1. MMC: mass medical culling • EDITORIAL  
   Page 879  
   The Lancet

2. Reforming research in China • EDITORIAL  
   Page 880  
   The Lancet

3. Promoting optimum management for ADHD • EDITORIAL  
   Page 880  
   The Lancet

4. ACUITY-PCI: one drug does not fit all • DISCUSSION  
   Pages 881-882  
   Ron Waksman

5. Cardiac arrest—guideline changes urgently needed • DISCUSSION  
   Pages 882-884  
   Gordon A Ewy

6. Effects of zinc supplementation on child mortality • DISCUSSION  
   Pages 885-886  
   Shinjini Bhatnagar

7. Objectification of physicians and loss of therapeutic power • DISCUSSION  
   Pages 886-888  
   Iona Heath and John Nessa

8. Podoconiosis: the most neglected tropical disease? • DISCUSSION  
   Pages 888-889  
   Gail Davey and Melanie Newport
9. COPD—call for papers • DISCUSSION
   Page 889
   Sabine Kleinert

10. Clinical update: the low-glycaemic-index diet • DISCUSSION
    Pages 890-892
    David S Ludwig

11. WHO Director-General announces senior team • NEWS
    Pages 893-895
    Hannah Brown

12. US guidelines seek to protect access to licensed technology • NEWS
    Page 896
    Michael McCarthy

13. Grieving for our children • BOOK REVIEW
    Pages 897-898
    George Rousseau

14. Dying of drugs • BOOK REVIEW
    Page 898
    Ernest Drucker

15. Margaret Chan: now is the time for WHO to achieve results • DISCUSSION
    Page 899
    Hannah Brown

16. James Allen Dvorak • PERSONAL REPORT
    Page 900
    Karen Masterson

17. A withdrawn prepublication • CORRESPONDENCE
    Page 901
    Peter B Dean

18. A withdrawn prepublication • CORRESPONDENCE
    Page 901
    Sabine Kleinert

19. A withdrawn prepublication • CORRESPONDENCE
    Pages 901-902
    Tom Jefferson
20. Wilson's disease • CORRESPONDENCE
Page 902
Lorenzo Leggio, Giovanni Gasbarrini and Giovanni Addolorato

21. Wilson's disease • CORRESPONDENCE
Page 902
JM Walshe

22. Wilson's disease • CORRESPONDENCE
Pages 902-903
Samar Harris, Harris VK Naina and Sameer Siddique

23. Subarachnoid haemorrhage • CORRESPONDENCE
Page 903
Yatindra Kumar Batra and Subramanyam Rajeev

24. Subarachnoid haemorrhage • CORRESPONDENCE
Pages 903-904
Ilonka Kreitschmann-Andermahr and Harald Jörn Schneider

25. Subarachnoid haemorrhage • CORRESPONDENCE
Page 904
Ian Holbrook, Robert Beetham, Anne Cruickshank, Geoff Keir and Ian Watson

26. Subarachnoid haemorrhage • CORRESPONDENCE
Page 904
Oscar Mlungu Jolobe

27. Stroke prevention: missed opportunities • CORRESPONDENCE
Pages 904-905
Sunku Hemanth Guptha, Pazhvoor Shibu and Peter Owusu-Agyei

28. Peace through medical education in Bosnia and Herzegovina • CORRESPONDENCE
Page 905
Dario Sambunjak and Vladimir J Šimunović

29. Sexual dysfunction in HIV infection • CORRESPONDENCE
Pages 905-906
Maria Paola Trotta, Adriana Ammassari, Rita Murri, Antonella d'Arminio Monforte and Andrea Antinori

30. Peer review and the Term Breech Trial • CORRESPONDENCE
Page 906
Susan Bewley and Andrew Shennan
31. Department of Error • ERRATUM
Page 906

32. Department of Error • ERRATUM
Page 906

33. Bivalirudin in patients with acute coronary syndromes undergoing percutaneous coronary intervention: a subgroup analysis from the Acute Catheterization and Urgent Intervention Triage strategy (ACUITY) trial • ARTICLE
Pages 907-919
Gregg W Stone, Harvey D White, E Magnus Ohman, Michel E Bertrand, A Michael Lincoff, Brent T McLaurin, David A Cox, Stuart J Pocock, James H Ware, Frederick Feit, et al.

34. Cardiopulmonary resuscitation by bystanders with chest compression only (SOS-KANTO): an observational study • ARTICLE
Pages 920-926

35. Effect of zinc supplementation on mortality in children aged 1–48 months: a community-based randomised placebo-controlled trial • ARTICLE
Pages 927-934
Sunil Sazawal, Robert E Black, Mahdi Ramsan, Hababu M Chwaya, Arup Dutta, Usha Dhingra, Rebecca J Stoltzfus, Mashavi K Othman and Fatma M Kabole

36. Bipolar disorder—focus on bipolar II disorder and mixed depression • REVIEW ARTICLE
Pages 935-945
Franco Benazzi

37. Management of functional somatic syndromes • REVIEW ARTICLE
Pages 946-955
Peter Henningsen, Stephan Zipfel and Wolfgang Herzog

38. Interpreting health statistics for policymaking: the story behind the headlines • REVIEW ARTICLE
Pages 956-963
Neff Walker, Jennifer Bryce and Robert E Black

39. An unexpected diagnosis after unstable angina • SHORT COMMUNICATION
Page 964
Georgia Tunnicliffe, Nigel Raymond and Mike Nowitz
MMC: mass medical culling

March has seen the meltdown of the new computerised system for junior doctors applying for specialist training in the UK NHS. The Medical Training Application Service (MTAS) was so flawed that some consultants refused to interview short-listed candidates. The Department of Health has U-turned, and applications submitted for the first round will now be given the opportunity to be reassessed. An independent review of the process in advance of the second round of applications in April is also underway. But, this is all too little too late. The fatally flawed MTAS should be suspended until the process has been comprehensively piloted and evaluated, and the Department of Health should use this opportunity to re-think the Modernising Medical Careers (MMC) curriculum.

The proposal for radical reforms in medical training was set out in a report by the Chief Medical Officer for England in 2002 and has been led by the Postgraduate Medical Education and Training Board, which is accountable to Parliament. Under the previous system, it took about 14 years to become a consultant. In 2005 this was streamlined under the MMC curriculum, with graduates completing a 2-year foundation programme, then immediately entering specialist training, lasting 3-7 years depending on specialty. MMC is currently in the transition period, where trainees from the former system are competing for specialist training posts against those in the new foundation programmes. The competition is fierce, with 30,000 applicants for 22,000 jobs.

MTAS is the central portal through which a trainee can apply for up to four specialist training posts. The first round of applicants was short-listed at the end of February, 2007. From the outset MTAS has been plagued by problems: the computerised system failed, and trainees complained that their practical experience and academic qualifications cannot be reflected adequately in the generic application form. Some interview dates clash with scheduled professional examinations, leaving applicants to make the difficult choice of which to attend.

Consultants have also expressed their lack of confidence in MTAS, with concerns that the most suitably qualified candidates have not been short-listed for interviews because of the computerised selection process, and that the application form is geared primarily to foundation trainees. These worries will not be addressed merely by reassessing first-round applications.

But the shambles of MTAS is only one aspect of a larger problem. MMC itself is failing. Trainers and trainees are concerned that the new system will produce consultants with inadequate experience. The fast-track production of consultants is compounded by the effect of the European Working Time Directive. Previously, surgeons underwent approximately 30,000 training hours before becoming consultants—this will now decrease to 6000. There is no clear guidance on how these MMC consultants are to gain adequate experience, or who will supervise their ongoing professional development.

MMC and MTAS have suffered as a result of inadequate planning and expediting a system that was not thoroughly tested. Increased numbers of medical students combined with poor workforce planning have resulted in a massive shortfall of jobs. Doctors have been systematically disempowered and are being herded through training to produce a consultant-led service. But, worryingly for patients, these doctors will not have the experience of today’s consultants. MMC—or “mass medical culling” as it has now become known—has betrayed a generation of doctors and the society they serve. Any other employer would cherish these trainees for their dedication, effort, and skills. Doctors in the NHS are now re-thinking their choice of career, with some moving abroad or changing profession.

The independent review of the current crisis is welcomed, although there is little time to implement changes for the second round of applications. There is no doubt that the former training system was in need of an update. However, MMC and MTAS should be re-evaluated, with the focus on training and employing skilled doctors to provide high-quality care of patients.

Doctors and the Royal Colleges should reclaim autonomy of training, uniting against the debacle that is MMC. Sadly, only 10% of trainees feel represented by the Royal Colleges and 6% by the British Medical Association. Remedy UK—a movement of doctors who lack confidence in the current reforms—is leading marches in protest against MMC/MTAS in London and Glasgow on March 17. At the time of going to press, 8000 doctors, including consultants, have signed up for the demonstrations. Perhaps the MMC/MTAS furore will be the stimulus to finally unite doctors in taking an active lead to value professionalism in the NHS.
Reforming research in China

After several allegations of scientific misconduct in China last year, 2007 appears to be starting better. The Ministry of Science and Technology has announced measures to address misconduct in state-funded programmes, and the respected Chinese Academy of Sciences published a Declaration of Scientific Ideology on Feb 26 and also plans a commission for scientific integrity. Such initiatives will be welcome both in China and abroad—provided that they are enforced with the same transparency and equity that they demand of researchers.

Why Chinese research should find itself in such a predicament is in part a result of the country’s phenomenal economic success. The Atlas of Ideas, a study of scientific research in Asia published by the UK think-tank DEMOS in January, 2007, summarises events. As in the heyday of US science in the 1960s, when the USA rallied to President Kennedy’s goal of sending a man to the moon, China now has resources, talent, and a vision: President Hu Jintao’s goal of becoming a research superpower by 2020. Most observers doubt that it will take that long. With funding increasing by 20% a year, China now has more investment in research and development than any other country except the USA.

But high investment brings high expectations. While all science needs adequate funding, to thrive it also requires openness, collaboration, and freedom. Unrealistic expectations can lead to stress, poor quality work, plagiarism, and fraud. All of these, however uncommon, have tarnished the reputation of science in China and show the need for more effective governance.

The Chinese Academy of Sciences’ announcement gives cause for optimism that the ancient tradition of science in China can realign its intellectual goals with contemporary social values, and that ethical principles can guide all stages of research. If grant awards are transparent, institutional review boards invite debate, participant autonomy is respected, and analysis and presentation of findings is accountable, then China has the opportunity to lead the world not only in research quantity, but also in quality.

Promoting optimum management for ADHD

Attention-deficit hyperactivity disorder (ADHD) in children—a complex disorder characterised by developmentally inappropriate levels of inattention, impulsivity, and hyperactivity—has received considerable attention in the lay press, partly because of controversy over its treatment with stimulant drugs. Last week, interest in ADHD treatment returned to the mainstream media following the publication of new research in the journal Health Affairs. The study showed that the global use of medications (including non-stimulant drugs) to treat ADHD in children aged 5–19 years has risen threefold between 1993 and 2003. In 1993, 31 countries belonging to the Organization for Economic Cooperation and Development had adopted the use of ADHD medications but by 2003 the number had grown to 55. Some nations, such as Italy, Sweden, and Japan, had lower use than would be expected by gross domestic product per head, whereas the USA, Canada, and Australia had higher than predicted use.

Possible explanations for the varying rates of medication use include direct-to-consumer advertising in the USA, differing diagnostic criteria for the disorder, and variations in national health systems.

Increases in the use of medications for ADHD raise alarms about the possible over-identification and over-treatment of the disorder. Stimulant drugs—the first-line treatment for severe cases of ADHD—have side-effects, they can be addictive, and their use can be abused. But under-recognition is also of concern in some countries, where cases actually diagnosed and treated by clinicians remain far below the population prevalence. Untreated ADHD can lead to poor self-esteem, academic achievement, peer and family relationships, and increases in accidental injuries.

Detailed, accurate, and comprehensive assessment by well-trained specialists is crucial for proper diagnosis of ADHD. Behavioural treatment should be used for less severe cases and can be used alongside medication in more severe cases. Identifying the cultural and economic factors that lead to over-treatment in some countries and under-treatment in others, will also be important for promoting optimum management of the disorder.
Interventional cardiologists have several choices of anticoagulation regimen for patients undergoing percutaneous coronary intervention. Bivalirudin is a direct thrombin inhibitor that reduces bleeding complications without increasing ischaemic events, compared with heparin plus glycoprotein IIb/IIIa inhibitors, in stable non-high-risk patients. In today’s *Lancet*, Gregg Stone and colleagues present a subgroup analysis of 7,789 patients from the Acute Catheterization and Urgent Intervention Triage strategy (ACUITY) trial, to investigate whether these properties can be expanded to a more complex subset of patients, including those with moderate and high-risk coronary syndromes.

The main findings suggest that, in patients with acute coronary syndrome undergoing contemporary percutaneous coronary intervention, bivalirudin alone can safely replace heparin plus glycoprotein IIb/IIIa inhibitors, with similar rates of ischaemia and less bleeding at 30 days; whereas patients given bivalirudin plus glycoprotein IIb/IIIa inhibitors had similar rates of bleeding and ischaemic events to those taking heparin plus glycoprotein IIb/IIIa inhibitors. These findings mirrored those of the ACUITY and the REPLACE-2 studies, but in a more complex population.

The study has deficiencies that preclude it from being definitive. Randomisation was not stratified by treatment assignment, since it was done before angiography. This approach, in conjunction with the use of an open-label design, leaves concern for potential imbalance of groups or treatment. Furthermore, different glycoprotein IIb/IIIa inhibitors and heparin agents were used in the control arm. The level of patients’ risk was not as high as in the SYNERGY or OASIS 5 trials, and patients with ST-elevation myocardial infarction were not included. Finally, the analysis was not powered for non-inferiority testing and therefore the comparability or non-inferiority status of bivalirudin remains a question. Although the troponin-negative patients and the patients who were pretreated with thienopyridine did well on bivalirudin monotherapy, this post-hoc analysis is not conclusive.

As in the previous trials of bivalirudin, the investigators combined safety (bleeding) and efficacy (ischaemic events) endpoints. Such combination is subject to ongoing criticism, since the nominal decrease in bleeding events that will offset an increase in ischaemia is unknown. Both endpoints have been reported to affect late clinical events, including mortality. The composition of safety and efficacy measures is not ideal, because drugs that are ineffective but safe can be made to seem better than effective drugs in a non-inferiority (and even superiority) trial. In bivalirudin studies, bleeding reduction always drives the composite endpoints, thus obscuring differences in efficacy. Investigators should report separate endpoints for safety and efficacy and refrain from adding them to comprise net clinical gain, especially at 30 days.

In the troponin-positive patients, ischaemic complications did not significantly differ between the heparin plus glycoprotein IIb/IIIa inhibitor group and the bivalirudin-only group, yet a 1% absolute increase (relative risk 1·12) in ischaemic composite events was noted. At the same time, however, the incidence of major bleeding was reduced in this subgroup by 2·8% (relative risk 0·59). The increase in the ischaemic composite was not statistically significant, but whether patients with a positive troponin or ST-elevation myocardial infarction would still benefit from glycoprotein IIb/IIIa inhibitors is unknown.
Comment

Antagonists remains uncertain. Yet, when glycoprotein IIb/IIIa antagonists are added to bivalirudin, the latter's edge over heparin with regard to bleeding reduction disappears.

About a third of the population in the subgroup analysed was not pretreated with thienopyridine. For the clopidogrel-preloaded patients, rates of bleeding were lower and rates of ischaemic events were similar on monotherapy with bivalirudin, compared with heparin plus glycoprotein IIb/IIIa inhibitors. In patients not on clopidogrel, bleeding was significantly reduced by 2.2% (relative risk of 0.61), but ischaemic events were increased by 2.8% (relative risk of 1.37) with bivalirudin monotherapy. This analysis was not prespecified in the protocol, and caution against overinterpretation is warranted. Whether bivalirudin plus P2Y12 blockade has a complementary role is the subject of continuing research, and until more conclusive evidence is available, any decision on bivalirudin use without thienopyridine is based on speculation.

High-risk patients may need an increased dose and duration of antiplatelet therapy. For patients loaded with thienopyridine immediately before or during percutaneous coronary intervention, the metabolised drug may not be available in the blood during the procedure. Furthermore, post-procedure, when the effect of bivalirudin has worn off (because of its short half-life), anticoagulation might not be sufficient in a high-risk patient. Whether these patients should be treated with glycoprotein IIb/IIIa inhibitors in addition to bivalirudin or heparin remains unclear.

The advantage of bivalirudin shown by these findings is attributable solely to reduced bleeding rate. Bivalirudin monotherapy can safely be expanded to patients with acute coronary syndromes undergoing percutaneous coronary intervention. However, for high-risk patients, other treatment options should not be excluded and the use of glycoprotein IIb/IIIa inhibitors should be considered, especially for those with positive troponin and for those not pretreated with thienopyridine.

The optimum anticoagulant treatment strategy for those with ST-elevation myocardial infarction remains unknown and perhaps one drug does not fit all. The 1-year data from this analysis may lend support to the practice of switching to bivalirudin monotherapy, but, in the meantime, we should acknowledge that other combinations could work as well. Other considerations, such as timing of invasive strategy, cost, complexity of the lesion, and the patient's characteristics, might determine the best approach, and should be studied prospectively with prespecified endpoints that differentiate between safety and efficacy.

Ron Waksman
Washington Hospital Center, Washington, DC 20010, USA
ron.waksman@medstar.net
I have received consulting fees from Biotronik, Medtronic, Boston Scientific, and research grants from the Medicine Company, GlasoSmithKline, Schering-Plough, and Sanofi-Aventis.


Cardiac arrest—guideline changes urgently needed

Cardiopulmonary resuscitation is traditionally defined as chest compression and ventilation. The need for chest compressions is unquestionable, but the need or advisability of intermittent ventilation for out-of-hospital non-respiratory primary cardiac arrest has been very controversial.1 The article by the SOS-KANTO study group in today’s Lancet is of the utmost importance because it provides evidence in human

See Articles page 920
beings that chest compression without ventilation is preferable for out-of-hospital cardiac arrest. This finding is an important piece of evidence that should lead to a prompt interim revision of the guidelines for out-of-hospital cardiac arrest. Eliminating the need for mouth-to-mouth ventilation will dramatically increase the occurrence of bystander-initiated resuscitation efforts and will increase survival.

A major problem with most studies of out-of-hospital cardiac arrest is that most patients studied have absolutely no chance of surviving. Therefore a better approach to resuscitation is almost impossible to show. Nevertheless, many studies have identified a subgroup of patients with the greatest chance for survival: those who have a witnessed cardiac arrest and a shockable rhythm when the emergency team arrives. Within this important subgroup in the SOS-KANTO study, 22% of those who received bystander-initiated chest-compression alone had a favourable neurological outcome, as compared to 10% of those who received 2000 AHA and ILCOR guideline-recommended chest compressions and mouth-to-mouth breathing.

From their inception, the standards and guidelines for a bystander’s response to out-of-hospital cardiac arrest have always emphasised the imperative of mouth-to-mouth ventilation, paradoxically called “rescue breathing.” While mouth-to-mouth ventilation may “rescue” an individual with respiratory arrest, this approach actually decreases the likelihood of a “rescue” in a much larger group of patients—those with a primary cardiac arrest.

