uk
OBITUARIES

John Michael Newsom-Davis

Enabled the understanding and treatment of myasthenia gravis

Some neurologists and neuroscientists have made outstanding contribution to basic science, and others to treatment. John Newsom-Davis did both, and his work is credited with transforming the lives of people with myasthenia gravis and related autoimmune brain diseases. He contributed to the discovery that myasthenia gravis was caused by antibodies and, with Anthony Pinching and Keith Peters, showed that plasmapheresis could produce a profound, though short lived, remission. He went on to develop effective treatment using immunosuppressive drugs.

John Newsom-Davis was born in Harpenden. His father, managing director of the Davis gas cooker company, and mother, a surgeon’s daughter, sent him to Sherborne School, which he didn’t much enjoy. He followed this with national service, where he learnt to pilot Meteor fighter planes, which gave him great pleasure and pride. He was offered a place at Cambridge to read English and history but changed his mind and decided to follow his grandfather into medicine, which meant cramming science.

He did his preclinical years at Pembroke College and his clinical training at the Middlesex. He qualified in 1960 and remained at the Middlesex as neurology registrar before moving in 1967 to the National Hospital for Nervous Diseases as resident medical officer and lecturer in neurology. Here he was influenced by Michael Kremer and Tom Sears.

He and Sears did experimental work on the physiological control of breathing, often using themselves as guinea pigs. He spent 1969–70 at Cornell Medical Center, New York, investigating the central pathways involved in breathing with Fred Plum.

He returned to England as consultant neurologist at Queen Square and the Royal Free and started the myasthenia research that defined his career. The disease was known to be caused by autoimmunity, and he pioneered plasmapheresis, which gave short term remission. After others discovered that the autoantibodies were directed against receptors at the neuromuscular junction, he characterised them and found that they are produced by the thymus in response to inappropriate stimuli. He and his team went on to uncover the immunology of various types of myasthenia, including rare inherited forms. He then demonstrated that these diseases could be controlled by immunosuppressive treatments.

In 1980 he became the first MRC Clinical Research professor at the Royal Free. In 1987 he was appointed to the Action Research chair of clinical neurology at Oxford, and his research team moved to Oxford with him. From then onwards, writes Angela Vincent in the Independent (18 Septem-

ber), he built up clinical neuroscience there, including establishing a world-leading centre for magnetic resonance functional brain imaging. His official retirement in 1999 made no difference to his activity, and he continued his research, and seeing patients.

Newsom-Davis was one of the few medical scientists to be made an FRS. He was a founder fellow of the Academy of Medical Sciences, a former president of the Association of British Neurologists (1999-2000), an honorary member of the American Academy of Neurology, and an elected foreign member of the Institute of Medicine of the US National Academy of Sciences. He received many medals and awards and was made CBE in 1996.

He was an active head of the biomedicai section of the British Association for the Advancement of Science and also gave his time generously to the Royal College of Physicians, Medical Research Council, Association of British Neurologists, Myasthenia Gravis Association, and Muscular Dystrophy Group. Under his editorship, the journal Brain was one of the first to introduce electronic processing and online publication.

In July 2005 he was awarded $5m from the US National Institute of Neurology Disorders and Stroke to lead an international trial to determine whether all patients with myasthenia gravis needed to undergo thymectomy. This is standard treatment, but no one knows whether it is justified in patients who do not have a thymus tumour. The trial compares prednisone alone with prednisone and thymectomy. He said at the start that previous studies were less than rigorous in patient selection and scientific design. He and his US collaborators aimed to recruit 200 patients in four continents and follow up each for three years.

As part of this study he visited a neurological clinic in Bucharest. He and his wife then left to visit ancient frescoed monasteries in Moldova, driving on dangerous roads. He was killed in a car crash.

He leaves a wife, Rosemary; two daughters and one son; and seven grandchildren.

Caroline Richmond

Stanley Norman Cole
Former general practitioner Guildford (b 1915; q Guy's Hospital, London, 1938), d 12 July 2007.
Stanley Norman Cole joined the Royal Navy at the start of the second world war, eventually serving on the aircraft carrier HMS Implacable bringing prisoners of war home from the Far East. After the war he was a registrar at Queen Charlotte's Hospital, London, taking this experience to general practice in Guildford, where he co-founded the Mount Alvernia maternity service. He and his partners built one of the first purpose built surgeries designed in the round. Stanley was a founding fellow of the Royal College of General Practitioners, and became provost of the South East Faculty in 1968. He retired from general practice in 1981. Predeceased by his wife, Margaret, in 2001, he leaves two daughters and four grandchildren.