Mouth-to-mouth ventilation for primary cardiac arrest is detrimental for several reasons. First, this requirement greatly decreases bystander-initiated resuscitation efforts, an important determinant of survival from out-of-hospital cardiac arrest. Second, studies have long reported that survival is better in individuals with cardiac arrest who receive chest compression only than it is in those in whom no bystander rescue efforts were started until the actual or simulated arrival of emergency personnel. Third, mouth-to-mouth ventilations by single bystanders requires inordinately long interruptions of essential chest compressions (figure). Fourth, during cardiac arrest, mouth-to-mouth or positive-pressure ventilation reduces the already marginal coronary and cerebral blood flow during cardiac arrest and resuscitation. This situation is made worse when forceful ventilation is given while the chest is being compressed. Fifth, with sudden unexpected primary cardiac arrest, ventilations are initially neither necessary nor logical, for with the onset of ventricular-fibrillation-induced arrest, the pulmonary veins, the left heart, and the entire arterial system are filled with oxygenated blood and the recommended ventilations

---

**Figure:** Aortic and right atrial pressures during chest compressions in anaesthetised swine model of cardiac arrest due to ventricular fibrillation

Coronary perfusion pressure (arrows) = difference between aortic (Ao, red) diastolic and right atrial (RA, blue) pressures. Cerebral perfusion pressure is difference between carotid systolic pressure and intracerebral venous pressure generated by each chest compression. Cerebral blood flow occurs because of closure of venous valves or collapse of veins at thoracic outlet with each chest compression, which results in positive pressure gradient across brain. Top: each 15 chest compressions are interrupted for about 16 s to simulate duration of interrupted chest compressions required for single bystander to deliver two mouth-to-mouth ventilations and return to chest compressions. Bottom: uninterrupted compressions, note arterial pressures are continuous, resulting in near continuous cerebral perfusion.
do not increase arterial saturation—they only further delay the onset of critical chest compressions.\(^1\)\(^2\)\(^3\) Sixth, mouth-to-mouth ventilation is not necessary in a significant number of victims of witnessed cardiac arrest because they initially gasp, and if chest compressions are started early and continued, many of these patients will continue to gasp and thereby provide physiological ventilation (eg, that with decreasing intrathoracic pressures that facilitates venous return to the chest).\(^4\)\(^5\)\(^6\) Seventh, survival from experimentally induced cardiac arrest is better with higher coronary perfusion pressures produced by forceful chest compressions.\(^7\)\(^8\)\(^9\) Eighth, in non-paralysed animals in cardiac arrest, survival is dramatically better with chest-compression-only resuscitation than with ventilations plus chest compressions, when chest compressions were interrupted for a realistic 16 s to provide the two mouth-to-mouth breaths between each set of 15 chest compressions (figure).\(^1\)\(^2\)\(^3\) The SOS KANTO study has now shown this result in human beings.\(^2\)

We have recommended cardiopulmonary resuscitation by bystander chest-compression-only for out-of-hospital cardiac arrest for years.\(^1\)\(^2\) More recently this approach has been incorporated into Cardiocerebral Resuscitation, a new approach to resuscitation of victims of cardiac arrest that eliminates early positive-pressure ventilation by emergency personnel, emphasises continuous chest compressions and improves survival.\(^1\)\(^2\)\(^3\)\(^4\)

A major flaw with the current, and all previous, guidelines for cardiac arrest is that they recommend the same approach of cardiopulmonary resuscitation for two entirely different clinical conditions: primary cardiac arrest where the arterial blood is well oxygenated at the time of the cardiac arrest, and respiratory arrest when the arterial blood is so severely desaturated that it contributes to hypotension and secondary cardiac arrest. We should continue, for now, to follow the newer guidelines of assisted ventilations and chest compressions for respiratory arrest (such as in drowning or drug overdose), but the guidelines should promptly be changed to chest-compression-alone for witnessed unexpected sudden collapse (a condition that is, in all probability, cardiac arrest).\(^3\)\(^4\)

The critically important findings by the SOS-KANTO group should lead to changes in guidelines. Advocating, encouraging, and teaching chest-compression-only for witnessed unexpected sudden collapse will dramatically increase bystander-initiated resuscitation efforts and thereby give these patients a better chance of survival when emergency personnel arrive. We should continue instructions in cardiopulmonary resuscitation for the equally important, but less frequent, occurrences of drowning and other forms of respiratory arrest.

Gordon A Ewy
Cardiology, University of Arizona College of Medicine, Tucson, AZ 85724, USA
gaewy@aol.com

I declare that I have no conflict of interest.

Effects of zinc supplementation on child mortality

More than 10 million children younger than 5 years die every year in the developing world, mostly from preventable infections. Malnutrition and micronutrient deficiencies are increasingly recognised as the main risk factor for childhood mortality in these countries. Micronutrients are important determinants of infection and their outcomes. Zinc is a crucial micronutrient because it affects various immune mechanisms and modulates host resistance to several pathogens. Zinc supplementation reduces morbidity from diarrhoea and pneumonia in high risk populations. A large body of evidence shows important therapeutic benefits with zinc administration during and after diarrhoea, and some studies also reported reduction in diarrhoea morbidity in the subsequent 2–3 months without further supplementation.

Assessment of the effect of improved zinc intake on child mortality is important for the development of effective public-health policy. In today’s Lancet, Sunil Sazawal and colleagues report an overall non-significant 7% reduction in mortality in a randomised trial of daily zinc supplementation in children aged 1–48 months residing in Pemba, Zanzibar. A significant reduction in mortality was noted in children aged 12–48 months which, although from a subgroup analysis, is plausible. The overall effect on mortality is less than what one might have expected, and is somewhat surprising if we consider the significant reduction in diarrhoea and pneumonia morbidity that has previously been seen in children older than 6 months. One would assume, but cannot be certain, that the mechanisms underlying reduction of severe morbidity in zinc-supplemented children must also be relevant for mortality reduction. Although Sazawal’s study was complex, it was well designed and has been effectively implemented and monitored. It was stimulated by smaller earlier trials that suggested some benefit of zinc supplementation on mortality (table). An integrated view of all mortality studies is complicated by variations in doses of zinc used, study populations, background morbidity, and administration of vitamin A supplements (table). Two studies focused on small-for-gestational age or low-birthweight infants, and one focused on children with diarrhoea.

Sazawal and colleagues’ study raises some important policy issues. A review of available studies of zinc supplementation on morbidity, growth, and mortality, including Sazawal’s study, is consistent with the view that zinc deficiency has important health consequences and needs attention. Could the varying doses of zinc used for these mortality studies have affected the results? Significant beneficial effects on morbidity were noted with large doses in some earlier studies, and these effects were not restricted to children with low baseline concentrations of zinc in plasma or serum. One is tempted to speculate that larger doses than those used in the trial in Zanzibar could be more beneficial in the reduction of mortality than was actually seen. An obvious lesson is that increased emphasis should be

<table>
<thead>
<tr>
<th>Study population</th>
<th>Dose and duration of zinc supplementation</th>
<th>Zinc formulation</th>
<th>Effect size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>154 full-term, small-for-gestational-age infants</td>
<td>5 mg per day, follow-up for 30–284 days</td>
<td>Zinc sulphate as syrup</td>
</tr>
<tr>
<td>Brazil</td>
<td>205 full-term low-birthweight infants</td>
<td>5 mg, 1 mg, or placebo daily for first 8 weeks of life, follow-up for 0–26 weeks</td>
<td>Zinc sulphate as syrup vs placebo: HR 0·53 (0·13–2·14); 1 mg vs placebo: HR 0·5 (0·13–2·01)</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>870 children aged 3–59 months with acute diarrhoea</td>
<td>20 mg per day for 14 days in addition to ORS</td>
<td>Zinc acetate as syrup</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1665 (unselected) children, aged 60 days to 12 months</td>
<td>70 mg once weekly</td>
<td>Zinc acetate as syrup</td>
</tr>
<tr>
<td>Burkina-Faso</td>
<td>709 (unselected) children, aged 6–31 months</td>
<td>12.5 mg per day, 6 days a week for 6 months</td>
<td>Zinc sulphate as tablets</td>
</tr>
<tr>
<td>Pemba</td>
<td>47546 (unselected) children aged 1–36 months</td>
<td>10 mg per day (5 mg per day in children aged &lt;12 months), given until age 48 months (mean 485 [SD 307] days)</td>
<td>Zinc sulphate as tablets</td>
</tr>
</tbody>
</table>

HR=adjusted hazard ratio, ORS=oral rehydration salts. *Adjusted rate ratio. †Relative risk.

Table: Randomised placebo-controlled clinical trials assessing effects of zinc supplementation on child mortality
given to studying specific mechanisms by which zinc acts, and whether the benefits are related to correction of zinc deficiency or to favourable effects of zinc at increasing concentrations. Such knowledge could guide definition of a rationale for a daily optimum dose. We should also examine the effect of daily zinc supplementation on plasma zinc concentrations, and the bioavailability of different doses and different formulations (zinc salts or syrup vs tablets).

Future research should also focus on low-birthweight infants, to define the burden of zinc deficiency and the potential benefits of survival through improved zinc intake.

In the interim, the present WHO strategy to focus on introduction of zinc for treatment of diarrhoea is an important step forward. The administration of zinc with oral rehydration salts for diarrhoea in the programme settings resulted in increased use of these salts, decreased use of antimicrobials and antidiarrhoeals, and reduction in hospital admissions. In the long-term, measures to improve zinc intake of children, such as improvement of the overall diet, supplementation, food fortification, and sub-selection of crops with improved zinc content need to be explored and assessed.

Shinjini Bhatnagar
Centre for Diarrhoeal Diseases and Nutrition Research, Department of Pediatrics, All India Institute of Medical Sciences, New Delhi 110 029, India
shinjini_bhatnagar@rediffmail.com

Objectification of physicians and loss of therapeutic power

Over the past 20 years, doctors have been sensitive to accusations of vested interest, laudably keen to descend from their pedestals, understandably reluctant to defend the importance of status, and surprisingly willing to accept the argument that increasing autonomy and power for patients means that the power and agency of doctors must be reduced. In reality, most human interactions increase the agency of both parties, and interactions between doctor and patient are no exception. The task of doctor and patient is to work together to achieve greater power and control over disease, and the agency of both parties is increased when doing so.

Within contemporary society, the ascendancy of market rhetoric has made it increasingly difficult to defend two major prerequisites of the therapeutic efficacy of physicians: professional authority, and a continuing relationship of mutual respect and trust between patient and doctor. The importance of the subjectivity of the doctor is being minimised, and, through a process described as objectification of the physician, the role of the doctor is being diminished to one of a competent technician who is interchangeable with any other with similar training. In a less explicit way, the same process is applied to patients such that they become interchangeable units of health need,
whose access to the health-care system is prioritised
over the need to sustain a relationship with a known
and trusted professional.

These trends serve the interests of the politically
powerful, because a system in which agents are
interchangeable is much easier to organise and to
control, and the interests of global capital, because
markets are maximised wherever consumers and em-
ployees can be standardised. Doctors find themselves
agents of medicalisation, promoting technological
and pharmacological treatments and procedures,1
instead of using the healing potential of an ongoing
doctor–patient relationship. The medicalisation of
human distress in the current epidemic of depression4
is just one example.

Medicine aims to restore a sense of wholeness such
that the sick body no longer impedes the aspirations
of the self. Technical expertise is essential, but not
sufficient, and doctors must also develop their capacity
for “mending through meaning”.5 This involves
purposeful dialogue between doctor and patient with
the aim of discovering a coherent explanation of illness
that makes sense to the affected person. This, in turn,
helps to reduce fear, uncertainty, and confusion, and
by making room for optimism, boosts the body’s innate
capacity for mending.

The meaning attached to the colour6 or form of drugs,
the reputation of medicines as a result of advertising7
or other types of publicity, the relationship between
patient and physician,8 the physician’s attitude to the
patient, and the physician’s confidence in the efficacy
of treatment can all have a measurable effect on
outcome.9 For generations, doctors have deliberately
cultivated continuing relationships with their patients
based on an intentional positive regard for every
patient, knowing intuitively and from experience that
relationships which both patient and doctor enjoy
produce better health outcomes.10

Trends that minimise the importance of the
subjectivity of both doctor and patient in conventional
medicine seem to be driving patients to seek therapeutic
relationships with practitioners of complementary
medicine: demand for such care has increased
substantially, leading to pressure for such alternatives to
be funded publicly by national health services. As a result,
technological and intersubjective therapeutic power
come divided between different roles, rather than
combined and maximised such that the intersubjective
can be used systematically to maximise the effectiveness
of scientific technologies of proven benefit.

The irony is that the current profligate destruction of
the therapeutic capacity of doctors is occurring just as
science itself is beginning to explain the mechanism of
areas of medical practice that those who disparage the
importance of continuity of care regard as unscientific
and therefore redundant. As psychoneuroimmunology
gradually unpicks the intricate interactions between
the immune system, the autonomic nervous system,
and the subjective experience of the individual, the
more we can understand the power of positive and
negative psychological states to promote or undermine
the healthy functioning of the body.11 Supportive
personal relationships—including, we must presume,
relationships between doctors and patients—diminish
negative emotions and promote health through a
positive effect on immune function.12

*Iona Heath, John Nessa
Caversham Group Practice, London NW5 2UP, UK (IH); and
Institute of General Practice, University of Bergen, Norway (JN)
iona.heath@dsl.pipex.com

We declare that we have no conflict of interest.

1 Nessa J, Malterud K. Tell me what’s wrong with me: a discourse analysis
approach to the concept of patient autonomy. J Med Ethics 1998; 24:
394–400.
2 Ainsworth-Vaughn N. Claiming power in doctor–patient talk. Oxford:
3 Moynihan R, Henry D. The fight against disease mongering: generating
pmed.0030191.
Podoconiosis (endemic non-filarial elephantiasis) is a geochemical disease in individuals exposed to red-clay soil derived from alkalic volcanic rock. It is a chronic and debilitating disorder that, although rarely seen outside endemic regions, is a serious public-health problem in at least ten countries in tropical Africa, Central America, and north India. It is unusual in being an entirely preventable non-communicable disease, and has been eradicated from Scotland, France, and the Canary Islands since footwear became routine.

In Ethiopia, where podoconiosis has been best described, prevalence is about 5% in areas of irritant soil. Podocnosis is thus more common than HIV infection in these areas. About 11 million people (18% of the national population) live in endemic areas in Ethiopia, and between 500 000 and 1 million people are affected nationwide. Up to 64% of affected individuals are from the most economically active age-groups, and direct and productivity costs of podoconiosis in a group of 1·5 million inhabitants have been estimated at US$16 million a year.

Social stigma towards people with podoconiosis is pronounced, leading to exclusion from school and religious and community gatherings, and a bar on marriage into non-affected families.

The epidemiological and clinical features are distinct and well described—the disorder occurs in barefooted farmers, weavers, and other occupational groups where exposure to volcanic soil is common. The disorder now occurs wherever irritant soils coexist with high altitude, high rainfall, and extremely low income, predominantly where subsistence farmers cannot afford shoes, socks, or even water to wash their feet.

In Ethiopia, where podoconiosis has been best described, prevalence is about 5% in areas of irritant soil. Podocnosis is thus more common than HIV infection in these areas. About 11 million people (18% of the national population) live in endemic areas in Ethiopia, and between 500 000 and 1 million people are affected nationwide. Up to 64% of affected individuals are from the most economically active age-groups, and direct and productivity costs of podoconiosis in a group of 1·5 million inhabitants have been estimated at US$16 million a year.

Social stigma towards people with podoconiosis is pronounced, leading to exclusion from school and religious and community gatherings, and a bar on marriage into non-affected families.

The epidemiological and clinical features are distinct and well described—the disorder occurs in barefooted farmers, weavers, and other occupational groups where exposure to volcanic soil is common. Affected populations tend to reside at high altitudes, with a geographical correlation with alkalic-type volcanoes. Men and women are equally affected, and although most patients develop signs and symptoms between the ages of 10 and 30 years, individuals as young as 4 years could show early signs. Not all individuals exposed to irritant soil develop podoconiosis, and family clustering has long been observed. Genetic studies show high heritability of the trait, and segregation analysis suggests the presence of an autosomal codominant major gene conferring susceptibility to podoconiosis.

Disease is bilateral, but asymmetrical, and almost always below the knee (figure). Silica particles absorbed through the foot are thought to induce an inflammatory response in the lymphatic vessels, leading ultimately to fibrosis and obstruction of the vessel lumen. This event leads initially to oedema of the foot and lower leg, which progresses to severe elephantiasis, resembling infection seen in *Wuchereria bancrofti*. However, podoconiosis occurs at altitudes higher than 1500 m, which exceeds that at which filarial transmission occurs, and both midnight blood sampling and filarial antigen testing have consistently excluded filariasis as the cause of podoconiosis.
Remarkable improvements can be achieved in the early and progressive stages of disease if foot hygiene with soap, water, and antiseptics is combined with daily use of robust shoes and socks. Podoconiosis management thus shares many features with that of filarial lymphoedema. Care at advanced stages is more difficult, and long-term outcomes seem to improve after simple nodulectomy rather than radical surgery (Charles’ operation). Occupational rehabilitation is vital to provide patients with a social role and a means to earn a living that does not involve contact with irritant soil.

Podoconiosis is a common, but barely recognised, non-communicable tropical disease. The communities in which podoconiosis occurs are by definition among the poorest in the world, and the disorder intensifies the existing economic, and social burdens.

*Gail Davey, Melanie Newport
Faculty of Medicine, Addis Ababa, Ethiopia (GD); and Brighton & Sussex Medical School, Brighton, UK (MN)

erurkar@ethionet.et

We are coprincipal investigators of a project investigating the genetics of podoconiosis, funded by the Wellcome Trust; otherwise we declare that we have no conflict of interest.


**COPD—call for papers**

Chronic obstructive pulmonary disease (COPD) is a common chronic disease that remains underdiagnosed and undertreated, and prevalence is increasing in many countries. To coincide with this year’s annual European Respiratory Society meeting in Stockholm, Sweden, on Sept 15–19, The Lancet is planning to publish a special issue highlighting COPD.

We therefore issue a call for research papers on any aspect of COPD. We are particularly interested in publishing results from clinical trials and other studies that will be presented at the meeting in Stockholm, but will also consider other research papers on COPD. If your study’s results have been accepted for presentation at the meeting, please let us know, so that we can plan publication, possibly online, to comply with any embargo policies.

Papers should be submitted online by May 18, 2007, at the latest, and the covering letter should state that the submission is in response to this call for papers.

Sabine Kleinert
The Lancet, London NW1 7BY, UK

To submit research go to [http://ees.elsevier.com/thelancet](http://ees.elsevier.com/thelancet)
Clinical update: the low-glycaemic-index diet

For most of the past half century, the US Government and many official health agencies have advocated a low-fat diet for the prevention and treatment of obesity, diabetes, and heart disease. The rationale seemed obvious: fat is energy dense and tasty, making this nutrient easily overconsumed; and saturated fat adversely affects LDL-cholesterol and insulin resistance. At the time, popular diet books argued that if you do not want fat on your body or in your coronary arteries, do not eat it, and instead fill up on carbohydrate. However, the effects of total fat on bodyweight and health have been called into question in recent years.1,2 And the prevalence of obesity has increased strikingly in the USA and Europe, as the proportion of fat in the diet has decreased. During the past decade, the pendulum swung far in the other direction with the very-low-carbohydrate diet. Clinical trials reported significant short-term weight loss on diets containing up to 60% energy from fat. Many diet books were published promising hunger-free weight loss by filling up on fat and avoiding carbohydrates. However, results from longer-term studies indicated substantial weight regain after 6 months,3 and the popularity of this approach is declining. Recently, the glycaemic index (GI) has attracted considerable interest with the publication of research linking this novel dietary factor to important health outcomes. Many current best-selling diet books (The South Beach diet, The zone diet, The new glucose revolution) advocate consumption of a low-GI diet. Is this yet another in a long succession of popular diets destined to rise and fall rapidly? Or will the low-GI diet have enduring merit?

Abandoning terms such as simple sugar and complex carbohydrate that do not correlate with postprandial glycaemia, the GI constitutes an empirical system for classifying carbohydrate-containing foods.4-6 GI is determined by measuring the 2-h incremental area-under-the-blood-glucose-curve after consuming a test food (containing 50 g available carbohydrate) relative to that of a control, either white bread or glucose. Most varieties of bread, rice, breakfast cereals, and potato products have a high GI because processing methods allow the starch to become fully hydrated and therefore rapidly hydrolysed into glucose in the human digestive tract. By contrast, non-starchy vegetables, legumes, nuts, and fruits have a low GI. Whole kernel and traditionally processed grain products, such as stoneground breads, steelcut oats, and pasta, tend to have a moderate GI. A related term, the glycaemic load (GL)—average GI multiplied by carbohydrate amount—takes into account differences in carbohydrate content among foods, meals, or diets.7

The metabolic requirement of the human brain for glucose provides a theoretical basis for understanding the importance of GI to health. Hormonal regulatory systems have evolved to maintain stable concentrations of blood glucose under various conditions, such as fasting or feasting, consumption of foods with varying nutrient composition, varying levels of physical activity, illness, and pregnancy. However, one environmental condition rarely encountered before the modern era is the wide availability of high-GI foods. Before the agricultural revolution, human beings did not often consume grain products and concentrated sugars.8 With domestication of cereal grains, the GI of human diets increased substantially. In the past few decades, prevailing diets in the USA and Europe have become even higher in GI and GL because of increases in carbohydrate consumption and the processing of that carbohydrate.

As previously described,9 a high-GI diet elicits a sequence of hormonal events that challenge glucose homeostasis. Shortly after a high-GI meal, blood insulin level rises higher than after a low-GI meal with similar nutrients. Conversely, a high-GI meal inhibits glucagon secretion. The strikingly increased insulin:glucagon ratio constitutes a powerful anabolic stimulus, promoting uptake of nutrients at liver, muscle, and fat, and suppressing hepatic glucose output. Within 60 min after a high-GI meal, blood glucose begins to fall, often reaching levels below fasting, and release of fatty acids from adipose tissue is suppressed. The combination of rapidly declining blood glucose and low concentrations of non-esterified fatty acids stimulates hunger and overeating, in the body’s attempt to restore the concentration of metabolic fuels to normal. In addition, the early postprandial hyperglycaemia and hyperinsulinaemia and the late postprandial hypoglycaemia and counter-regulatory hormone response could adversely affect body composition, and increase risk for diabetes, cardiovascular disease, and cancer.

Since Jenkins proposed the GI in 1981,4 several hundred observational analyses, clinical trials, and animal studies addressing this concept have been published. Most report
beneficial effects of a low-GI/GL diet on health, although a significant minority report no health effects of GI or GL. Of particular relevance, virtually no studies have found health benefits with a high-GI/GL diet. The webtable summarises clinical outcomes that have been examined. The main areas of investigation have been: diabetes-related (glycated blood proteins, C-peptide, fasting blood glucose and insulin, insulin resistance, second-meal effect, β-cell function, adiponectin, risk of type 2 diabetes, gestational diabetes risk, quality of life); cardiovascular disease-related (blood pressure, serum lipids, systemic inflammation, oxidative stress, coagulopathy, metabolic syndrome, atherosclerosis risk, myocardial infarction risk, stroke risk, duration of hospital stay); obesity-related (hunger, bodyweight, body fatness, energy expenditure, substrate oxidation); cancer-related (breast, ovary, uterus, prostate, gastrointestinal tract); and perinatal-related (large-for-gestational-age infant, neural tube defects). Other topics of study have included fatty liver, gallbladder disease, eye disease, dental carries, epilepsy, mental function, and exercise performance. Existing reviews give more detailed analyses.9,10 Further evidence of the benefits of delaying carbohydrate absorption comes from studies of α-glucosidase inhibitors in the prevention of diabetes and heart disease.11

The WHO/FAO (Food and Agriculture organisation of the UN) and some health agencies in Europe, Canada, and Australia advocate consumption of a low-GI diet.12–15 However, no governmental agency or major professional association in the USA does so. Indeed, the clinical relevance of GI remains the subject of a debate centred on three points.

First, are the data consistent? Since the inception of the field 25 years ago, researchers’ understanding of the relevant methodological issues and physiological mechanisms has evolved substantially. Therefore some inconsistency among findings is to be expected. Specific causes of inconsistency between studies include: measurement error attributable in part to use of too few participants for determination of GI; confounding, such as by other dietary factors; misclassification in observational research; lack of treatment fidelity in clinical trials, with some failing to show significant difference in GI between treatment groups; inadequate study power; and physiological differences involving, for example, how much insulin an individual secretes after carbohydrate ingestion.16,17

Second, how does glycaemic response change when several foods are eaten in combination? Some researchers report that the GI of individual foods eaten alone does not predict blood glucose response after those foods are eaten as part of a mixed meal.18,19 Others argue that GI predicts glycaemic response with excellent accuracy under usual conditions (except for meals with very-low-carbohydrate content), when appropriate methodology is used.9,20 In any event, the issue may be moot if observational and interventional studies show improvement in health in individuals consuming self-prepared low-GI compared with high-GI diets over the long term. Moreover, the mixed-meal effect can be used to advantage, for example by eating high GI foods (a bagel) together with protein, fat, and fibre (peanut butter) to lower glycaemic response.

Third, is a low-GI diet practical? A concern is that the concept of GI may be too complicated to be useful for the general public. The GI of a food can change depending on production and preparatory methods, and some lower GI foods do not appear to be particularly healthy (eg, ice cream). Indeed, a healthy diet could never be selected from a simple numbering system, be it GI or any other single dietary factor. Few advocates of a low-fat diet would encourage consumption of sugar-sweetened soft-drinks because those beverages have 0 g fat. Nevertheless, the principles of a low-GI diet can be summarised straightforwardly (figure). This dietary

![Figure: Low-glycaemic load pyramid](See Online for webtable)
pattern is consistent with current recommendations to increase consumption of vegetables, fruits, and minimally-processed grains, and has the advantages of being high in fibre and low in energy density.

Diets that restrict one major nutrient, either fat or carbohydrate, have produced poor long-term results. These diets might actually be counterproductive if they divert attention away from potentially more effective measures. The low-GI diet, with its focus on carbohydrate quality rather than quantity, aims to address an underlying physiological cause of diseases arising from excessive swings in postprandial glycaemia. Because this diet does not restrict either fat or carbohydrate, it may also be more behaviourally sustainable. Although the data are variable, most published studies report beneficial effects of lowering GI and virtually no study suggests potential for harm (by contrast with low-fat and very-low-carbohydrate diets that can adversely affect some risk factors for cardiovascular disease). Additional well-controlled and adequately powered studies are needed to examine the long-term effects of a low-GI diet on human health. Pending results of those studies, the clinician should consider a low-GI diet to be a prudent approach to the prevention and treatment of diabetes, heart disease, and obesity.

David S Ludwig
Department of Medicine, Children’s Hospital, Boston, MA 02115, USA
david.ludwig@childrens.harvard.edu

I am the author of a book on childhood obesity (Ending the food fight: guide your child to a healthy weight in a fast food/fake food world, Boston: Houghton Mifflin, 2007). I received support from grants 1R01 DK59240, 1R01 DK73428, and 1R01 DK73025 from the US National Institutes of Diabetes and Digestive and Kidney Diseases, and from the Charles H Hood Foundation. I thank Cara Ebbeling for help creating the figure.
WHO Director-General announces senior team

Last month, WHO Director-General Margaret Chan announced the names of her senior team. Among the 11 Assistant Director-General appointments and reappointments are three that will head new organisational divisions. Hannah Brown takes a look at the people and the changes.

Anarfi Asamoa-Baah
Deputy Director-General
In one of her first moves as Director General, Margaret Chan reinstated the position of Deputy Director-General, not used as a full-time position since August, 1992, with the appointment of Anarfi Asamoa-Baah from Ghana, announced ahead of the Assistant Director-General positions on Jan 10. Asamoa-Baah joined WHO in 1998 as a senior policy adviser to then Director-General Gro Harlem Brundtland. He progressed through a series of top jobs in the organisation, including Executive Director for External Relations and Executive Director for Health Technology and Pharmaceuticals, before being appointed Assistant Director-General for Communicable Diseases and then for HIV/AIDS, Tuberculosis, and Malaria in October, 2005. Before joining WHO, Asamoa-Baah had risen through the public-health ranks in his native Ghana, reaching the position of Director of Medical Services for Ghana in 1997. He is a medical doctor with a degree from the Ghana Medical School in Korle-Bu, and has done postgraduate stints in both the UK and USA, including gaining qualifications in community health, health economics, and public policy. According to WHO, Asamoa-Baah’s role as Deputy Director-General will involve assisting Margaret Chan in managing the organisation and raising WHO’s profile and leadership in health development.

Ala Alwan
Assistant Director-General—Health Action in Crises
Alwan will head one of three new clusters created in Chan’s first round of reorganisation at WHO. The new top-tier division elevates the Health Action in Crisis department, which Alwan has directed since 2005 when he took over from David Nabarro, who was seconded to New York as UN Systems Coordinator for Avian and Pandemic Influenza. Alwan came to Geneva from Iraq, his home country, where he was minister of health and education between 2003 and 2005, in the crucial years after the start of the Iraq war. Before taking this government job, Alwan was WHO’s country representative in Jordan; he has also held WHO posts in Geneva as Director of Noncommunicable Diseases Prevention and Director of Noncommunicable Diseases Management, in Oman as WHO’s country representative, and in the Eastern Mediterranean Regional Office as Director of the Division of Health Systems Development. After graduating in medicine from the University of Alexandria, Alwan moved to the UK to pursue further training and qualifications and then practised medicine in Scotland before returning to Iraq. Alwan held several academic, clinical, and public-health positions in Iraq, including professor and dean of the Faculty of Medicine at Mustansiriya University in Baghdad.

Catherine Le Galès-Camus
Assistant Director-General—Noncommunicable Diseases and Mental Health
Reappointed by Chan to the position she held under Lee Jong-wook, Catherine Le Gâles-Camus has watched over the Noncommunicable Diseases and Mental Health Cluster through a significant strengthening of WHO’s mandate and activities in these areas. During Gâles-Camus’s tenure, the WHO Framework Convention on Tobacco Control opened for signing; the World Health Assembly adopted the WHO Global Strategy on Diet, Physical Activity and Health and 13 other noncommunicable disease-related resolutions; and WHO launched the first global report on road traffic accidents and the first child growth standards. A French national, Galès-Camus has a PhD in economics from the University of Paris and did subsequent research into the economics of prevention; measurement of the state of health in populations, and the economic appraisal of medical technology. Before joining WHO, she was a scientific adviser to the Director-General of Health in France. She says she is strongly committed to promoting mental health and the reduction of the burden associated with mental and neurological disorders, including substance abuse.

Tim Evans
Assistant Director-General—Information, Evidence and Research
Tim Evans’ new cluster of Information, Evidence and Research superficially seems to have undergone only a slight change from its previous title of Evidence and Information for Policy, which includes WHO’s publishing operations. This group, which Evans had led since he was appointed by Lee Jong-wook in 2003, helps to “build the evidence base to help improve the process for health systems policy decisions”. Evans came to WHO after 6 years at the Rockefeller Foundation, where he was responsible for the development and implementation of programme strategy aimed at redressing disparities in health. Before that appointment, Evans held clinical
positions at Brigham and Women’s Hospital between 1992—after qualifying in medicine at McMaster University in his home country of Canada—and 1997, with a 2-year break to take up an assistant professor post at the Harvard School of Public Health. Evans initially studied for a Bachelor of Social Sciences degree from the University of Ottawa, Canada, and a DPhil in Agricultural Economics from the University of Oxford, UK. Evans was a founding board member of both the Global Forum for Health Research and the Global Alliance for Vaccines and Immunisation, now the GAVI Alliance.

David Heymann  
Assistant Director-General—Communicable Diseases  
David Heymann’s temporary charge of the Communicable Diseases cluster, which he took over from Chan when she put herself forward for the Director-General post, has been now been made permanent, although he retains the title of Representative of the Director-General for Polio Eradication. Heymann has had a long career association with WHO, which began with a 2-year stint in India working on the smallpox eradication programme in 1974–75. He spent the subsequent 13 years working as a medical epidemiologist in various countries of sub-Saharan Africa on behalf of the US Centers for Disease Control and Prevention. During this time, Heymann participated in the investigation of the first outbreak of ebola virus in 1976. Heymann first gained a BA from Pennsylvania State University before getting an MD from Wake Forest University and a Diploma in Tropical Medicine and Hygiene from the London School of Hygiene and Tropical Medicine, UK. He joined WHO headquarters in Geneva in 1988, and held several different posts in communicable disease control, including Executive Director of the Communicable Diseases Cluster, before becoming Assistant Director General.

Elizabeth Mason  
Daisy Mafubelu  
Assistant Director-General—Family and Community Health  
Since the departure of former Assistant Director-General Joy Phumaphi, who moved to the World Bank to become Vice President and Head of Human Development in February, Elizabeth Mason has been in charge of the Family and Community Health cluster, ad interim. Director of Child and Adolescent Health and Development since September 2004, Mason joined WHO in 1993, initially in the inter-country office for southern Africa. She held several posts within WHO’s African regional office, including as an adviser on the integrated management of childhood illness, before moving to the Geneva headquarters in 2004. Next month, however, Mason is due to be replaced permanently as Assistant Director-General for Family and Community Health by Daisy Mafubelu, currently health attaché for the Permanent Mission of South Africa to the UN. Mafubelu initially trained as a nurse, and spent 13 years working in that field, but she also has a degree in business administration and a post-graduate diploma in health management. She held several positions within the South African government before her move to Geneva.

Hiroki Nakatani  
Assistant Director-General—AIDS, Tuberculosis and Malaria  
Hiroki Nakatani came to WHO from the Japanese government where he was Director-General at the Department of Health and Welfare of Disabled Persons. According to WHO, during a long career at the ministry he acquired extensive technical experience in public health, including tuberculosis and HIV/AIDS, as well as immunisation, non-communicable diseases, health promotion, health emergencies, and health workforce development. His senior-level responsibilities included a focus on administration, management, and organisational and legislative development. Nakatani also participated actively in many international health initiatives. He was a member of the G8+Mexico Global Health Security Working Group and chaired the Chemical Events Working Group for 2 years. From 1988 to 1993, he was seconded to WHO in Geneva as a scientist in the Division of Development of Human Resources for Health. Nakatani received his MD from the Keio University School of Medicine in 1977 and his PhD from the Department of Hygiene and Public Health of Keio University in 2001.

Anders Nordström  
Assistant Director-General—Health Systems and Services  
Stepping back into an Assistant Director-General position after leading WHO as Acting Director-General for the 7 months before Chan’s term began on Jan 4, Anders Nordström is now heading the third new cluster of Health Systems and Services. This new department will focus on scaling-up health services, financing, information systems and human resources, and will aim to “provide global technical leadership throughout WHO”. Nordström’s first stint as an Assistant Director-General for General Management began in July 2003, when he was charged with making WHO a more effective and efficient organisation through strengthening internal processes, management, and accountability. However, his previous experience involves extensive field work. He graduated as a medical doctor from the Karolinska Institute in 1988 and took courses in political science and development studies at Stockholm University before taking a job with the Swedish Red Cross in Cambodia. Nordström also worked with the International Committee of the
Red Cross in Iran and then for the Swedish International Development Co-operation Agency for 12 years. He was the first person to take the helm of the newly created Global Fund to Fight AIDS, Tuberculosis and Malaria in 2002, during which time he helped establish its basic structure and organisation.

Namita Pradhan
Assistant Director-General—General Management

The job of Assistant Director-General for General Management was the role Nordström left to take up. Acting Director-General reins. Namita Pradhan comes to the role after involvement with WHO’s operations and strategic planning as Director of Planning, Resource Coordination and Performance Monitoring. She has worked on introducing results-based management to the organisation and was involved in the development of the 11th General Programme of Work 2006–15 and budget programming, including strategies to improve resource mobilisation for WHO’s activities. Pradhan started her career in the Indian Administrative Service in the late 1970s, including a period as Additional Commissioner for Family Welfare, in which she oversaw a World Bank-funded programme of health care and family planning in Mumbai. Other government appointments included positions in the ministries of women and child development and as Director of International Health in the Ministry of Health. After joining WHO’s Indonesia country office in 1996, Pradhan worked on strengthening health systems before moving to the Geneva headquarters as an adviser first to Gro Harlem Brundtland and then to Lee Jong-wook.

Susanne Weber-Mosdorf
Assistant Director-General—Sustainable Development and Healthy Environments

Susanne Weber-Mosdorf remains the Director-General’s Representative on European Union affairs in addition to the sustainable development and environmental role she has held since last year. A former German civil servant, Weber-Mosdorf studied economics, legal sciences, and politics, as well as administrative and management sciences at the University of Konstanz, Germany, before working at the Council of Europe in Strasbourg and studying further at the Ecole Nationale d’Administration in Paris. She has held several positions in government and industry across many sectors and disciplines in Germany including 4 years as Mayor of Kirchheim. Directly before joining WHO, she was Director-General of the Federal Ministry of Health and Social Security in Berlin, which included responsibility for international issues including health and social policy. During this time, Weber-Mosdorf was also the head of the German delegation to the World Health Assembly and the WHO Regional Committee for Europe.

Howard Zucker
Assistant Director-General—Health Technology and Pharmaceuticals

Howard Zucker’s unchanged portfolio includes anticounterfeiting efforts, intellectual property and health, access to medicines, applying technology to solve public-health problems, and traditional medicines. In addition to his Assistant Director-General role he is also Representative for the Director-General on Intellectual Property, Innovation, and Public Health. Zucker came to WHO from the US government where he was Deputy Assistant Secretary of Health at the US Department of Health and Human Services but before moving up the administrative ranks he gained a first degree and MD from McGill University and George Washington University School of Medicine, respectively. Zucker’s clinical training was in paediatric anaesthesiology; he held the post of Assistant Professor at Yale University School of Medicine, and Associate Professor at Columbia University College of Physicians and Surgeons, Adjunct Associate Professor at Cornell University Medical School and was also on the faculty at the National Institutes of Health. Zucker is a member of the Bar of the Supreme Court of the USA, a member of the Council on Foreign Relations, and was a public health “high-level expert” for NATO.

More continuity than change

Margaret Chan says she thought long and hard about who to appoint in her senior team. The authority granted to her by the World Health Assembly means she could have brought in an entirely new set of subordinates, if she had so wished. But, says Chan, after consulting widely on the campaign trail, “I did not hear a single member state telling me that they want major change”. The result is that many of Lee Jong-wook’s close advisers and colleagues keep their roles, with only minor changes of responsibility. “The team of senior management brought in by Dr Lee has established networks and relationships...So now I would like to maintain the stability and continuity,” she says.

In addition to the Assistant Director-General reappointments, Chan has also publicly confirmed the roles of four other high-level staff: Denis Aitken, formerly Lee Jong-wook’s Chef de Cabinet, is now Representative of the Director-General for a new Partnerships and UN Reform programme; Bill Kean will continue as Executive Director of the Director-General’s Office; Liu Peilong will continue as an Adviser to the Director-General; Andrey Pirogov will continue as Assistant Director-General of the WHO Office at the UN in New York; and finally, Ian Smith, one of Lee’s close advisers, will continue as an Adviser to the Director-General.

Hannah Brown

See Perspectives page 893 for a profile on Margaret Chan
US guidelines seek to protect access to licensed technology

When universities grant exclusive licences to commercial companies for the discoveries of their scientists, they may unwittingly hamper university research. Now new licensing guidelines aim to address the problem and protect the public interest. Michael McCarthy reports.

When universities license their researchers’ inventions, they should negotiate agreements that allow for the widest dissemination of the new technology, according to a set of guidelines released on March 6 by 11 top US universities in conjunction with the Association of American Medical Colleges (AAMC). Universities should “be mindful of their primary mission to use patents to promote technology development for the benefit of society”, the guidelines say.

University scientists have sometimes found that they cannot use key research techniques and materials because universities had granted exclusive licences to commercial companies that either denied access to the technology or have demanded prohibitively high fees for its use. In some cases, licence agreements have prevented researchers from publishing their results or collaborating with other scientists. Increasingly, such licences are covering technologies, particularly those used in genomics and proteomics, that are early in the development process and are often essential for basic research.

In the past, some universities have signed technology transfer agreements that have “unwittingly given away the store”, said David Korn, AAMC senior vice president who helped craft the new guidelines, which came out of a meeting of research officers and technology transfer officers held at Stanford University in Palo Alto, California last year.

Arthur Bienenstock, former dean of research at Stanford, who convened the meeting, said that technology transfer officers are usually not researchers and “may not think about the long-term implications of licensing agreements they are signing”.

The guidelines, entitled In the Public Interest: Nine Points to Consider in Licensing University Technology, encourage universities to make technology transfer licences as non-exclusive as possible and provide legal language to help them draft such documents.

The guidelines recognise that in order for a technology to be developed and brought to market exclusive licensing agreements are often necessary, but that when the university has leverage it should negotiate agreements with “retained rights” clauses that reserve the right for scientists at universities, non-profits, and governmental organisations to use both “tangible research materials”, such as biological materials and chemical compounds, and “intangible materials”, such as software, databases, and know-how.

“Whenever possible, whenever humanly possible, the university that has to write an exclusive licence in order to develop a property should retain the right to use the discovery for its own researchers and to share it with other university researchers”, Korn said.

Because licensees sometimes decide not to develop a technology for business reasons, the guidelines recommend that the licences be structured to guarantee that discoveries are developed and put to use. Licences can also be drafted with “mandatory sublicensing” provisions so that should unmet market or public-health needs arise, the technology can be licensed to third parties willing to address those needs, according to the guidelines.

The guidelines also caution universities about entering into licensing agreements that guarantee the licensee access to “future improvements” and “follow-on inventions” related to the original discovery. Such obligations “may effectively enslave a faculty member’s research programme to the company”, the guidelines warn. “In the rare case where a licensee is granted rights to improvement patents, it is critical to limit the scope of the grant so that it does not impact uninvolved researchers and does not extend indefinitely into the future.”

The guidelines also call on universities to include provisions in licences that will allow their researchers’ inventions to be applied to address unmet needs of neglected patient populations and regions, such as therapeutics, diagnostics, and agricultural technologies for the developing world.

Universities should put the public interest first when they are granting licences for their researchers’ discoveries, said Korn, “it’s important to keep in mind why the university exists in the first place”.

Michael McCarthy
The history of childhood has been transformed from the time when French historian Philippe Ariès published his landmark book L’enfant et la vie familiale sous l’Ancien Régime in 1960. Translated into English as Centuries of Childhood in 1962 this book soon became the gospel view of children in the past. Parental indifference to the premature deaths of young children was one of Ariès’ key concerns. If your children are destined to die before their fifth birthday, why become too intimately attached to them? Nowadays not a week goes by, it seems, without national commissions and government task forces issuing admonishing reports about our own children’s welfare. We claim that every child counts yet enact little to substantiate the rhetoric. In the richest developed countries, children of the poor continue to be neglected while those of the rich have plenty to eat and are smothered in consumables. Yet mortality rates of both rich and poor children in the West are at their lowest in history. However, in most developing countries the overall neglect of children is so grim that death rates have not been reduced.

Here enter historical demographer Robert Woods with his “experiment” to chart “three parallel histories: demographic rates, artistic images, and poetic language” relating to childhood death. The idea is to study degrees of emotional indifference by focusing on one emotional sphere—grief—viewed through these three distinct disciplines. Woods stands on firm ground when surveying child mortality in the major Western countries (Eastern Europe and the rest of the world are omitted), buttressed by up-to-the-minute scholarship, fleet statistics, and the accepted view about a steep drop in mortality rates at the end of the 18th century. He is adroit in decoding intricate patterns of child mortality, but out of his depth when analysing visual and literary materials. The 16 paintings he enlists seem unrepresentative of broad attitudes to grief and he says little new about them.