Malcolm Read

Nigel Henry Harris
Former orthopaedic surgeon St Charles and St Mary's Hospital, London (b 1924; q Cambridge/Middlesex Hospital 1947; FRCS), d 8 July 2007.
Nigel Harris pioneered hip and knee replacement, and was editor of the Postgraduate Textbook of Clinical Orthopaedics. He established the first NHS sports clinic in 1972, and was surgeon to Arsenal Football Club and the Football Association. As chair of the planning committee, he enabled the development of the Queen Elizabeth the Queen Mother Wing at St Mary's. Nigel joined the Academy of Experts in the 1980s, became one of the first medical fellows, and co-edited the standard text on medical negligence. A founding trustee of the Police Rehabilitation Trust, he championed health matters in 28 letters in the Times. He leaves a wife, Elizabeth, and two sons. A memorial service will be held on 12 November 2007.

Andrew Harris

Michael Stuart Brighton Hollinrake
General practitioner Northwich, Cheshire (b 1944; q Galway, National University of Ireland, 1969; BAO, DObstRCOG), died from acute coronary thrombosis on 5 April 2007. Educated at Stonyhurst College from the age of 8, Michael Hollinrake studied medicine in the Republic of Ireland, returning to the United Kingdom on graduation. After house jobs in Littlerborough and Rochdale and pathology at the University Hospital of South Manchester he settled in general practice in Cheshire in 1972. In 2002 he semi-retired, becoming a locum general practitioner. A devout Catholic, he enjoyed, in recent years, voluntary medical work in Lourdes. He leaves a wife, Aileen; two daughters; and one grandson.

Aileen Hollinrake (Jennings), Patricia Hollinrake Sarah Hollinrake

Colin Protheroe
Former consultant psychiatrist St Nicholas Hospital, Newcastle upon Tyne (b 1929; q Sheffield 1953, MD, FRCPEd, FRCPsych), died from diabetic complications on 11 August 2007. A fine rugby player, Colin Protheroe was selected for Welsh International Schoolboys. During his national service in Berlin he was the British medical officer at Spandau prison and looked after high ranking Nazi prisoners. His MD thesis on puerperal psychosis was a major contribution. In his hospital work he was a superb administrator and a realist who, though he had strong views, was prepared to compromise. Colin was also BMA representative. After retirement from the NHS he worked for many years as a member of the Mental Health Review Tribunal. Predeceased by his wife, Ann, in 2006, he leaves three children and three grandchildren.

H A McClelland, T A Kerr

Geoffrey Sheers
Former consultant chest physician Plymouth General Hospital, and civil consultant Royal Navy (b 1913; q Cambridge 1938; MD), d 17 June 2007.
Soon after training at St Thomas' Hospital, Geoffrey Sheers joined the Royal Naval Volunteer Reserve and was posted to Egypt, where he saw active service as a surgeon lieutenant, and was torpedoed twice. In 1946 a spell as a patient at Papworth Hospital with tuberculosis led to his specialising in chest medicine. His main interest was in occupational chest diseases, and he published studies on granite and kaolin workers. His contribution to the understanding of the effects of asbestos in dockyard workers led to his appointment as a civil consultant to the navy, and he continued to be consulted by workers after retirement from the NHS. He leaves two sons and five grandchildren.

Roger Sheers

Julian Edmund Christopher Tower
Former general practitioner Headcorn, Kent (b 1920; q Cambridge/St Thomas’ Hospital 1953; MRCP), died from cerebral metastases (unknown primary) on 5 June 2007. Julian Tower read classics at Cambridge for one year and sang in King’s College Chapel choir before he was called up in 1940, serving in the Royal Navy in HMS Valiant and HMS Shah. On demobilisation in 1946 he returned to Cambridge to read medicine. He went into general practice in 1955, retiring in 1985. Julian was passionate about music. He conducted Cranbrook and District Choral Society from 1960 to 1985. In retirement he was on the committee of Westmorland Music Council, and sang in Kendal South Choir. He kept up his fluent French with a local group. Predeceased by a son, he leaves a wife, June; three children; and eight grandchildren.

Monica Baynes, June Tower

John Ernest Michael Whitehead
Former director Public Health Laboratory Service (b 1920; q Cambridge/St Thomas’ 1945; FRCPath), died from Parkinson’s disease on 10 August 2007.
After junior posts in bacteriology at St Thomas’ and a postgraduate travelling fellowship in Copenhagen, Michael Whitehead joined the fledgling Public Health Laboratory Service, first at Sheffield, then as director at Coventry. He was appointed director of the service in 1981, having been deputy since 1976. He managed the service through financial constraint and the completion of the new development at Colindale. In 1983 Michael was elected vice president of the Royal College of Pathologists. After retirement, he was specialist adviser to the Social Services and Agricultural Committees of the House of Commons, notably in the much publicised affair of salmonella infection in poultry. Predeceased by his wife, Betty, in 1996, he leaves two children and six grandchildren.