“The range of emotions attached to the premature death of children in history has been so varied that it cannot be delimited to words, still less to a single genre. Formerly “we moderns continue to grieve the premature loss of our own children only and turn a deaf ear to those of less fortunate families in the developing world”.

Historians and anthropologists of these matters, including Ariès, have immersed themselves in far-flung diaries, journals, letters, and memoirs to understand this most modulated, if also elusive, expression of loss. But Woods glues himself to poetry for reasons he never explains, a privileged and coded discourse controlled by literacy, education, class, and social exclusion. Perhaps some authority advised him to read poems instead of more mundane forms of writing, or his demographic bias propelled him to think he landed on “statistical sources” by the identification of 69 poems written in English over 400 years because they contain “keywords”, from which he produces his “vocabulary of grief”. Missing are the all-important contexts capable of endowing this welter of statistical detail with meaning. Woods’ discussion omits their relevant traditions, conventions, conceits, metaphors, and tone, and the crucial biographical personalities of the authors. It is curious to delimit oneself so idiosyncratically or suppose that just a few keywords could unlock the secrets of parental grieving.

His approach instils little confidence in interdisciplinary methods of historical and cultural analysis, where the interloper persuades the other discipline by speaking their language first and then converting it. A more fitting interdisciplinary approach would have dwelled on the comparative versions of grief in different societies: Western and Eastern, civilised and primitive. Yet cultural relativism and religious variations outside Christianity seem not to exist within Woods’ intellectual orbit.

Woods’s approach neither changes the understanding of parents and children in history nor makes a difference to the contemporary malaise about the crushing need to reduce child mortality in developing countries. A more productive approach would perhaps have been to select two or three historical moments in the 18th and 19th centuries when child mortality rates sharply declined, and studied the responses to precipitous decline through “numbers, pictures and words”. This may have begun to explain why we moderns continue to grieve the premature loss of our own children only and turn a deaf ear.
The printed journal includes an image merely for illustration

to those of less fortunate families in the developing world. Alternatively, Woods could have devised other statistical models to measure forms of greed at key moments, and explore whether greed patterns changed among those whose children were now suddenly surviving. Greed, rather than any other “emotion”, prevents us in 2007 from assisting the dying children of developing nations. Greed is, indeed, the new “indifference” and may become the most interesting emotion of the 21st century.

Nor does Woods shed light on the incommensurability of the forms of contemporary global grief over the death of children. African children are smothered with love by their mothers according to their local African cultural systems. Poverty and disease rather than parental apathy kills these children, and when they die their parents grieve profoundly, albeit differently from their more affluent Western counterparts. By contrast, many Western children are swaddled in sentiment and bribed by materialism, while our experts continue to confirm that their childhoods are wrecked, their early lives ruined by greed and runaway consumption. Despite this state of affairs the survival rates of Western children are higher today than at any time in recorded history. When every survey of contemporary child mortality assures us that it continues to be unacceptably high in the 60 countries identified in Millennium Development Goal 4, it is a shame that such a capable historical demographer as Woods has scattered his energy so disappointingly.

George Rousseau
geroge.rousseau@ntlworld.com

In brief

Book Dying of drugs

Although the history of human drug use reaches back millennia, the notion of addiction itself—ie, the characterisation of heavy drug use as a disease—did not occur until about 200 years ago: first for alcohol and later for the far less prevalent opiates. But even once so identified and ensconced in the official catalogues of disease, addiction is still disputed by many in favour of more social or psychological constructions. Not surprisingly, comprehensive studies of the global epidemiology of drug use remain a novelty. No coherent summary of these most fundamental data—incidence, prevalence, morbidity, and mortality rates—have previously been available about illicit drugs, and even now measuring this illicit behaviour is no simple task.

Accordingly, Mortality Amongst Illicit Drug Users is a welcome sight. Shane Darke, Louisa Degenhardt, and Richard Mattick carefully synthesise the mass of literature on drug-related mortality, delineate its key domains, and provide a rich bibliography for future investigators. They estimate 200 000 annual deaths attributable to illegal drug use among the more than 200 million users worldwide. These deaths occur disproportionately among young men and the 8 million users who inject opiates. The major causes of death are: overdose (25%); disease, especially HIV and hepatitis C virus infection (25%); suicide (10%), which is ten times the rate of comparable non-users; and the sequelae of violence and trauma common to illicit drug scenes. Each topic is independently examined and usefully linked to data on a range of interventions that can reduce mortality, including effective treatment, harm reduction strategies that reduce risk among active users, and peer access to naloxone for overdoses. Although “supply reduction” can save lives, so can “supply enhancement”; programmes that offer heroin in Switzerland, The Netherlands, and Canada are also saving lives.

In concluding, the authors ask “Why should society care about drug-related deaths?” Why indeed? A century of moral panics and social crusades that demonise drug users hasn’t helped. Many health professionals still see drug problems as self-inflicted, allowing the usual clinical, ethical, and moral responsibilities to be ignored. And organised political efforts (including international treaties) to legislate drugs out of existence by criminalising drug users are a big part of the problem: imprisonment dramatically increases the risk of overdose in the weeks after release. Now that the HIV/AIDS pandemic has created “a perfect storm” of lethal consequences of drug use, it is vividly apparent how far the public-health approach to drugs is from realised. This book’s documentation of drug-related mortality—so much of it preventable with existing methods readily available to us—is a testament to the poverty of current drug policies and bears stark witness to the price paid for these failures.

Ernest Drucker
emdrucker@earthlink.net
Profile
Margaret Chan: now is the time for WHO to achieve results

These days, the job of WHO Director-General is as much about redefining the notion of health as it is about managing an international organisation. Gro Harlem Brundtland, who held office from 1998 to 2003, led this trend by framing health as an economic issue with her acclaimed Commission on Macroeconomics and Health. Her successor, Lee Jong-wook, promoted the view of health as a product of wide-ranging social determinants during his 3-year term. But Margaret Chan, who took office on Jan 4 this year, looks set to push through what could be the organisation’s most important conceptual change. For Chan’s term, health is all about security. Chan described her position to The Lancet in an interview at her Geneva offices. “We are putting health on the security agenda because of the vulnerability due to new and emerging infections”, she explains. “The notion of universal vulnerability is extremely important. If you do not have a system in place to protect what I call collective defence, no matter how strong or how well prepared the UK or US is, the weakest point is in a country very far away.”

There is more to this idea than urging member states to take sensible precautions and share responsibility. Chan is acutely aware that she has taken office at a time when health is the focus of much unprecedented political attention. “It is a matter for us to be smart to maintain that political commitment and deploy resources wisely to achieve results”, she says. Chan believes that fulfilling her responsibility requires asking some tough questions about WHO’s activities and culling those that have outlived their usefulness. In her business-like way, Chan says WHO should refocus on its “core competencies”—for example, implementation of the international health regulations—and avoid “internal inefficiencies” or duplication of activities that other organisations are already doing. “There are so many other key players [in global health], so we have to focus in areas that if WHO does not take on nobody else will”, she says.

Now is a good time for this kind of soul-searching. WHO’s 60th birthday is coming up next year, but this year’s World Health Assembly will again be discussing many of the same issues that concerned the first assembly in 1948. So how does Chan plan to move the organisation forward? “In some areas we are still quite traditional and conventional. In order to be relevant, we really need to rethink”, she says. And this includes dumping some of the oft-repeated excuses for WHO’s poor performance. “People always say we don’t have enough money, we don’t have enough this, that, and the other”, says Chan. “Yes, resources are important, but we also have to ask ourselves should we be doing so many things”, she says.

Pragmatism is key to Chan’s leadership style, as shown by her willingness to question WHO’s traditional role. She is, for example, prepared to rethink the part played by private providers in the delivery of health care. In her straight-talking manner, she tells The Lancet that her main aim before taking office was to identify the key things that member states want from WHO. During her hectic campaign schedule, which took in visits to 30 countries and meetings with 40–50 health ministers, Chan put together her vision of currently unmet needs. Tackling the pervasive problems of conflict emerged as a key priority, and Chan has responded by elevating the Health Action in Crisis programme to cluster level. “At any one time 40 countries in WHO—and we have 193 member states—are in situations where they are in crisis, either in conflict or recovering from conflict. That is a big number, so I need to realign the priorities with the needs of our member states”, Chan says. The other major management change so far is to split up the responsibilities of the former Evidence and Information for Policy cluster into two new divisions: a cluster on Health Systems and Services, led by former Acting Director General Anders Nordstrom; and one dedicated to Information, Evidence, and Research, led by Tim Evans. This change reflects member states’ concerns about the limited effectiveness of vertical programmes in countries with weak health systems and demands for evidence on which to base policy decisions.

But while promising to redefine WHO’s relations with the outside world, Chan has been careful to keep the upheaval inside WHO to a minimum. “Dr Lee’s sudden death took many people by surprise and that was a huge change for the organisation and that was a big shock for our member states”, she says solemnly. And, in the absence of a demand from member states for a big change, Chan has decided to keep the “team of excellent people” Lee brought in just 3 years ago—of whom Chan herself was a part—and build on, rather than waste, the time and effort they have invested in making contacts and building relationships. “I know them, I worked with them, and I respect them”, she says.

For now, Chan’s main focus is on achieving results and she is aware that this will mean changing the way WHO does business. “Why did I use the health of African people and the health of women [as indicators of WHO’s success]? Because I know that the data that we have been collecting is not very good”, she says. “The bottom line is we should not squander the political will and the resources on hand. We really need to produce results.”

Hannah Brown
hannah@two-cultures.com
James Allen Dvorak

Pioneering malaria researcher. He was born on Oct 26, 1935, in Racine, WI, USA, and died on Feb 5, 2007, of a brain haemorrhage in Bethesda, MD, USA, aged 71 years.

More than a half century ago James Dvorak threatened to throw his only microscope into the moat of Japan’s Imperial Palace if the Japanese woman he loved refused his marriage proposal. By agreeing she saved a tool that came to inspire landmark studies on malaria and Chagas’ disease. Dvorak, a senior scientist at the US National Institute of Allergy and Infectious Diseases (NIAID), in Bethesda, MD, USA, was known best for his groundbreaking microscopic innovations that allowed him to videotape malaria parasites as they invaded red blood cells. His work, which gave new insight into one of the world’s most deadly microbes, made the cover of Science magazine in 1975.

“If a picture is worth a thousand words, a video is worth even more”, said Thomas Wellem, chief of NIAID’s Laboratory of Malaria and Vector Research, where Dvorak ran the Biochemical and Biophysical Parasitology Section. Dvorak captured the parasites with unprecedented clarity in a live cell chamber he designed. His techniques came before scientists had figured out how to keep malaria parasites alive in vitro. “To see one of these beasts bursting out of a red blood cell and then track it as it interrogates the surface of a fresh red blood cell—watching it decide whether that will be the next one it wants to invade—well, it was just remarkable”, said Wellem. The footage quickly became essential viewing for anyone in the field.

From that work Dvorak and Louis Miller, now chief of NIAID’s Malaria Vaccine Development Branch, used Dvorak’s live cell chamber to identify an important protein known as the Duffy antigen. They showed that this antigen serves as a gateway through which Plasmodium vivax gains entry into red blood cells. “This period, working daily and nightly with Jim, was a wonderful experience for both of us”, recalls Miller. Their discovery led to work that showed that nearly all Africans lack the Duffy antigen, explaining why black people tend to be immune to P vivax—an evolutionary phenomenon that does not extend to P falciparum.

Dvorak was working on several projects when he died. The most important involved the use of atomic force microscopy to view surface changes on red blood cells infected with malaria. In this he was considered a maverick because he was so visual. He made topographical maps of infected blood cells, creating clear images of the sticky nodes malaria parasites produce on the cells’ outside membranes and examined how these nodes prevent infected cells from reaching the spleen.

Alongside his work on malaria, Dvorak was a renowned expert on Chagas’ disease. In Brazil he tracked the parasites’ many strains, leading to a realisation that the disease had many more manifestations than originally thought. “He showed that a lot of people are infected with a broad sweep of symptoms, from nothing to heart attacks”, said Alan Sher, chief of NIAID’s Laboratory of Parasitic Diseases. “He was probably a genius, and so dedicated to his work.”

As a young man Dvorak served for 3 years in the US Army. After graduating from the University of Tennessee in 1963, he got his doctorate in molecular biology and biophysics from the University of Pennsylvania in 1968. His first job was at the National Institutes of Health, where he would remain for the rest of his life. A motorcycle enthusiast rarely seen without his signature black biker boots, he is remembered as warm but demanding and always generous towards young scientists. “He considered our future with his heart”, said Fuyuki Tokumasu, who came from Japan to work for Dvorak.

Phred Dvorak said her father’s passions for work, friends, and family were seamless. During a high school family sports day, while other father-daughter teams competed, she and her father walked the school grounds talking about parasites. He was, she said, a man completely devoted to the things he loved. Nothing exemplified this more than Dvorak’s initial reaction to his wife of nearly 50 years. They met in Tokyo while he was in the Army. On their first date he asked her to marry him. She refused many times before he threatened to toss his microscope into the palace moat. Dvorak is survived by his wife, Hisako Komori Dvorak; his daughter, Phred; his two sons, Vojin and Anton; a sister; and two grandchildren.

Karen Masterson

masterson_wang@yahoo.com
The Lancet in place of the data published in should be accepted as accurate assumptions and approximations, on a series of multiple, erroneous that their own calculations, based by Gøtzsche. Finally, they claim and in another previously cited in a publication which they cite withdrawal, but it is described its withdrawal—flaws that had apparently initially escaped the reviewers. When attempting to estimate the number of cancers which should have been diagnosed in the Swedish Two-County mammography screening trial, Gøtzsche and colleagues misquoted the end date of the 11-year trial by nearly 4 years, although the correct dates appear in one of their cited references. An error of this magnitude would render their calculations worthless.

More seriously, they failed to account for the precise, stepwise, district-by-district nature of the randomisation process published for the trial. Getzschke and colleagues refer to this as “new information” received after their manuscript’s withdrawal, but it is described in a publication which they cite and in another previously cited by Gotzschke. Finally, they claim that their own calculations, based on a series of multiple, erroneous assumptions and approximations, should be accepted as accurate in place of the data published in The Lancet for the trial.

Pointing an accusing finger at peer reviewed published medical research gains attention, but this particularly clumsy attempt was uncovered at a late stage of prepublication. It is to the credit of the EJC that the manuscript was withdrawn.

I declare that I have no conflict of interest.

Peter B Dean
debian@tyks.fi
Faculty of Medicine, University of Turku, PB 52, 20521 Turku, Finland

In its meeting of December, 2006, the Committee on Publication Ethics (COPE) discussed The Lancet’s decision to publish Peter Gøtzsche’s account of his experience with the European Journal of Cancer. COPE concluded that The Lancet was correct to bring this case into the public domain and noted, like the International Committee of Medical Editors, that published work—electronically or otherwise—should not be removed without appropriate correction or retraction. Retraction or removal is a very serious matter for authors and their institutions and should not happen without due process, which COPE judged was lacking in this case.

Sabine Kleinert
sabine.kleinert@lancet.com
Vice-Chair, Committee on Publication Ethics, c/o The Lancet, London NW1 7BY, UK

My first reaction on reading the Comment by Peter Gøtzsche and colleagues about their experience at the hands of the European Journal of Cancer was one of rage. What struck me is not that there might be a debate about the work of Gøtzsche and colleagues and their challenge to the reliability of the Two-County Trial, but that it was conducted in such a one-sided and occult fashion. Violent letters pressuring the Editor-in-Chief of the European Journal of Cancer have not been published, anonymous referees’ reports have been fed in a selective and fragmented fashion to the authors, and the paper was removed from the journal’s website without warning.

There are two issues that need airing here. First, is there any ethical or scientific basis to anonymous refereeing? Those who think that ethics are not an issue should imagine the analogy of unsigned prescriptions. Scientifically, the answer is also no. There is no evidence that anonymising referees’ reports makes any difference to the quality of the content or indeed of what is published. So, why do it?

Second, how are we doing in our contemporary research scene? Not well I would say. The Comment that follows Gøtzsche and colleagues’ account of research censorship and debate stifling is about straightforward fabrication (the Sudbo case). Until such time as the whole issue of perverse incentives is taken seriously by the scientific community, please let us not be indignant about plagiarism, redundant publications, salami slicing, citation cartels, data fabrication, and other similar practices. We have reaped what we have sown. Let us instead try to open up the peer review process. This might help in making ethical decisions, where the influence of lobbies, activism, ideology, and lucre are at least recognised.

I personally know P C Gøtzsche, and respect his work and his integrity. This might have biased my judgment.

Tom Jefferson
tom.jefferson.tom@gmail.com
Cochrane Acute Respiratory Infections Group, Via Adige 28a, 00061 Anguillara Sabazia, Italy

Correspondence

Submissions should be made via our electronic submission system at http://ees.elsevier.com/thelancet/
Wilson’s disease

Aftab Ala and colleagues (Feb 3, p 397) report that studies on Wilson’s disease patients with the homozygous H1069Q mutation have been used to investigate possible genotype–phenotype relations, but that an association has not been universally reported.

Although other genetic or environmental factors could influence the phenotypes of Wilson’s disease, a possible genotype-phenotype correlation has been suggested by clinical studies reporting other less common Wilson’s disease gene mutations. On this point, family studies could be of great interest since members of the same family should have the same additional genetic and environmental factors. Our laboratory did a family study showing a new missense homozygous mutation (T1288R) of the ATP7B gene in three brothers with Wilson’s disease; all three presented with hepatic symptoms (H2 Ferenczi score phenotype).

Another family study by Firneisz et al. showed a different phenotype in members of the same family with different mutations. The mother (H1069Q/E1064K) presented with late-onset neurological symptoms, whereas her son (E1064K/E1064K) presented with early-onset, rapidly progressing acute liver failure requiring orthotopic liver transplantation.

Although further confirmation is needed, different mutations of ATP7B could be related to specific clinical presentations. This feature could be of great interest in Wilson’s disease in terms of clarification of the pathological mechanisms and clinical features.

We declare that we have no conflict of interest.

*Lorenzo Leggio, Giovanni Gasbarrini, Giovanni Addolorato
g.addolorato@rm.unicatt.it
Institute of Internal Medicine, Catholic University of Rome, Largo A Gemelli 8, 00168 Rome, Italy

In their admirable Seminar on Wilson’s disease, Aftab Ala and colleagues give a valuable and up-to-date summary of our latest knowledge on the genetics of this disorder, but the section on clinical aspects and treatment calls for comment.

In addition to the presenting symptoms they describe, Wilson’s disease can also present with osteoarticular problems and renal disease, although these are rare. Sunflower cataracts can be visible to the naked eye and do not necessarily require slit-lamp examination to reveal their presence. The Kayser-Fleischer ring always appears first as a crescent at 12 o’clock on the cornea, next being manifest at 6 o’clock, before joining up to form a complete ring. Pigment gallstones should also be looked for: they are probably a legacy of earlier haemolysis and are often symptomless.

I am less than happy about two of Ala and colleagues’ therapeutic recommendations. First, penicillamine should always be given before meals, not after. I showed in 1967 that when penicillamine is given before the administration of a test dose of radiocopper (64-Cu) it mobilises much more of the metal than when given afterwards. Presumably the penicillamine, present in the plasma, can chelate the copper before it becomes protein-bound. If the drug (and this will also apply to trientine) is given after the copper has been absorbed and handled by the liver, it is much less vulnerable to chelation. Second, I have never been able to follow the logic of treating a patient with a metal (zinc) and a chelating agent—one will combine with the other so neither can be fully effective. If this strategy is followed, it must be made very clear to the patient that the different drugs must be taken many hours apart. The combination of zinc and trientine can induce a serious sideroblastic anaemia and the result is dramatic worsening of neurological signs.

I declare that I have no conflict of interest.

J M Walsh
penicillamine@waitrose.com
Broom Lodge, SS High Street, Hemingford Grey, Huntingdon PE28 9BN, UK

In their Seminar on Wilson’s disease, Aftab Ala and colleagues omitted a few clinical observations.

We have found that Kayser-Fleischer rings are more common in patients who present with neurological symptoms as the initial manifestation of Wilson’s disease (23 of 24) than in those with hepatic symptoms at presentation (30 of 36). It is therefore essential that a thorough examination of the eye is done in any patient who presents with neurological and neuropsychiatric symptoms.


*Correspondence

We have also shown that those presenting with liver symptoms had a fivefold increase in mortality compared with those who had neurological symptoms at presentation. Liver transplantation would therefore be more appropriate in this group.

We declare that we have no conflict of interest.

Samar Harris, *Harris V K Naina, Sameer Siddique
naina.harris@mayo.edu

College of Medicine, Mayo Clinic, Rochester, MN 55905, USA (SH, HVKN); and Department of Gastroenterology, Ipswich Hospital, Ipswich, UK (SS)


Subarachnoid haemorrhage

We thank Jan van Gijn and colleagues for an extensive Seminar on subarachnoid haemorrhage (Jan 27, p 306). In patients with aneurysmal and non-aneurysmal subarachnoid haemorrhage, various scores are commonly used for initial assessment and to predict outcomes. These scores are not mentioned by van Gijn and colleagues.

The Hunt and Hess scale is usually used in the assessment of aneurysmal subarachnoid haemorrhage, but can also be suitable for non-aneurysmal perimesencephalic haemorrhage. A higher grade correlates with a worse outcome. The World Federation of Neurological Surgeons (WFNS) scale is a widely used clinical scale for aneurysmal subarachnoid haemorrhage based on the Glasgow coma scale—a reliable and familiar method to assess consciousness. The WFNS scale is more popular than the Hunt and Hess scale because the latter measures variables that are not of prognostic significance and has interobserver variability. Prognosis is inversely proportional to the score, similarly to the Hunt and Hess scale.

The Fisher and Claassen scales reliably grade the amount of blood seen on initial CT scanning of the head. The Fisher scale identifies patients with thick cisternal blood, but does not differentiate between intraventricular and intracerebral haemorrhage. However, the Claassen scale assesses intraventricular and intracerebral haemorrhage separately and shows that both are independent predictors of poor outcome, as is a thick subarachnoid clot.

We declare that we have no conflict of interest.

Yatindra Kumar Batra,* Subramanyam Rajeev
docrajeevs@gmail.com

Department of Anaesthesiology and Intensive Care, Postgraduate Institute of Medical Education and Research, Chandigarh 160012, India


In their comprehensive Seminar on subarachnoid haemorrhage, Jan van Gijn and colleagues do not address the relevance of the disorder as a clinically significant cause of hypothalamic-pituitary dysfunction.

The description of a high incidence of hypothalamic lesions in the wake of aneurysmal subarachnoid haemorrhage dates back more than 40 years. Correspondingly, several recent studies have shown that patients are at substantial risk of hypothalamic-pituitary dysfunction after subarachnoid haemorrhage. Five studies including a total of 122 patients who underwent endocrine testing several months to years after aneurysmal subarachnoid haemorrhage have identified hypothalamic-pituitary deficits with a pooled frequency of 47%. Thus, this complication could be as common as other long-term sequelae of subarachnoid haemorrhage such as psychosocial dysfunction, anosmia, and neuropsychological deficits.

Endocrine dysfunction could conceivably contribute to medical as well as neurobehavioural sequelae of subarachnoid haemorrhage. Indeed, in our own observation, neuroendocrine dysregulation is associated with lower quality of life scores and more depressive symptoms in patients investigated more than 1 year after the acute event.

Many patients who have had a subarachnoid haemorrhage experience neurological and psychosocial sequelae that might mask the sometimes subtle signs of hypopituitarism. Current clinical experience—as well as the missing mention of this topic in van Gijn and colleagues’ Seminar—indicates that most cases of hypopituitarism after subarachnoid haemorrhage remain undiagnosed. Physicians who treat patients with subarachnoid haemorrhage should be aware of this condition since it represents one recognisable and treatable potential cause of impaired recovery and long-term morbidity in survivors of subarachnoid haemorrhage.

IK-A has received speaker fees, travel grants, and research grants from Pfizer and Novo Nordisk, and travel grants from Ipsen. She is a member of the German KIMS advisory board. HJS has received speaker fees and travel grants from Pfizer, and travel grants from Lilly, Novo Nordisk, and Serono.

*Ilonka Kreitschmann-Andermahr, Harald Jörn Schneider
ilonka.kreitschmann@post.rwth-aachen.de

Department of Neurosurgery, RWTH Aachen University, Pauwelsstrasse 30, 52074 Aachen, Germany (IK-A); and Max Planck Institute for Psychiatry, Internal Medicine/Endocrinology and Clinical Chemistry, Munich, Germany (HJS)


www.thelancet.com Vol 369 March 17, 2007 903
We were concerned to read in the Seminar by Jan van Gijn and colleagues on subarachnoid haemorrhage that “experienced neurologists and neurosurgeons can confidently exclude xanthochromia by visual inspection alone”.

In the reference quoted in the Seminar to support this statement, an arbitrary cut-off of 0·05 absorbance units (AU) at 450–460 nm was used despite no evidence to back up this level.

We have highlighted two cases in which spectrophotometry of clear and colourless CSF showed minor increases in CSF bilirubin, which then led to catheter angiography and increases in CSF bilirubin, which then and colourless CSF showed minor evidence to back up this level.

The ability to assess xanthochromia in CSF is highly subjective, particularly when other pigments such as oxyhaemoglobin are also present. We feel that spectrophotometry provides the best available method for investigation of CSF in CT-negative suspected subarachnoid haemorrhage.

Although headache is undoubtedly one of the most characteristic manifestations of subarachnoid haemorrhage, the diagnosis, where appropriate, should still be entertained even in the absence of this symptom.

For example, if the initial manifestation of subarachnoid haemorrhage is sudden loss of consciousness, and on recovery of consciousness the patient subsequently becomes mentally confused, the fact that he denies headache should not necessarily rule out the diagnosis. In one series, comprising 1962 patients with recently ruptured aneurysm, 24% reported no headache at the time of admission. In another report involving a presumably mentally alert patient who had not lost consciousness, the initial symptom was vomiting, and this preceded the gradual onset of headache by at least 24 h.

We declare that we have no conflict of interest.

Ian Holbrook, Robert Beetham, Anne Cruickshank, Geoff Keir, Ian Watson
ian.b.holbrook@york.nhs.uk

Department of Clinical Biochemistry, York Hospital, York, YO31 8HE, UK (IH); Department of Clinical Biochemistry, Frenchay Hospital, Bristol, UK (RB); Biochemistry Department, Southern General Hospital, Glasgow, UK (AK); Department of Neuroimmunology, National Hospital for Neurology and Neurosurgery, London, UK (GK); and Department of Clinical Biochemistry, University Hospital Aintree, Liverpool, UK (IW)

Stroke prevention: missed opportunities

We welcome Nils Wahlgren and colleagues’ exciting findings on thrombolysis for acute ischaemic stroke (Jan 27, p 275). However, although thrombolysis can reduce the effect of major disability in the small minority of stroke patients admitted within 3 h of their symptoms, effective and vigorous primary and secondary preventive measures will have the most impact on reducing the overall burden of stroke.

As part of our stroke audit cycle, we studied patients with recurrent ischaemic stroke to assess the prevalence of undertreated risk factors and also the incidence of new vascular risk factors after their admission. 81 (24%) of 331 stroke inpatients in 2005 had a recurrent ischaemic stroke. Only 17 of 52 patients with pre-existing hypertension had achieved the target blood pressure set in the guidelines for secondary prevention of stroke; 10 of 37 with cholesterol concentrations of more than 3·5 mmol/L after their initial stroke were not on a statin; 13 of 81 patients were still smoking; 47 of 81 had not had carotid dopplers after their previous stroke; 12 of 81 had newly diagnosed hypertension; and a further 12 of 81 had new hypercholesterolaemia.

Most patients with recurrent ischaemic stroke had multiple risk factors identified at the time of their first stroke, but most of these risk factors were either undertreated or untreated. A smaller number of patients had newly identified vascular risk factors at admission. We agree with the essence of your Editorial that interventions to reduce hypertension, poor diet, and tobacco use would save more lives globally by preventing strokes than the use of thrombolytics, antiplatelet drugs, and neuroprotective agents.

We declare that we have no conflict of interest.

Peace through medical education in Bosnia and Herzegovina

12 years after the Dayton Peace Agreement ended the 1992–95 war in Bosnia and Herzegovina, the country is still burdened with two major problems: ethnic and political division, and economic transition. Medical teachers found a way to tackle both at the same time.

A SWOT (strengths, opportunities, weaknesses, threats) analysis showed that, despite their political, ethnic, and religious differences, all five schools of medicine in Bosnia and Herzegovina (Sarajevo, Banja Luka, Tuzla, East Sarajevo, and Mostar) had much in common: they were all under pressure to work together. Serbs, Muslims, and Croats regularly gathered at informal meetings or dinners, eating and drinking at the same tables, laughing and joking, even singing traditional songs—often to the amusement of their colleagues from previously divided by the war, had a chance to communicate with each other, dispelling some prejudices and regaining belief that it is possible to work together.

The example of the schools of medicine of Bosnia and Herzegovina shows that higher education can be a favourable arena for peace promotion. Financial incentive can serve as a catalyst for establishing and maintaining trust and good-will. The conclusion is that peace can be promoted indirectly, through formal education and professional engagement, not necessarily by pressing the “opposing” sides to talk about reconciliation and sign peace declarations.

We declare that we have no conflict of interest.

*Sunku Hemant H Guptha, Pazzhoor Shibu, Peter Owusu-Agyei sunku.guptha@pbh-tr.nhs.uk

Edith Cavell Hospital, Breton Gate PE3 9GZ, UK


Sexual dysfunction in HIV infection

We welcome Rosemary Basson and The Lancet’s proposal that sexuality in chronic illness should no longer be ignored (Feb 3, p 350). Particularly, we would like to focus on sexual dysfunction among people with HIV infection. In this context several factors could contribute to sexual dysfunction: psychological and emotional issues, endocrine alterations, peripheral and autonomic neuropathy, comorbidity such as cardiovascular diseases or sexually transmitted diseases, and side-effects of antiretroviral therapy or other concomitant medications.

Little scientific attention has been dedicated to this topic and clinical investigation is often intuitively based, clinicians commonly being reluctant to assess sex-related issues with their patients for many reasons including fear of awkward situations, lack of training about sex-related counselling, and time constraints. Furthermore, HIV-infected patients are not always inclined to disclose problems of sexual activity because of the specific relevance of sex-related issues to this infection.

Undisclosed sexual dysfunction among HIV-infected people can have specific consequences. First, patient-reported sexual dysfunction is associated with non-adherence to antiretroviral therapy. Since non-adherent patients are more likely to have higher HIV RNA concentrations in semen or cervical secretions (with the risk of harbouring drug-resistant virus) and to engage in unprotected sex, they could cause
more frequent transmission of drug-resistant HIV strains. Second, men with alterations in sexual activity might take treatment for erectile dysfunction outside of medical prescription, raising the likelihood of pharmacokinetic interaction with antiretrovirals.

In the light of clinicians’ need to optimise treatment adherence and to ensure an adequate quality of life for HIV-infected people, identification and treatment of sexual dysfunction together with close risk-reduction counselling should be encouraged, not only for the sake of the individual but also for the sake of public health.

We declare that we have no conflict of interest.

*Maria Paola Trotta, Adriana Ammassari, Rita Murri, Antonella d’Arminio Monforte, Andrea Antinori*trotta@inmi.it

National Institute for Infectious Diseases “L. Spallanzani” IRCCS, Via Portuense 292, 00161 Rome, Italy (MPT, AA, AA); Clinic of Infectious Diseases, Catholic University, Rome, Italy (RM); and Clinic of Infectious and Tropical Diseases, S Paolo University Hospital, Milan, Italy (AdAM)


Peer review and the Term Breech Trial

We refer Lancet readers to continuing debate about the Term Breech Trial, published in this journal in 2000. The paper had a powerful impact worldwide, not least because the trial was stopped early (owing to a higher event rate than expected) and was fast-tracked by The Lancet. The intervention of planned caesarean benefits the term breech baby (perinatal death prevented in 1% of cases) but harms the mother (relative risk of immediate severe maternal morbidity 1·29, 95% CI 1·03–1·61). Thus, it is unsurprising that there should be a continuing desire to question, refine, and follow up the information from the trial, particularly when 2-year follow-up data show no residual maternal harm and no effect on neurodevelopmental outcomes.

However, there seems to be a recurrent theme of criticism of the trial followed by rebuttal, with the authors impugning critics of their trial as suffering from prior beliefs and disappointment in the findings. They also shelter behind the respectability of the peer-review process. But did this fast-tracked peer-review process provide an adequate shield? One peer reviewer (SB) recommended, in view of the importance and likely impact of the results, that it was not fast-tracked (unless and until detailed queries were addressed, particularly around the differential findings and implications for resource-rich and poor countries, where much of the controversy has raged). The Lancet, presumably on the basis of other favourable reviews, chose to go ahead with fast-track publication.

We have watched the subsequent scientific debate with concern. We are worried about the sanctity in which peer review is held and used to defend this research; investigators and editors usually do their best and in this case were supportive of their findings. Peer review is, however, subject to all the pitfalls of any judgment process. In retrospect, fast track in particular might only be appropriate with unanimous support from review. Further consideration of the points might have reduced the subsequent controversy.

Medical journals are becoming more transparent because this is thought to protect scientific integrity. Is it time to make peer review more transparent?

SB and AS belonged to the group of uncertain clinicians and were supporters of the trial. They offer external cephalic version followed by elective caesarean but would also support informed women to have vaginal breech delivery if they choose or arrive in advanced labour.

*Susan Bewley, Andrew Shennan susan.bewley@gss.tsthames.nhs.uk

Department of Gynaecology, Guy’s and St Thomas’ NHS Foundation Trust, 10th Floor North Wing, St Thomas’ Hospital, Lambeth Palace Road, London SE1 7EH, UK

Bivalirudin in patients with acute coronary syndromes undergoing percutaneous coronary intervention: a subgroup analysis from the Acute Catheterization and Urgent Intervention Triage strategy (ACUIty) trial

Gregg W Stone, Harvey D White, E Magnus Ohman, Michel E Bertrand, A Michael Lincoff, Brent T McLaurin, David A Cox, Stuart J Pocock, James H Ware, Frederick Feit, Antonio Colombo, Steven V Manoukian, Alexandra J Lansky, Roxana Mehran, Jeffrey W Moses, for the Acute Catheterization and Urgent Intervention Triage Strategy (ACUIty) trial investigators

Summary

Background The aim of this study was to assess anticoagulation with the direct thrombin inhibitor bivalirudin during percutaneous coronary intervention in individuals with moderate and high-risk acute coronary syndromes.

Methods 13 819 individuals in the Acute Catheterization and Urgent Intervention Triage strategy (ACUITY) trial were prospectively randomly assigned to receive heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors, bivalirudin plus glycoprotein IIb/IIIa inhibitors, or bivalirudin alone. Of these individuals, 7789 underwent percutaneous coronary intervention after angiography. The effect of the three regimens on the primary 30-day endpoints of composite ischaemia (death, myocardial infarction, or unplanned revascularisation for ischaemia), major bleeding, and net clinical outcomes (composite ischaemia or major bleeding) was assessed in this subgroup. Analyses were done by intention to treat. This trial is registered with ClinicalTrials.gov, with the number NCT00093158.

Findings Of the individuals who underwent percutaneous coronary intervention, 2561 received heparin plus glycoprotein IIb/IIIa inhibitors, 2609 received bivalirudin plus glycoprotein IIb/IIIa inhibitors, and 2619 received bivalirudin alone. 26 (0.3%) individuals dropped out or were lost to follow-up. There was no significant difference in the proportion of individuals with composite ischaemia, major bleeding, or net clinical outcomes at 30 days between those who received bivalirudin plus glycoprotein IIb/IIIa inhibitors and those who received heparin plus glycoprotein IIb/IIIa inhibitors (composite ischaemia: 243 [9%] patients vs 210 [8%] patients, p=0.016; major bleeding: 196 [8%] patients vs 174 [7%] patients, p=0.32; net clinical outcomes: 389 [15%] patients vs 341 [13%] patients, p=0.1). Rates of composite ischaemia were much the same in those who received bivalirudin alone and those who received heparin plus glycoprotein IIb/IIIa inhibitors (230 [9%] patients vs 210 [8%] patients, p=0.45); however, there were significantly fewer individuals who experienced major bleeding among those who received bivalirudin alone than among those who received heparin plus glycoprotein IIb/IIIa inhibitors (92 [4%] patients vs 174 [7%] patients, p=0.0001, relative risk 0.52, 95% CI 0.40–0.66), resulting in a trend towards better 30-day net clinical outcomes (303 [12%] patients vs 341 [13%] patients, p=0.057; 0.87, 0.75–1.00).

Interpretation Substitution of unfractionated heparin or enoxaparin with bivalirudin results in comparable clinical outcomes in patients with moderate and high-risk acute coronary syndromes treated with glycoprotein IIb/IIIa inhibitors in whom percutaneous coronary intervention is done. Anticoagulation with bivalirudin alone suppresses adverse ischaemic events to a similar extent as does heparin plus glycoprotein IIb/IIIa inhibitors, while significantly lowering the risk of major haemorrhagic complications.

Introduction Management of patients with acute coronary syndromes (ie, unstable angina or non-ST-segment elevation myocardial infarction) with a routine invasive strategy (early angiography followed by coronary revascularisation when appropriate) improves long-term survival while reducing the rates of re-infarction and rehospitalisation for recurrent unstable angina. In the USA and Europe, percutaneous coronary intervention before hospital discharge is done in about 40% of all patients with non-ST-segment elevation acute coronary syndromes and in 55% of those undergoing cardiac catheterisation. Complication rates are increased when percutaneous coronary interventions are done in patients with thrombotic ischaemic syndromes compared with when they are done in patients with stable coronary artery disease. Numerous clinical trials have been done to identify the optimum pharmacological regimen to enhance outcomes of such interventions in patients with acute coronary syndromes. Current guidelines recommend initiation of either unfractionated or low molecular mass heparin before angiography in such patients to reduce thrombin generation and clot propagation. The guidelines also recommend administration of aspirin, a thienopyridine platelet antagonist (started either before angiography or shortly after the percutaneous coronary intervention), and a glycoprotein IIb/IIIa inhibitor (started either upstream in all patients before angiography, or deferred for selective
initiation in the catheterisation laboratory just before the percutaneous coronary intervention).6–9 The combination of these potent antithrombotic and antiplatelet agents, although effective in suppressing ischaemic adverse events related to the percutaneous coronary intervention, also results in high rates of haemorrhagic complications, the occurrence of which have been strongly associated with early and late mortality.10–16

Bivalirudin is a direct-acting synthetic antithrombin that—by contrast with unfractioned and low molecular mass heparin—is active against clot-bound thrombin. The drug does not activate platelets, does not bind to plasma proteins, does not cause heparin-induced thrombocytopenia, and has linear pharmacokinetics with a short half-life of 25 min.17,18 Among 6010 patients with coronary artery disease undergoing percutaneous coronary intervention in the double-blind Randomized Evaluation of PCI Linking Angiomax to Reduced Clinical Events (REPLACE-2) trial, bivalirudin monotherapy was found to be non-inferior to unfractioned heparin plus the routine use of glycoprotein IIb/IIIa inhibitors in suppressing periprocedural ischaemic events; major and minor bleeding complications were also markedly reduced.19 However, high-risk patients with acute coronary syndromes undergoing percutaneous coronary interventions were excluded from REPLACE-2, including those with angiographic thrombus and those in whom glycoprotein IIb/IIIa inhibitors were required; the efficacy of bivalirudin monotherapy in such patients is unknown. Moreover, whether the addition of glycoprotein IIb/IIIa inhibitors to bivalirudin would safely result in a further decrease in ischaemic complications in high-risk patients is also unknown.

In the Acute Catheterization and Urgent Intervention Triage strategy (ACUITY) trial, 13819 patients with moderate and high-risk acute coronary syndromes undergoing an early invasive strategy were prospectively randomised to one of three antithrombin regimens: heparin (unfractioned or enoxaparin) plus glycoprotein IIb/IIIa inhibitors, bivalirudin plus glycoprotein IIb/IIIa inhibitors, or bivalirudin alone.20,21 Following angiography, percutaneous coronary interventions were done in 7789 patients (56%). Here, the results from this subgroup are presented.

### Methods

#### Patients

The design of the ACUITY trial has been described previously.20 In brief, to enrol a study population with moderate and high-risk acute coronary syndromes, patients aged 18 years or older with symptoms of unstable angina lasting for 10 min or more within the preceding 24 h were deemed to be eligible for enrolment if one or more of the following criteria were met: new ST-segment depression or transient elevation of 1 mm or more; raised troponin I, T, or creatine kinase MB isozyme; known coronary artery disease; or all four other unstable angina risk criteria as defined by the TIMI study group.21 Major exclusion criteria

### Figure 1: Trial profile
 included acute ST-segment elevation myocardial infarction or shock; bleeding diathesis or major bleeding episode within 2 weeks; thrombocytopenia; calculated creatinine clearance less than 30 mL per min; recent administration of abciximab, warfarin, fondaparinux, fibrinolytic agents, bivalirudin, or two or more doses of low molecular mass heparin; and allergy to study drugs or iodinated contrast that could not be adequately premedicated. The study was approved by the institutional review board or ethics committee at each participating centre, and all patients gave written, informed consent.