Stephen Whitehead
Reports by the media of the suicide of a celebrity may have led to an increase in people killing themselves, according to a Taiwanese study in the International Journal of Epidemiology (2007 Sep 28 doi: 10.1093/ije/dym196). The number of suicides increased in the four weeks after the media reported the death of a television actor (relative risk 1.17, 95% confidence interval 1.04 to 1.31). The relative risk was greater among men (1.30, 1.14 to 1.50) and for people using the same method (hanging) as the actor (1.51, 1.25 to 1.83).

Two groups of scientists have independently discovered that the content of nitric oxide rapidly decreases in bank (stored) blood, which may reduce the flow of blood in some people (Proceedings of the National Academy of Sciences 2007 Oct 11 doi: 10.1073/pnas.0707958104 and doi: 10.1073/pnas.0708160104). A nitric oxide signal carried by haemoglobin induces vascular relaxation where necessary, but with low concentrations of nitric oxide this is unlikely to happen, increasing the risk of a dangerous drop in blood flow. The scientists also found that adding nitric oxide to stored blood could restore vasodilation.

Young people who have dislocated their shoulder have the reduced shoulder conventionally immobilised in internal rotation. But a randomised controlled trial of immobilisation in either internal or external rotation, after an initial shoulder dislocation, points to external rotation being the better position to reduce the risk of recurrence of dislocation (Journal of Bone and Joint Surgery (Am) 2007;89:2124-31). The relative risk reduction of a recurrence was 38.2% in the external rotation group and 46.1% in patients aged less than 30.

Sometimes it’s difficult to tell the difference between the states of “burnout” and “prolonged fatigue” among workers. An observational study in the QJM concludes that although there is considerable overlap between the two states and they can occur simultaneously, there are also some important and relevant differences when looking at work and health factors (2007;100:617-27). Having both at the same time results in a predictably worse outcome.

Does smoking influence the progression of multiple sclerosis or the accumulation of disability caused by the illness? Apparently not, according to a paper in Neurology (2007;69:1515-20). Although the incidence of multiple sclerosis is rising among smokers, whether the relapsing-remitting sort or the sort that progresses steadily, smoking plays no part in the course of the disease.

To explore the existence of biomarkers of dairy fat intake that correlate with heart disease, data from 166 of the women who had participated in the nurses’ health study were matched with 327 controls (American Journal of Clinical Nutrition 2007;86:929-37). Multivariate analysis showed that women with greater plasma and erythrocyte concentrations of fatty acids with a carbon chain of 15:0 had a significantly higher risk of ischaemic heart disease. The other useful biomarker for dairy fat intake was the fatty acid trans isomer trans 16:1n-7.

Endovascular stenting for traumatic aortic rupture compares well with repair by open surgery when patients are haemodynamically stable. A retrospective review of 15 consecutive cases in one institution performed over six years found that stenting was a feasible option and could be performed in a wide range of cardiac patients (Journal of Thoracic and Cardiovascular Surgery 2007;134:897-901). It’s a shorter procedure and requires less heparin and blood products. The duration of artificial ventilation; time spent in the intensive care unit; and overall hospital stay were no different than with open surgery, however.

Young women undergoing chemotherapy or radiotherapy often run the risk of becoming infertile from the treatment, and they usually resort to assisted reproductive technology to protect their future child bearing potential. Analogues of gonadotrophin releasing hormone are also given to protect the ovaries, but a review in The Oncologist says that there’s no conclusive evidence for either the safety or the efficacy of such treatment for gonadal injury caused by chemotherapy or radiotherapy (2007;12:1055-66).

Does the amount of sleep children have influence their growth? More than 300 children were followed up at 12, 18, and 24 months after birth and then every year until the age of 10 years (Pediatrics 2007;120:e769-76). Duration of sleep varied widely at all ages, but most children showed long term stability in their sleep patterns. Sleep duration and height or weight at any age showed no association.

Skin can give away all kinds of clues about our health. In the case of two patients described in CMAI their trunks were covered with itchy papules of keratotic, follicular lesions that were finally diagnosed as “phrynoderma” (2007;177:855-6). Both men were shown to be deficient in various minerals and vitamins, and the phrynoderma attributed to poor nutrition. Both were treated with keratolytic ointments and put on strict but healthy diets, which included vitamin supplements. Complete regression of the papules was achieved in three to six months.