**Procedures**

Randomisation was done with a random number generator in blocks of six stratified by site and by the use of or intent to administer a thienopyridine before angiography. Patients were assigned with an interactive voice response system to one of three antithrombin regimens started immediately after randomisation: heparin (either unfractionated or enoxaparin) plus a glycoprotein IIb/IIIa inhibitor (the control group), bivalirudin plus a glycoprotein IIb/IIIa inhibitor, or bivalirudin alone. Unfractionated heparin was administered as an intravenous bolus of 60 IU/kg plus infusion of 12 IU/kg per hour to target an activated partial thromboplastin time of 50–75 s before angiography and an activated clotting time of 200–250 s during percutaneous coronary intervention. Enoxaparin 1 mg/kg twice daily delivered subcutaneously was administered before angiography, with an additional 0.3 mg/kg or 0.75 mg/kg intravenous bolus given before percutaneous coronary intervention if the most recent subcutaneous dose had been given more than 8 or 16 hours earlier, respectively. Bivalirudin was initiated with an intravenous bolus of 0.1–0.5 mg/kg and an infusion of 0.25–0.5 mg/kg per hour. An additional intravenous bolus of 0.5 mg/kg was administered before percutaneous coronary intervention, and the infusion was increased to 1.75 mg/kg per hour. All antithrombin agents were routinely discontinued per protocol at the completion of angiography or percutaneous coronary intervention. Of note, antithrombin monitoring either before or during percutaneous coronary intervention was not done in patients treated with enoxaparin or bivalirudin.

To determine the optimum strategy for use of glycoprotein IIb/IIIa inhibitors, patients assigned to a glycoprotein IIb/IIIa inhibitor-based regimen were randomised again in a 2x2 factorial design to upstream glycoprotein IIb/IIIa inhibitors or bivalirudin monotherapy for severe breakthrough ischaemia, and during the percutaneous coronary intervention in bivalirudin monotherapy patients for prespecified procedural complications.

Angiography was performed by protocol within 72 h after randomisation, after which the decision was declared and recorded for primary treatment either with a percutaneous coronary intervention, coronary artery bypass graft surgery, or medical management. Aspirin (300–325 mg orally or 250–500 mg intravenously) was administered daily during hospitalisation. The initial the percutaneous coronary intervention.

---

**Table 1: Baseline characteristics of patients undergoing percutaneous coronary intervention**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Heparin (unfractionated or enoxaparin plus glycoprotein IIb/IIIa inhibitors) (n=2561)</th>
<th>Bivalirudin plus glycoprotein IIb/IIIa inhibitors (n=2609)</th>
<th>Bivalirudin alone (n=2619)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>63 (25–91)</td>
<td>62 (21–95)</td>
<td>63 (30–92)</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>1880 (73%)</td>
<td>1919 (74%)</td>
<td>1919 (73%)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>703/2543 (28%)</td>
<td>713/2555 (27%)</td>
<td>721/2603 (28%)</td>
</tr>
<tr>
<td>Insulin requiring</td>
<td>205/2543 (8%)</td>
<td>208/2555 (8%)</td>
<td>224/2603 (9%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1673/2546 (66%)</td>
<td>1690/2594 (65%)</td>
<td>1714/2611 (66%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1409/2519 (56%)</td>
<td>1440/2555 (56%)</td>
<td>1436/2566 (56%)</td>
</tr>
<tr>
<td>Current smoker</td>
<td>770/2507 (31%)</td>
<td>797/2553 (31%)</td>
<td>795/2571 (31%)</td>
</tr>
<tr>
<td>Previous myocardial infarction</td>
<td>761/2506 (30%)</td>
<td>764/2549 (30%)</td>
<td>758/2565 (31%)</td>
</tr>
<tr>
<td>Previous percutaneous coronary intervention</td>
<td>973/2541 (38%)</td>
<td>978/2585 (38%)</td>
<td>973/2596 (40%)</td>
</tr>
<tr>
<td>Previous coronary artery bypass graft surgery</td>
<td>442/2555 (17%)</td>
<td>450/2605 (17%)</td>
<td>468/2613 (18%)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>84 (73–96)</td>
<td>84 (74–96)</td>
<td>84 (75–95)</td>
</tr>
<tr>
<td>Renal insufficiency*</td>
<td>457/2407 (19%)</td>
<td>453/2465 (18%)</td>
<td>441/2475 (18%)</td>
</tr>
<tr>
<td>Raised cardiac biomarker at baseline</td>
<td>1547/2376 (65%)</td>
<td>1555/2441 (64%)</td>
<td>1625/2450 (66%)</td>
</tr>
<tr>
<td>Raised baseline troponin</td>
<td>1436/2215 (65%)</td>
<td>1447/2310 (63%)</td>
<td>1513/2285 (66%)</td>
</tr>
<tr>
<td>Baseline ST-segment deviation ≥1 mm</td>
<td>907/2559 (35%)</td>
<td>958/2608 (37%)</td>
<td>923/2618 (35%)</td>
</tr>
<tr>
<td>Raised cardiac biomarker at baseline or ST-segment deviation</td>
<td>1874/2441 (77%)</td>
<td>1886/1506 (75%)</td>
<td>1931/2508 (77%)</td>
</tr>
<tr>
<td>TIMI risk score**</td>
<td>0–2 377/2242 (17%)</td>
<td>377/2228 (15%)</td>
<td>376/2346 (16%)</td>
</tr>
<tr>
<td></td>
<td>3–4 1179/2242 (52%)</td>
<td>1272/2228 (55%)</td>
<td>1248/2346 (53%)</td>
</tr>
<tr>
<td></td>
<td>5–7 706/2242 (31%)</td>
<td>699/2238 (30%)</td>
<td>722/2346 (31%)</td>
</tr>
</tbody>
</table>

Data are median (IQR) or n/N (%), except for age, where data are median (range). *Calculated creatinine clearance with the Cockcroft-Gault equation =60 mL per min.
Table 2: Study medications

<table>
<thead>
<tr>
<th>Study medications</th>
<th>Heparin (unfractionated or enoxaparin) plus glycoprotein Iib/IIIa inhibitors (n=2561)</th>
<th>Bivalirudin plus glycoprotein Iib/IIIa inhibitors (n=2609)</th>
<th>Bivalirudin alone (n=2619)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admission to PCI (hours)*</td>
<td>19.7 (6.3–30.5)</td>
<td>19.3 (6.3–29.2)</td>
<td>19.6 (6.3–30.7)</td>
</tr>
<tr>
<td>Admission to randomisation (hours)</td>
<td>5.8 (1.8–15.0)</td>
<td>5.4 (1.9–14.8)</td>
<td>5.3 (1.9–14.4)</td>
</tr>
<tr>
<td>Randomisation to antithrombin</td>
<td>0.4 (0.1–0.9)</td>
<td>0.6 (0.4–1.1)</td>
<td>0.6 (0.3–1.0)</td>
</tr>
<tr>
<td>study drug initiation‡ (hours)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antithrombin during PCI</td>
<td>4.1 (1.5–19.0)</td>
<td>3.9 (1.3–19.0)</td>
<td>4.1 (1.4–20.0)</td>
</tr>
<tr>
<td>Glycoprotein IIb/IIIa inhibitor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>admission before angiography§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin use or administration</td>
<td>1293/2554 (51%)</td>
<td>36/2606 (1%)</td>
<td>36/2614 (1%)</td>
</tr>
<tr>
<td>before PCI</td>
<td>1238/2554 (48%)</td>
<td>14/2606 (0.5%)</td>
<td>11/2614(0.4%)</td>
</tr>
<tr>
<td>Eptifi batide</td>
<td>23/2554 (0.9%)</td>
<td>2556/2606 (98%)</td>
<td>2567/2614 (98%)</td>
</tr>
<tr>
<td>Thienopyridine use or administration before PCI</td>
<td>2488/2540 (98%)</td>
<td>2530/2591 (98%)</td>
<td>2546/2597 (98%)</td>
</tr>
<tr>
<td>Glycoprotein Iib/IIIa inhibitor</td>
<td>230 (9%)</td>
<td>221 (8%)</td>
<td>10 (0%)</td>
</tr>
<tr>
<td>administration before randomisation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycoprotein Iib/IIIa inhibitor</td>
<td>1299 (51%)</td>
<td>1329 (51%)</td>
<td>18 (0.7%)</td>
</tr>
<tr>
<td>administration before angiography§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eptifi batide</td>
<td>814 (32%)</td>
<td>840 (32%)</td>
<td>11 (0.4%)</td>
</tr>
<tr>
<td>Ticlopidine</td>
<td>473 (18%)</td>
<td>476 (18%)</td>
<td>6 (0.2%)</td>
</tr>
<tr>
<td>Abciximab</td>
<td>12 (0.5%)</td>
<td>13 (0.5%)</td>
<td>1 (0.04%)</td>
</tr>
<tr>
<td>Glycoprotein Iib/IIIa inhibitor</td>
<td>2473 (97%)</td>
<td>2523 (97%)</td>
<td>238 (9%)</td>
</tr>
<tr>
<td>administration during PCI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eptifi batide</td>
<td>1526 (60%)</td>
<td>1588 (61%)</td>
<td>119 (5%)</td>
</tr>
<tr>
<td>Ticlopidine</td>
<td>492 (19%)</td>
<td>513 (20%)</td>
<td>20 (0.8%)</td>
</tr>
<tr>
<td>Abciximab</td>
<td>455 (18%)</td>
<td>425 (16%)</td>
<td>99 (4%)</td>
</tr>
</tbody>
</table>

Data are median (IQR), n/N (%), or n (%). PCI=percutaneous coronary intervention. *Data are not normally distributed, so median durations of subgroup do not add up to the duration of the parent group. †In patients not receiving the study antithrombin drug before randomisation. ‡In patients undergoing PCI. §Includes those patients in whom a glycoprotein Iib/IIIa inhibitor was administered after randomisation, but before angiography. ¶Three patients received both ticlopidine and abciximab.

Three primary 30-day endpoints were prespecified: composite ischaemia (death from any cause, myocardial infarction, or unplanned revascularisation for ischaemia), major bleeding (not related to coronary artery bypass graft surgery), and net clinical outcomes (composite ischaemia or major bleeding). The component definitions of the primary endpoints have been described in detail elsewhere. Major bleeding was defined as intracranial or intraocular bleeding, access site haemorrhage requiring intervention, 5 cm or more diameter haematoma, reduction in haemoglobin of 40 g/L or more without or with an overt bleeding source, reoperation for bleeding, or blood product transfusion. Stent thrombosis occurring within 30 days was defined as angiographic thrombus or subacute closure within the stented vessel at the time of clinically driven angiography for ischaemia, or in the absence of documented angiographic stent patency, any death not attributed to a non-cardiac cause, or any Q-wave myocardial infarction. A clinical events committee blinded to treatment assignment adjudicated all primary endpoint events and stent thrombosis, requiring original source document verification.

### Statistical analysis

The analyses presented here were a prespecified study of patients in the ACUITY trial triaged after angiography to management with percutaneous coronary interventions. Categorical variables were compared by χ² or Fisher’s exact test. Continuous variables were compared by the non-parametric Wilcoxon rank sum test. The power and sample size determination methodology for the ACUITY trial has been described previously. A minimum sample size for the percutaneous coronary intervention cohort was not prespecified. Given the use of a common control and two experimental groups, a two-sided α=0.025 was used for superiority testing. Formal non-inferiority testing was not done. All primary categorical binary event rate analyses were done in the intention-to-treat group, with no patient lost to follow-up excluded. A secondary analysis was done with time-to-event data (for which patients were censored at the time of study withdrawal or at last follow-up), displayed with Kaplan-Meier methodology and compared with the log-rank test. The effect of treatment assignment on the binary event rates of the composite ischaemia, major bleeding, and net clinical outcome endpoints in patients who underwent percutaneous coronary intervention were tested in multiple prespecified subgroups, with formal interaction testing done. All statistical analyses were done with SAS version 8.2.

This trial is registered with ClinicalTrials.gov, with the number NCT00093158.
Role of the funding source
The trial was sponsored and funded by the Medicines Company and Nycomed. The sponsors were involved in study design and in data collection, analysis, and interpretation, along with the principal investigator and the executive and steering committees.20,21 The corresponding author had full access to all the data in the study. The manuscript was prepared by the corresponding author and revised by all coauthors. The authors controlled the decision to submit the paper for publication. The sponsors provided the opportunity for a non-binding review of the manuscript before submission.

Results
Of 13 819 enrolled patients, 7789 (56%) were managed with a percutaneous coronary intervention strategy; 1539 (11%) with medical therapy (figure 1). Percutaneous coronary intervention was attempted in a median time of 19–5 hours after admission in 2561 (56%) of the patients assigned to heparin plus glycoprotein IIb/IIIa inhibitors, 2609 (57%) of those assigned to bivalirudin plus glycoprotein IIb/IIIa inhibitors, and 2619 (57%) of those assigned to bivalirudin alone. Baseline characteristics were much the same in the three groups undergoing percutaneous coronary intervention (table 1). Median patient age was 63 years, 5698 (73%) were male, and 2137 (28%) had diabetes. Non-ST-segment elevation myocardial infarction (raised baseline creatine kinase MB isozyme or troponin) was present in 4728 (65%) patients; 2539 (35%) had unstable angina.

About half of the patients in the control group underwent percutaneous coronary intervention with unfractionated heparin, while half received enoxaparin (table 2). Of the patients receiving unfractionated heparin, the median maximum activated clotting time during percutaneous coronary intervention was 239 (IQR 211–291) seconds. Eptifibatide was the most common glycoprotein IIb/IIIa inhibitor during percutaneous coronary intervention with bivalirudin either with or without glycoprotein IIb/IIIa inhibitors compared with heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors (table 4). Similarly, there were no significant differences in the rates of angiographic complications at procedure end between the three groups undergoing percutaneous coronary intervention with bivalirudin either with or without glycoprotein IIb/IIIa inhibitors compared with heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors (table 2; figure 2).

By blinded core laboratory analysis, the baseline and post-procedural rates of TIMI flow23 and myocardial blush were not significantly different in patients undergoing percutaneous coronary intervention with bivalirudin either with or without glycoprotein IIb/IIIa inhibitors compared with heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors (table 5). Similarly, there were no significant differences in the rates of angiographic complications at procedure end between the three treatment groups. There were no reported occurrences of catheter-related thrombus in any of the three groups.

Of the patients undergoing percutaneous coronary intervention, the proportion of individuals with composite ischaemia, major bleeding, and net clinical outcomes at 30 days were not significantly different in those who received bivalirudin plus glycoprotein IIb/IIIa inhibitors compared with those who received heparin plus glycoprotein IIb/IIIa inhibitors (table 5 and figure 2; composite ischaemia: 243 [9%] patients vs 210 [8%] patients, p=0·16; relative risk [RR] 1·14, 95% CI 0·95–1·36; major bleeding: 196 [8%] patients vs 174 [7%] patients, p=0·32;
that was unrelated to coronary artery bypass graft surgery was seen in fewer patients receiving bivalirudin monotherapy compared with those receiving heparin with glycoprotein IIb/IIIa inhibitors, even when haematomas larger than 5 cm in diameter were removed from the endpoint definition (81/2619 [3%] vs 136/2561 [5%], p=0·0001). Stent thrombosis within 30 days of stent implantation occurred in 103 (1·4%) patients, with similar rates among the three treatment groups (table 5).

The treatment effects of bivalirudin monotherapy were consistent across multiple prespecified subgroups, with no significant interactions noted with regard to age, gender, diabetic status, baseline risk measures (including raised troponins, ST-segment deviation, or TIMI unstable angina risk score), antithrombin crossovers, upstream versus deferred glycoprotein IIb/IIIa inhibitor randomisation, and time to intervention (figure 3, figure 4, and figure 5). There was a non-significant trend towards an interaction effect based on the timing of thienopyridine administration (p=0·08); the incidence of 30-day composite ischaemic events was much the same in the bivalirudin monotherapy

### Table 4: Core angiographic laboratory analysis

<table>
<thead>
<tr>
<th></th>
<th>Heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors</th>
<th>Bivalirudin plus glycoprotein IIb/IIIa inhibitors</th>
<th>p value*</th>
<th>Bivalirudin alone</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TIMI flow at baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/1</td>
<td>1389 vessels/1183 patients</td>
<td>1467 vessels/1250 patients</td>
<td>0·47</td>
<td>1421 vessels/1212 patients</td>
<td>0·48</td>
</tr>
<tr>
<td>2</td>
<td>185 (13%)</td>
<td>182 (12%)</td>
<td></td>
<td>193 (14%)</td>
<td>0·84</td>
</tr>
<tr>
<td>3</td>
<td>134 (10%)</td>
<td>140 (10%)</td>
<td>0·92</td>
<td>141 (10%)</td>
<td>0·71</td>
</tr>
<tr>
<td>Corrected TIMI frame count‡</td>
<td>1070 (77%)</td>
<td>1145 (78%)</td>
<td>0·52</td>
<td>1085 (76%)</td>
<td>0·67</td>
</tr>
<tr>
<td><strong>TIMI flow after PCI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/1</td>
<td>1384 vessels/1177 patients</td>
<td>1461 vessels/1441 patients</td>
<td>0·24</td>
<td>1419 vessels/1211 patients</td>
<td>0·29</td>
</tr>
<tr>
<td>2</td>
<td>24 (2%)</td>
<td>15 (1%)</td>
<td></td>
<td>29 (2%)</td>
<td>0·29</td>
</tr>
<tr>
<td>3</td>
<td>1339 (97%)</td>
<td>1425 (98%)</td>
<td>0·21</td>
<td>1353 (95%)</td>
<td>0·06</td>
</tr>
<tr>
<td>Corrected TIMI frame count‡</td>
<td>32·4 (16·5)</td>
<td>31·9 (17·3)</td>
<td>0·34</td>
<td>32·1 (16·9)</td>
<td>0·48</td>
</tr>
<tr>
<td><strong>Baseline myocardial blush grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/1</td>
<td>1211 vessels/1058 patients</td>
<td>1304 vessels/1133 patients</td>
<td>0·18</td>
<td>1239 vessels/1080 patients</td>
<td>0·93</td>
</tr>
<tr>
<td>2</td>
<td>206 (17%)</td>
<td>196 (15%)</td>
<td></td>
<td>209 (17%)</td>
<td>0·93</td>
</tr>
<tr>
<td>3</td>
<td>186 (15%)</td>
<td>201 (15%)</td>
<td>0·97</td>
<td>208 (17%)</td>
<td>0·34</td>
</tr>
<tr>
<td>Myocardial blush grade after PCI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/1</td>
<td>819 (68%)</td>
<td>907 (70%)</td>
<td>0·30</td>
<td>822 (66%)</td>
<td>0·50</td>
</tr>
<tr>
<td>3</td>
<td>1145 vessels/1003 patients</td>
<td>1252 vessels/1092 patients</td>
<td>0·12</td>
<td>1168 vessels/1020 patients</td>
<td>0·84</td>
</tr>
<tr>
<td>Corrected TIMI frame count‡</td>
<td>24·3 (12·8)</td>
<td>24·1 (11·3)</td>
<td>0·92</td>
<td>23·6 (11·8)</td>
<td>0·17</td>
</tr>
<tr>
<td><strong>Adverse angiographic events after PCI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=1181 patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New or increased thrombus</td>
<td>18 (2%)</td>
<td>11 (0·9%)</td>
<td>0·15</td>
<td>9 (0·7%)</td>
<td>0·07</td>
</tr>
<tr>
<td>Abrupt closure</td>
<td>2 (0·2%)</td>
<td>3 (0·2%)</td>
<td>0·50</td>
<td>0·55</td>
<td></td>
</tr>
<tr>
<td>No reflow</td>
<td>0 (0%)</td>
<td>2 (0·2%)</td>
<td>0·17</td>
<td>3 (0·2%)</td>
<td>0·09</td>
</tr>
<tr>
<td>Distal embolisation</td>
<td>10 (0·8%)</td>
<td>5 (0·4%)</td>
<td>0·16</td>
<td>12 (1%)</td>
<td>0·72</td>
</tr>
<tr>
<td>Spasm</td>
<td>6 (0·5%)</td>
<td>7 (0·6%)</td>
<td>0·85</td>
<td>13 (1·1%)</td>
<td>0·12</td>
</tr>
<tr>
<td>Dissection</td>
<td>8 (0·7%)</td>
<td>11 (0·9%)</td>
<td>0·56</td>
<td>10 (0·8%)</td>
<td>0·68</td>
</tr>
<tr>
<td>Perforation</td>
<td>1 (0·1%)</td>
<td>1 (0·1%)</td>
<td>0·97</td>
<td>0 (0%)</td>
<td>0·31</td>
</tr>
</tbody>
</table>

Data are n (%) unless stated otherwise. PCI=percutaneous coronary intervention. TIMI=thrombolysis in myocardial infarction. *Comparison between bivalirudin plus IIb/IIIa inhibitors and heparin (unfractionated or enoxaparin) plus IIb/IIIa inhibitors. †Comparison between bivalirudin alone and heparin (unfractionated or enoxaparin) plus IIb/IIIa inhibitors. Data are cineframes, mean (SD), where lower numbers represent more rapid blood flow in the coronary artery to reach standardised distal landmark.
and heparin plus glycoprotein IIb/IIIa inhibitor groups when thienopyridines were administered before percutaneous coronary intervention (145/1789 [8%] vs 145/1722 [8%]; RR 0·96, 0·77–1·20; figure 3); by contrast, such events tended to be more frequent with bivalirudin monotherapy in patients receiving a thienopyridine after percutaneous coronary intervention (73/762 [10%] vs 57/766 [7%]; 1·29, 0·92–1·79) and in those not receiving a thienopyridine at all (10/42 [24%] vs 4/45 [9%]; 2·68, 0·91–7·89). There were fewer cases of major bleeding with bivalirudin monotherapy than with heparin plus glycoprotein IIb/IIIa inhibitors, irrespective of thienopyridine pretreatment (figure 4). Among 1338 troponin-positive patients who received clopidogrel pretreatment (figure 4), Among 1338 troponin-positive patients who received clopidogrel pretreatment, rates of composite ischaemia were much the same with bivalirudin monotherapy as they were with heparin plus initiation of glycoprotein IIb/IIIa inhibitors in the catheterisation laboratory, but there was less major bleeding (composite ischaemia: 8% in both groups [79/975 and 31/383, respectively]; 1·00, 0·67–1·49; major bleeding: 42/975 [4%] vs 31/383 [8%]; 0·53, 0·34–0·83; figure 3 and figure 4).

By core laboratory analysis, thrombus was present on the baseline angiogram in 712 (19%) of 3664 patients, including 222 patients assigned to heparin plus glycoprotein IIb/IIIa inhibitors, 241 assigned to bivalirudin with glycoprotein IIb/IIIa inhibitors, and 249 patients assigned to bivalirudin monotherapy. In these patients the respective 30-day rates of composite ischaemia were 12% (26/222 patients), 8% (20/241 patients), and 13% (33/249 patients; p=0·22 for comparison of heparin plus glycoprotein IIb/IIIa inhibitors vs bivalirudin with glycoprotein IIb/IIIa inhibitors; p=0·61 for comparison of heparin plus glycoprotein IIb/IIIa inhibitors vs bivalirudin alone), and the 30-day rates of major bleeding unrelated to coronary artery bypass graft surgery were 7% (16/222 patients), 6% (15/241 patients), and 3% (7/249 patients; p=0·67 for comparison of heparin plus glycoprotein IIb/IIIa inhibitors vs bivalirudin with glycoprotein IIb/IIIa inhibitors; p=0·03 for comparison of heparin plus glycoprotein IIb/IIIa inhibitors vs bivalirudin alone). The 30-day rates of net clinical outcomes in patients with

<table>
<thead>
<tr>
<th></th>
<th>Heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors (n=2561)</th>
<th>Bivalirudin plus glycoprotein IIb/IIIa inhibitors (n=2609)</th>
<th>p value* Bivalirudin alone (n=2619)</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net clinical outcomes</td>
<td>341 (13%)</td>
<td>389 (15%)</td>
<td>0·10</td>
<td>303 (12%)</td>
</tr>
<tr>
<td>Composite ischaemia</td>
<td>210 (8%)</td>
<td>243 (9%)</td>
<td>0·16</td>
<td>230 (9%)</td>
</tr>
<tr>
<td>Death from any cause</td>
<td>23 (0·9%)</td>
<td>30 (1·2%)</td>
<td>0·37</td>
<td>28 (1·2%)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>144 (6%)</td>
<td>171 (7%)</td>
<td>0·16</td>
<td>170 (6%)</td>
</tr>
<tr>
<td>Q-wave</td>
<td>25 (1·2%)</td>
<td>25 (1·2%)</td>
<td>0·95</td>
<td>30 (1·2%)</td>
</tr>
<tr>
<td>Non Q-wave</td>
<td>119 (5%)</td>
<td>147 (6%)</td>
<td>0·11</td>
<td>141 (5%)</td>
</tr>
<tr>
<td>Unplanned revascularisation for ischaemia</td>
<td>81 (3%)</td>
<td>96 (4%)</td>
<td>0·31</td>
<td>85 (3%)</td>
</tr>
<tr>
<td>PCI</td>
<td>61 (2%)</td>
<td>76 (3%)</td>
<td>0·23</td>
<td>65 (2%)</td>
</tr>
<tr>
<td>CABG</td>
<td>21 (0·8%)</td>
<td>24 (0·9%)</td>
<td>0·70</td>
<td>23 (0·9%)</td>
</tr>
<tr>
<td>Stent thrombosis</td>
<td>31 (1%)</td>
<td>39 (1%)</td>
<td>0·38</td>
<td>33 (1%)</td>
</tr>
<tr>
<td>Major bleeding (non-CABG related)</td>
<td>174 (7%)</td>
<td>196 (8%)</td>
<td>0·32</td>
<td>92 (4%)</td>
</tr>
<tr>
<td>Intracranial</td>
<td>1 (0%)</td>
<td>1 (0%)</td>
<td>0·99</td>
<td>1 (0%)</td>
</tr>
<tr>
<td>Retropertitoneal</td>
<td>18 (0·7%)</td>
<td>23 (0·9%)</td>
<td>0·47</td>
<td>5 (0·2%)</td>
</tr>
<tr>
<td>Access site haemorrhage</td>
<td>84 (3%)</td>
<td>94 (4%)</td>
<td>0·52</td>
<td>25 (1%)</td>
</tr>
<tr>
<td>Requiring intervention or surgery</td>
<td>15 (0·6%)</td>
<td>22 (0·8%)</td>
<td>0·27</td>
<td>8 (0·3%)</td>
</tr>
<tr>
<td>Haematoma &gt;5 cm</td>
<td>74 (3%)</td>
<td>80 (3%)</td>
<td>0·71</td>
<td>23 (0·8%)</td>
</tr>
<tr>
<td>Haemoglobin decrease &gt;30 g/L with overt source</td>
<td>71 (3%)</td>
<td>75 (3%)</td>
<td>0·82</td>
<td>37 (1%)</td>
</tr>
<tr>
<td>Haemoglobin decrease &gt;40 g/L with no overt source</td>
<td>25 (1%)</td>
<td>29 (1%)</td>
<td>0·63</td>
<td>24 (0·9%)</td>
</tr>
<tr>
<td>Blood transfusion</td>
<td>77 (3%)</td>
<td>103 (4%)</td>
<td>0·09</td>
<td>44 (2%)</td>
</tr>
<tr>
<td>Reoperation for bleed</td>
<td>1 (0%)</td>
<td>5 (0·2%)</td>
<td>0·11</td>
<td>3 (0·1%)</td>
</tr>
<tr>
<td>Minor bleeding (non-CABG related)</td>
<td>665 (26%)</td>
<td>740 (28%)</td>
<td>0·053</td>
<td>393 (15%)</td>
</tr>
<tr>
<td>TIMI scale bleeding‡</td>
<td>200 (8%)</td>
<td>222 (9%)</td>
<td>0·36</td>
<td>118 (5%)</td>
</tr>
<tr>
<td>TIMI major bleeding</td>
<td>58 (2%)</td>
<td>63 (2%)</td>
<td>0·72</td>
<td>22 (0·8%)</td>
</tr>
<tr>
<td>TIMI minor bleeding</td>
<td>193 (8%)</td>
<td>213 (8%)</td>
<td>0·40</td>
<td>109 (4%)</td>
</tr>
<tr>
<td>Any thrombocytopenia (acquired)§</td>
<td>189 (7%)</td>
<td>207 (8%)</td>
<td>0·45</td>
<td>176 (7%)</td>
</tr>
</tbody>
</table>

CABG=coronary artery bypass graft surgery. PCI=percutaneous coronary intervention. TIMI=thrombolysis in myocardial infarction. *Comparison between bivalirudin plus glycoprotein IIb/IIIa inhibitors and heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors. †Comparison between bivalirudin alone and heparin (unfractionated or enoxaparin) plus glycoprotein IIb/IIIa inhibitors. ‡Ecchymoses, epistaxis, gastrointestinal, genitourinary, puncture site, haemopericardium, pulmonary, or other; §Platelet count <150 000 cells per µL in patients without baseline thrombocytopenia.

Table 5: Clinical outcomes at 30 days
Figure 2: Time-to-event curves for composite ischaemia, and major bleeding not related to coronary artery bypass graft surgery, and net clinical outcomes

The primary endpoints were measured at 30 days (range 25–30); thus follow-up is reported to 35 days. P values bypass graft surgery, and net clinical outcomes estimates based on time to event data and log rank p values vary slightly from the binary event rate data and 

Discussion

This subgroup analysis of ACUITY indicates that, in patients with moderate and high-risk acute coronary syndromes undergoing contemporary percutaneous coronary interventions with drug-eluting and bare metal stents, the inclusion of bivalirudin as part of a glycoprotein IIb/IIIa inhibitor-based procedural anticoagulation regimen results in similar net clinical outcomes as does the use of unfractionated heparin or enoxaparin. Furthermore, the use of bivalirudin alone in these patients results in suppression of periprocedural ischaemic events similar to that achieved with unfractionated heparin or enoxaparin with routine IIb/IIIa inhibitors, but markedly lower risk of major and minor bleeding complications. Thus, as a pharmacological adjunct in patients with acute coronary syndromes undergoing percutaneous coronary intervention using contemporary drugs and devices, bivalirudin monotherapy is comparable with the combination of heparin plus glycoprotein IIb/IIIa inhibitors in the prevention of periprocedural ischaemic complications, while significantly lowering the risk of major and minor bleeding, as well as the need for blood transfusions.

In the large-scale, double-blind randomised Bivalirudin Angioplasty Trial, bivalirudin monotherapy was found to reduce the occurrence of periprocedural composite ischaemia (death, myocardial infarction, or repeat revascularisation) and major bleeding, compared with unfractionated heparin, in patients with acute coronary syndromes undergoing percutaneous coronary intervention.23 However, this trial was done in the early 1990s, before the widespread use of stents, clopidogrel, and glycoprotein IIb/IIIa inhibitors, each of which is now considered standard of care in acute coronary syndromes. The PROTECT-TIMI-30 trial randomised 857 patients with non-ST-segment elevation acute coronary syndromes to unfractionated heparin plus eptifibatide, enoxaparin plus eptifibatide, or bivalirudin monotherapy. The primary endpoint of coronary flow reserve after percutaneous coronary intervention was significantly improved with bivalirudin in this study, although myocardial blush was more often normalised with an eptifibatide-based regimen. There was no significant difference in the composite endpoint of death or myocardial infarction at 48 h between the groups, although TIMI scale bleeding and the need for blood transfusions were markedly reduced with bivalirudin monotherapy, in much the same way as seen in the ACUITY trial.23 The results presented here thus confirm and extend the results of these earlier trials.

The results with bivalirudin monotherapy in patients with moderate and high-risk acute coronary syndromes undergoing percutaneous coronary intervention in ACUITY—an open-label randomised trial—are much the same as those seen in the double-blind randomised REPLACE-2 trial, in which patients with stable coronary syndromes and lower risk acute coronary syndromes were enrolled.24 In REPLACE-2, bivalirudin monotherapy (with provisional use of glycoprotein IIb/IIIa inhibitors in about
7% of patients) resulted in a non-significant 9% relative increase in composite ischaemia at 30 days and a significant 41% relative decrease in in-hospital major bleeding compared with unfractionated heparin plus routine administration of either abciximab or eptifibatide. In the present trial, bivalirudin monotherapy (with provisional use of glycoprotein IIb/IIIa inhibitors in about 9% of patients) resulted in a non-significant 7% relative increase in the risk of composite ischaemia and a significant 48% relative decrease in the risk of major bleeding at 30 days compared with heparin (unfractionated or enoxaparin) plus routine administration of either abciximab, eptifibatide, or tirofiban. One should note that the definition of major bleeding was different in REPLACE-2.

Figure 6 shows the results of a pooled analysis of three prospective, contemporary randomised trials in which bivalirudin monotherapy with provisional use of glycoprotein IIb/IIIa inhibitors was compared with heparin (unfractionated or enoxaparin) plus the routine use of glycoprotein IIb/IIIa inhibitors in patients undergoing percutaneous coronary intervention for either moderate or high-risk acute coronary syndromes (the percutaneous coronary intervention subgroup of ACUITY and PROTECT-TIMI-30) or for more stable coronary syndromes (REPLACE-2). There is a non-significant trend toward a 9% increase in composite ischaemia and significant 45% and 44% reductions in TIMI scale bleeding and the need for blood transfusions, respectively, with bivalirudin monotherapy, independent of clinical syndrome acuity.

Coronary blood flow velocity and myocardial perfusion after percutaneous coronary interventions in patients with acute coronary syndromes have previously been shown to correlate with peri-procedural ischaemic outcomes after the percutaneous coronary intervention. Angiographic complications and residual thrombus after such interventions in patients with acute coronary syndromes also have negative prognostic implications. In the ACUITY blinded angiographic core laboratory substudy, no significant differences were seen in the rates of TIMI flow, TIMI frame counts, myocardial blush, or angiographic complications after percutaneous coronary intervention in patients randomly assigned to heparin with IIb/IIIa inhibitors compared with those assigned to bivalirudin with or without glycoprotein IIb/IIIa inhibitors. This analysis of 3664 patients thus provides a mechanistic foundation for the finding of similar rates of 30-day ischaemia with the three adjunctive pharmacological regimens.

The observation of lower rates of major bleeding after percutaneous coronary intervention with bivalirudin monotherapy than with heparin plus glycoprotein IIb/IIIa inhibitors is potentially important, since numerous studies have established a strong link between iatrogenic haemorrhagic complications and transfusions in acute coronary syndromes and percutaneous coronary interventions with subsequent mortality. Major bleeding is also a strong determinant of in-hospital cost after
Table: Comparison of patients randomly assigned to heparin plus glycoprotein IIb/IIIa inhibitors vs bivalirudin monotherapy

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Bivalirudin alone</th>
<th>Unfractionated heparin/ enoxaparin plus glycoprotein IIb/IIIa inhibitors</th>
<th>RR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (n=5180)</td>
<td>3.5%</td>
<td>6.8%</td>
<td>0.52 (0.40-0.66)</td>
<td>..</td>
</tr>
<tr>
<td>Age &lt;65 years (n=2829)</td>
<td>2.4%</td>
<td>5.0%</td>
<td>0.48 (0.32-0.71)</td>
<td>0.63</td>
</tr>
<tr>
<td>Age &gt;65 years (n=2351)</td>
<td>4.9%</td>
<td>8.8%</td>
<td>0.55 (0.41-0.76)</td>
<td>0.06</td>
</tr>
<tr>
<td>Men (n=3779)</td>
<td>2.5%</td>
<td>4.9%</td>
<td>0.51 (0.36-0.72)</td>
<td>1.00</td>
</tr>
<tr>
<td>Women (n=1401)</td>
<td>6.3%</td>
<td>11.3%</td>
<td>0.53 (0.37-0.75)</td>
<td>..</td>
</tr>
<tr>
<td>Diabetes (n=1424)</td>
<td>4.6%</td>
<td>8.5%</td>
<td>0.54 (0.36-0.81)</td>
<td>0.86</td>
</tr>
<tr>
<td>No diabetes (n=3722)</td>
<td>3.1%</td>
<td>6.2%</td>
<td>0.51 (0.37-0.69)</td>
<td>..</td>
</tr>
<tr>
<td>CrCl ≥60 mL per min (n=3984)</td>
<td>2.9%</td>
<td>5.5%</td>
<td>0.52 (0.38-0.71)</td>
<td>0.70</td>
</tr>
<tr>
<td>CrCl &lt;60 mL per min (n=898)</td>
<td>7.0%</td>
<td>11.8%</td>
<td>0.59 (0.39-0.89)</td>
<td>..</td>
</tr>
<tr>
<td>US (n=2800)</td>
<td>3.7%</td>
<td>6.2%</td>
<td>0.60 (0.43-0.84)</td>
<td>0.20</td>
</tr>
<tr>
<td>Non-US (n=2380)</td>
<td>3.2%</td>
<td>7.5%</td>
<td>0.44 (0.30-0.63)</td>
<td>..</td>
</tr>
<tr>
<td>Troponin raised (n=2949)</td>
<td>4.2%</td>
<td>7.0%</td>
<td>0.59 (0.44-0.80)</td>
<td>0.11</td>
</tr>
<tr>
<td>Troponin normal (n=1551)</td>
<td>2.5%</td>
<td>6.8%</td>
<td>0.36 (0.22-0.61)</td>
<td>..</td>
</tr>
<tr>
<td>ST deviation (n=1830)</td>
<td>4.9%</td>
<td>8.6%</td>
<td>0.57 (0.40-0.81)</td>
<td>0.54</td>
</tr>
<tr>
<td>No ST deviation (n=3347)</td>
<td>2.8%</td>
<td>5.8%</td>
<td>0.48 (0.34-0.67)</td>
<td>..</td>
</tr>
<tr>
<td>TIMI risk score 0–2 (n=753)</td>
<td>2.7%</td>
<td>5.6%</td>
<td>0.48 (0.23-1.00)</td>
<td>..</td>
</tr>
<tr>
<td>3–4 (n=2427)</td>
<td>3.2%</td>
<td>5.8%</td>
<td>0.55 (0.38-0.81)</td>
<td>0.84</td>
</tr>
<tr>
<td>5–7 (n=1428)</td>
<td>4.4%</td>
<td>9.2%</td>
<td>0.48 (0.32-0.73)</td>
<td>..</td>
</tr>
<tr>
<td>Thienopyridine administration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-PCI (n=1528)</td>
<td>3.6%</td>
<td>7.2%</td>
<td>0.50 (0.37-0.67)</td>
<td>..</td>
</tr>
<tr>
<td>Post-PCI (n=1528)</td>
<td>3.7%</td>
<td>5.5%</td>
<td>0.67 (0.42-1.07)</td>
<td>0.56</td>
</tr>
<tr>
<td>None (n=87)</td>
<td>0.0%</td>
<td>8.9%</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Upstream IB/IIa inhibitor (n=3691)</td>
<td>3.5%</td>
<td>7.0%</td>
<td>0.50 (0.38-0.67)</td>
<td>..</td>
</tr>
<tr>
<td>Deferred IB/IIa inhibitor (n=3908)</td>
<td>3.5%</td>
<td>6.6%</td>
<td>0.53 (0.40-0.71)</td>
<td>NA</td>
</tr>
<tr>
<td>Randomisation to PCI, tertiles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early (+2–7 h) (n=1696)</td>
<td>3.4%</td>
<td>7.0%</td>
<td>0.48 (0.31-0.75)</td>
<td>..</td>
</tr>
<tr>
<td>Intermediate (+2–16.5 h) (n=1719)</td>
<td>3.2%</td>
<td>5.6%</td>
<td>0.57 (0.36-0.90)</td>
<td>0.87</td>
</tr>
<tr>
<td>Late (+16.5 h) (n=1719)</td>
<td>4.1%</td>
<td>7.6%</td>
<td>0.54 (0.36-0.80)</td>
<td>..</td>
</tr>
<tr>
<td>Antithrombin crossovers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No prior antithrombin (n=1687)</td>
<td>3.2%</td>
<td>6.2%</td>
<td>0.52 (0.33-0.82)</td>
<td>..</td>
</tr>
<tr>
<td>Consistent therapy (n=3026)</td>
<td>3.2%</td>
<td>6.9%</td>
<td>0.46 (0.31-0.69)</td>
<td>NA</td>
</tr>
<tr>
<td>Crossover (n=1883)</td>
<td>3.7%</td>
<td>6.8%</td>
<td>0.52 (0.32-0.92)</td>
<td>..</td>
</tr>
<tr>
<td>DES only implanted (n=3090)</td>
<td>3.6%</td>
<td>6.3%</td>
<td>0.57 (0.41-0.78)</td>
<td>..</td>
</tr>
<tr>
<td>BMS only implanted (n=1666)</td>
<td>3.0%</td>
<td>8.1%</td>
<td>0.37 (0.24-0.58)</td>
<td>0.13</td>
</tr>
<tr>
<td>No stent implanted (n=376)</td>
<td>5.8%</td>
<td>6.5%</td>
<td>0.90 (0.41-1.98)</td>
<td>..</td>
</tr>
</tbody>
</table>

Figure 4: Selected subgroup analyses for the 30-day rates of major bleeding
Comparison of patients randomly assigned to heparin plus glycoprotein IIb/IIIa inhibitors vs bivalirudin monotherapy, displayed as relative risk (RR; black boxes) with 95% CI (horizontal lines). p value is the p value for the interaction between the variable and the relative treatment effect. BMS=bare metal stent. CrCl=baseline creatinine clearance. DES=drug-eluting stent. PCI=percutaneous coronary intervention. CABG=coronary artery bypass graft surgery. NA=interaction test not applicable.

percutaneous coronary interventions. Data presented here suggest that there was not only less access site bleeding, but also fewer non-access site major bleeding complications, less TIMI major and minor bleeding, and a reduction in the need for blood transfusions with bivalirudin monotherapy than with a glycoprotein IIb/IIIa inhibitor-based regimen. In REPLACE-2, the reduction in major bleeding in concert with similar rates of 30-day ischaemia resulted in a trend towards lower 1-year mortality in patients undergoing percutaneous coronary interventions with bivalirudin monotherapy rather than with heparin plus glycoprotein IIb/IIIa inhibitors. Longer term follow-up from the present study is required to confirm these findings.

Bivalirudin monotherapy suppressed periprocedural ischaemic events in numerous subgroups examined in much the same way as did heparin plus glycoprotein IIb/IIIa inhibitors, including in high-risk patients such as those with TIMI risk score 5–7, baseline ST-segment deviation, and positive troponin levels. In the ISAR-REACT-2 trial, the addition of abciximab initiated in the cardiac catheterisation laboratory to clopidogrel-pretreated troponin-positive patients further reduced the 30-day rate of ischaemic complications compared with heparin anticoagulation alone. By contrast, in the present study, the 30-day rates of composite ischaemia were identical in those individuals who received bivalirudin monotherapy compared with those who received heparin plus glycoprotein IIb/IIIa inhibitors initiated in the catheterisation laboratory in clopidogrel-pretreated troponin-positive patients; major bleeding was significantly lower. The efficacy of bivalirudin monotherapy was also independent of age, sex, diabetes, renal function, previous antithrombin use, timing of glycoprotein IIb/IIIa inhibitor administration, and stent type. In patients in whom a thienopyridine was not administered before the percutaneous coronary intervention, a non-significantly higher 30-day rate of composite ischaemia was seen with bivalirudin monotherapy than with heparin plus a glycoprotein IIb/IIIa inhibitor. This difference was most marked in the small subset of patients in whom a thienopyridine was never administered. However, given the number of subgroups examined and the borderline significance of this finding (p=0.08 for interaction), caution against over-interpretation is warranted, in view of the risk of a spurious finding. Finally, there was less major bleeding with bivalirudin monotherapy than in the control group in all subgroups examined, including patients with and without thienopyridine pretreatment.

No subgroup was identified in which the outcomes of patients randomly assigned to bivalirudin plus glycoprotein IIb/IIIa inhibitors were better than with heparin plus glycoprotein IIb/IIIa inhibitors or bivalirudin alone, with the possible exception of patients with thrombus. Among 712 randomised patients in whom the angiographic core laboratory determined thrombus to be present before the percutaneous coronary intervention, the 30-day rate of
composite ischaemia tended to be lowest with bivalirudin plus glycoprotein IIb/IIIa inhibitors (8% vs 12% and 13% respectively), although this finding was not significant. However, rates of major bleeding were lowest in patients randomised to bivalirudin monotherapy. Thus, although definitive recommendations cannot be provided, in patients with acute coronary syndromes started on upstream bivalirudin in whom subsequent angiography reveals thrombus, the addition of abciximab or eptifibatide before percutaneous coronary intervention could reduce ischaemic complications to a minimum, although at the cost of increased bleeding.

The general limitations of the ACUITY trial have been discussed in detail elsewhere. One should note that randomisation occurred before angiography, and study drugs were administered at a median of 4 h before the percutaneous coronary intervention. Moreover, the percutaneous coronary intervention subgroup represents a subset of 56% of all patients enrolled in ACUITY, and randomisation was not stratified by treatment assignment. Since the complexity of the study precluded blinding the numerous pharmacological agents administered, the potential for group imbalance or treatment bias is of concern. However, percutaneous coronary interventions were done in a similar proportion of patients in each randomised group, and the procedural techniques used were also well balanced; the baseline clinical and angiographic characteristics of the three groups who underwent percutaneous coronary intervention (including the time from admission to percutaneous coronary intervention) were similarly matched; and the rate of provisional use of glycoprotein IIb/IIIa inhibitors for ischaemic complications was almost identical to that in the blinded REPLACE-2 trial, suggesting that there was no substantial bias despite the open-label design, and that the results of the study as reported would be unlikely to be affected by propensity or multivariable analyses. Second, to represent the global standard of care and everyday use of these agents have never been directly compared for this application, the use of both was allowed. Similarly, both abciximab and double-bolus eptifibatide, when initiated in the catheterisation laboratory in patients with acute coronary syndromes, reduce ischaemic complications, and in the absence of a comparative trial, both are recommended and widely used for this purpose. Third, the analysis of the ACUITY percutaneous coronary intervention cohort was under-powered for non-inferiority...
testing, and subgroup analysis was not adjusted for multiple comparisons; as such, the present results should be considered to be hypothesis-generating. Fourth, the present trial included patients with a calculated creatinine clearance of 30 mL per min or more, a range for which dose adjustment of bivalirudin is not required. Further study to determine whether the results of the present trial can be extended to patients with more severe renal insufficiency is warranted; in such patients, the half-life of bivalirudin is substantially prolonged, thus necessitating a reduced infusion rate. Finally, the results from an ongoing formal cost-effectiveness analysis, as well as the 1-year data from ACUITY will be useful in providing a thorough perspective on the relative benefits of different pharmaceutical regimens in patients with acute coronary syndromes.

Treatment imperatives for patients with moderate and high-risk acute coronary syndromes include early angiography followed by revascularisation when appropriate, which is most commonly accomplished with stent-based percutaneous coronary intervention. The optimum pharmacological regimen to support coronary stenting in acute coronary syndromes will provide optimum protection from periprocedural ischaemic complications as well as adverse haemorrhagic events. Administration of upstream aspirin, clopidogrel, and bivalirudin (using glycoprotein IIb/IIIa inhibitors only for refractory procedural thrombotic complications) will provide the highest likelihood for event-free survival in patients with moderate and high-risk acute coronary syndromes in whom percutaneous coronary interventions are done.

**Figure 6: Meta-analysis of contemporary trials**

(A) Composite ischaemia. (B) TIMI bleeding (major or minor). (C) Blood transfusion. Includes trials in which patients undergoing percutaneous coronary intervention were randomised to heparin plus the routine use of glycoprotein IIb/IIIa inhibitors or bivalirudin monotherapy plus the provisional use of glycoprotein IIb/IIIa inhibitors. Event rates are at 30 days for the ACUITY percutaneous coronary intervention subgroup and REPLACE-2, and death or myocardial infarction from PROTECT-TIMI-30. TIMI bleeding = TIMI major or TIMI minor bleeding. **B** Blood transfusions—one or more unit of blood product transfusion not related to coronary artery bypass graft surgery.

<table>
<thead>
<tr>
<th>Articles</th>
<th>Bivalirudin alone</th>
<th>Unfractionated heparin or enoxaparin plus glycoprotein IIb/IIIa inhibitors</th>
<th>RR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td>ACUITY-PCI</td>
<td>230/2619 (8.8%)</td>
<td>230/2561 (8.2%)</td>
<td>1.07 (0.90-1.28)</td>
</tr>
<tr>
<td>REPLACE-2</td>
<td>227/2975 (7.6%)</td>
<td>211/2990 (7.0%)</td>
<td>1.08 (0.90-1.30)</td>
<td>0.56 (0.46-0.71)</td>
</tr>
<tr>
<td>PROTECT-TIMI-30</td>
<td>25/284 (8.8%)</td>
<td>38/273 (6.6%)</td>
<td>1.33 (0.82-2.15)</td>
<td>0.19 (0.37-0.68)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>482/5878 (8.2%)</td>
<td>459/6124 (7.5%)</td>
<td>1.09 (0.97-1.24)</td>
<td>0.55 (0.46-0.66)</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>ACUITY-PCI</td>
<td>118/2619 (4.5%)</td>
<td>200/2561 (7.8%)</td>
<td>0.58 (0.46-0.72)</td>
</tr>
<tr>
<td>REPLACE-2</td>
<td>58/2993 (1.9%)</td>
<td>117/3008 (3.9%)</td>
<td>0.50 (0.37-0.68)</td>
<td>0.11 (0.02-0.84)</td>
</tr>
<tr>
<td>PROTECT-TIMI-30</td>
<td>1/284 (0.4%)</td>
<td>18/555 (3.0%)</td>
<td>0.05 (0.01-0.19)</td>
<td>0.55 (0.46-0.66)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>177/5896 (3.0%)</td>
<td>335/6424 (5.2%)</td>
<td>0.55 (0.46-0.66)</td>
<td>0.55 (0.46-0.66)</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>ACUITY-PCI</td>
<td>44/2619 (1.7%)</td>
<td>72/2561 (2.8%)</td>
<td>0.56 (0.39-0.81)</td>
</tr>
<tr>
<td>REPLACE-2</td>
<td>50/2993 (1.7%)</td>
<td>76/3008 (2.5%)</td>
<td>0.66 (0.46-0.94)</td>
<td>0.66 (0.46-0.94)</td>
</tr>
<tr>
<td>PROTECT-TIMI-30</td>
<td>1/284 (0.4%)</td>
<td>25/573 (4.4%)</td>
<td>0.08 (0.05-0.59)</td>
<td>0.08 (0.05-0.59)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>95/5896 (1.6%)</td>
<td>178/6424 (2.8%)</td>
<td>0.56 (0.44-0.71)</td>
<td>0.56 (0.44-0.71)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bivalirudin better</th>
<th>Heparin plus glycoprotein IIb/IIIa inhibitor better</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**Confl ict of interest statement**

G W Stone has received consulting fees from the Medicines Company, Boston Scientifi c, Guidant, Abbott, Volcano, St Jude, and BMS Imaging and lecture fees from the Medicines Company, Nycomed, Sanofi-Aventis, Guidant, Medtronic, and Abbott. H D White has received consulting fees and lecture fees from the Medicines Company and Sanofi-Aventis and grant support from the Medicines Company, Sanofi-Aventis, Procter & Gamble, Schering-Plough, Eli Lilly, Alexion, Merck, Neunen Pharmaceuticals, GlaxoSmithKline, Pfi zer, Roche, Fournier Laboratories, and Johnson & Johnson. E M Ohman has received consulting fees from Invoise, Savaroc, Liposiscience, Response Biomedical, the Medicines Company, Datascop, and AbioMed, has equity interest in Medtronic, and has grant support Bristol-Myers Squibb, Sanofi-Aventis, Schering-Plough, Millennium Pharmaceuticals, Eli Lilly, and Berlex. M E Bertrand has received consulting fees and lecture fees from Nycomed, Servier, and Sanofi-Aventis. A M Lincoff has received grant support from the Medicines Company, Eli Lilly, Sanofi-Aventis, Bristol Myers Squibb, and Centocor. D A Cox has received consulting fees and lecture fees from the Medicines Company, Boston Scientifi c, Guidant, St Jude, and Cordis. S J Pocock has received consulting fees from the Medicines Company. J H Ware has received consulting fees from the Medicines Company, BMS, InfraRedx, and Schering Plough. F Feit has equity interests in the Medicines Company, Johnson & Johnson, and Millennium Pharmaceuticals, and has received consulting fees from the Medicines Company. S V Manoukian has received lecture fees from the Medicines Company, Cordis, and Boston Scientifi c. J W Moses has received consulting fees from Johnson & Johnson and is on the speaker’s bureau for Astra Zeneca. B T McLaurin, A Colombo, and A J Lansky declare that they have no confl ict of interest.

**Acknowledgments**

The trial was funded by the Medicines Company (Parsippany, NJ, USA) and Nycomed (Rokskilde, Denmark).

**References**


Cardiopulmonary resuscitation by bystanders with chest compression only (SOS-KANTO): an observational study

SOS-KANTO study group

Summary

Background
Mouth-to-mouth ventilation is a barrier to bystanders doing cardiopulmonary resuscitation (CPR), but few clinical studies have investigated the efficacy of bystander resuscitation by chest compressions without mouth-to-mouth ventilation (cardiac-only resuscitation).

Methods
We did a prospective, multicentre, observational study of patients who had out-of-hospital cardiac arrest. On arrival at the scene, paramedics assessed the technique of bystander resuscitation. The primary endpoint was favourable neurological outcome 30 days after cardiac arrest.

Findings
4068 adult patients who had out-of-hospital cardiac arrest witnessed by bystanders were included; 439 (11%) received cardiac-only resuscitation from bystanders, 712 (18%) conventional CPR, and 2917 (72%) received no bystander resuscitation. Any resuscitation attempt was associated with a higher proportion having favourable neurological outcomes than no resuscitation (5·0% vs 2·2%, p<0·0001). Cardiac-only resuscitation resulted in a higher proportion of patients with favourable neurological outcomes than conventional CPR in patients with apnoea (6·2% vs 3·1%; p=0·0195), with shockable rhythm (19·4% vs 11·2%, p=0·041), and with resuscitation that started within 4 min of arrest (10·1% vs 5·1%, p=0·0221). However, there was no evidence for any benefit from the addition of mouth-to-mouth ventilation in any subgroup. The adjusted odds ratio for a favourable neurological outcome after cardiac-only resuscitation was 2·2 (95% CI 1·2–4·2) in patients who received any resuscitation from bystanders.

Interpretation
Cardiac-only resuscitation by bystanders is the preferable approach to resuscitation for adult patients with witnessed out-of-hospital cardiac arrest, especially those with apnoea, shockable rhythm, or short periods of untreated arrest.

Introduction
Cardiopulmonary resuscitation (CPR) consisting of chest compression plus mouth-to-mouth ventilation done by bystanders is a major element in the so-called chain of survival for people with cardiac arrest.5,6 Although bystander CPR improves likelihood of survival,7 reports8-14 have shown that bystander CPR was attempted for less than a-third of patients who had collapsed. Surveys have identified reluctance of bystanders to undertake mouth-to-mouth ventilation as a substantial barrier to CPR attempts.15-17 This reluctance is partly caused by fear of transmission of infectious diseases. Despite the remote chance of such infection, fears about disease transmission are common in the present era of universal precautions.18-20 Another barrier to bystanders attempting CPR is the complexity of the technique as presently taught.21-23 In CPR guidelines, cardiac-only resuscitation by bystanders is recommended in dispatcher-assisted resuscitation or if a rescuer is unwilling or unable to do mouth-to-mouth ventilation.24-26 However, this technique is not generally known, recommended, or taught to the public. Since few clinical studies have focused on the efficacy of cardiac-only resuscitation,25,27-30 we sought to compare the outcomes for patients who underwent cardiac-only resuscitation or conventional CPR by bystanders. If cardiac-only resuscitation is as effective as conventional CPR by bystanders, rescuers might be more willing and able to provide this simple intervention than they are at present. Furthermore, clinical studies have established that interruptions to chest compressions during out-of-hospital cardiac arrest are common, even by trained emergency medical staff.24-26 Studies suggest...
that such interruptions can have lethal consequences.\textsuperscript{15–17} Kern and colleagues\textsuperscript{26} showed that cardiac-only resuscitation results in substantially better survival without neurological impairment 24 h after cardiac arrest than conventional CPR.

We therefore assessed the effect of cardiac-only resuscitation by bystanders to assess the characteristics of bystander resuscitation and asked additional questions about the technique during bystander resuscitation.

Methods

Participants

A survey of survivors of out-of-hospital cardiac arrest in the Kanto region of Japan (SOS-KANTO) was done by the Association for Acute Medicine of Kanto and included 58 emergency hospitals and emergency medical service units. Between Sept 1, 2002, and Dec 31, 2003, patients who had out-of-hospital cardiac arrest witnessed by bystanders and who were subsequently transported by paramedics to emergency hospitals participating in SOS-KANTO were included in the study. Exclusion criteria were: patients younger than 18 years of age; further cardiac arrest after the arrival of paramedics; documented terminal illness; presence of a do-not-resuscitate order; and bystander resuscitation without documented chest compressions.

Procedures

The study was a prospective, multicentre, observational trial that followed Utstein style reporting guidelines.\textsuperscript{3} The study was approved by the SOS-KANTO Research Ethics Board, and the requirement for informed consent was waived according to Japanese government guidelines.\textsuperscript{26} Paramedics observed the technique of bystander resuscitation and asked additional questions of bystanders to assess the characteristics of resuscitation.\textsuperscript{12} The technique during bystander resuscitation was identified as cardiac-only resuscitation, conventional CPR, pulmonary-only resuscitation, unidentified resuscitation technique (including change of technique), or chest compression not documented. Chest compression rate and depth were not assessed. The person attempting bystander resuscitation was classified as a lay person with basic CPR training, a lay person assisted by a dispatcher, a lay person without either training or dispatcher assistance, or an off-duty health worker. All event times were measured by the dispatch centre clock, and times of collapse and first bystander resuscitation attempts were obtained from the bystanders. Cardiac arrest was defined as the cessation of cardiac mechanical activity, manifesting as unresponsiveness, apnoea (or gasping breathing), and absence of pulse. The arrest was presumed to be from a cardiac cause, unless it was known to have been caused by a non-cardiac cause including trauma, drowning, and asphyxia.\textsuperscript{1}

Resuscitation attempts were documented by both paramedics and attending physicians according to Utstein style reporting guidelines.\textsuperscript{1} Data for individual patients were entered into a database by SOS-KANTO members at each hospital and were independently cross-checked twice by different investigators. Original data were made available to the data and safety monitoring committee for independent scrutiny.

Study endpoints

The primary endpoint was favourable neurological outcome 30 days after cardiac arrest, defined as a Glasgow-Pittsburgh cerebral-performance category of 1–3.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Any bystander resuscitation</th>
<th>No bystander resuscitation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>68 (55–80)</td>
<td>68 (57–78)</td>
<td>0.8450</td>
<td></td>
</tr>
<tr>
<td>Male sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>778/1151 (68%)</td>
<td>2022/2917 (69%)</td>
<td>0.2848</td>
<td></td>
</tr>
<tr>
<td>Cardiac cause</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>806/1151 (70%)</td>
<td>1836/2917 (63%)</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Location of cardiac arrest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home or other residence</td>
<td>598/1151 (53%)</td>
<td>1730/2879 (62%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Public, indoors</td>
<td>413/1151 (36%)</td>
<td>623/2879 (22%)</td>
<td></td>
</tr>
<tr>
<td>Public, outdoors</td>
<td>127/1151 (11%)</td>
<td>466/2879 (16%)</td>
<td></td>
</tr>
<tr>
<td>First physical findings at arrival of EMS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gapping breathing</td>
<td>123/1151 (11%)</td>
<td>166/2917 (6%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Pupilometer (mm\textsuperscript{*})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.0 (5.0–6.0)</td>
<td>6.0 (5.0–6.0)</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Initial cardiac rhythm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VF/pulseless VT</td>
<td>329/1151 (29%)</td>
<td>549/2917 (19%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>PEA</td>
<td>239/1151 (21%)</td>
<td>755/2917 (26%)</td>
<td></td>
</tr>
<tr>
<td>Asystole</td>
<td>594/1151 (51%)</td>
<td>1613/2917 (55%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Any bystander resuscitation</th>
<th>No bystander resuscitation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defibrillatory shock</td>
<td>440/1151 (38%)</td>
<td>839/2917 (29%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Tracheal intubation</td>
<td>1119/1151 (97%)</td>
<td>2804/2917 (96%)</td>
<td>0.0901</td>
</tr>
<tr>
<td>Epinephrine</td>
<td>1047/1151 (91%)</td>
<td>2681/2917 (92%)</td>
<td>0.366</td>
</tr>
<tr>
<td>Hypothermia</td>
<td>28/1151 (2%)</td>
<td>48/2917 (2%)</td>
<td>0.0949</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Medical history</th>
<th>Any bystander resuscitation</th>
<th>No bystander resuscitation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronary heart disease</td>
<td>122/1118 (11%)</td>
<td>287/2878 (10%)</td>
<td>0.4570</td>
</tr>
<tr>
<td>Hypertension</td>
<td>178/1118 (16%)</td>
<td>424/2878 (15%)</td>
<td>0.4392</td>
</tr>
<tr>
<td>Heart failure</td>
<td>35/1118 (3%)</td>
<td>106/2878 (4%)</td>
<td>0.3588</td>
</tr>
<tr>
<td>Diabetes</td>
<td>139/1118 (12%)</td>
<td>332/2878 (12%)</td>
<td>0.5207</td>
</tr>
</tbody>
</table>

**Table 1: Baseline characteristics of the patients**

Data are median (IQR) or numerator/total number (%). Calculations based on available data. AED=automated external defibrillator. CPR=cardiopulmonary resuscitation. EMS=emergency medical services. PEA=pulseless electrical activity. VF=ventricular fibrillation. VT=ventricular tachycardia.\textsuperscript{1} The pupil was measured by EMS workers using one patient’s checklist with pupillometer rulers, and was recorded for 1080 patients in the any resuscitation group and for 2764 in the no bystander resuscitation group. \textsuperscript{1}Time recorded for 1073 in the any resuscitation group and 2716 in the no bystander resuscitation group. \textsuperscript{1}Time recorded for 1052 in the any resuscitation group and 2861 in the no bystander resuscitation group. \textsuperscript{1}Time recorded for 1073 in the any resuscitation group and 2764 in the no bystander resuscitation group.
Table 2: Baseline characteristics of the patients receiving bystander CPR

<table>
<thead>
<tr>
<th>Cardiac-only resuscitation</th>
<th>Conventional CPR</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>69 (55–80)</td>
<td>68 (55–80)</td>
</tr>
<tr>
<td>Male sex</td>
<td>116/439 (72%)</td>
<td>461/712 (53%)</td>
</tr>
<tr>
<td>Cardiac cause</td>
<td>305/439 (69%)</td>
<td>501/712 (70%)</td>
</tr>
<tr>
<td>Location of cardiac arrest</td>
<td>253/432 (59%)</td>
<td>345/706 (49%)</td>
</tr>
<tr>
<td>Home or other residence</td>
<td>126/432 (29%)</td>
<td>277/706 (40%)</td>
</tr>
<tr>
<td>Public, indoors</td>
<td>43/432 (10%)</td>
<td>84/706 (12%)</td>
</tr>
<tr>
<td>Public, outdoors</td>
<td>133/439 (30%)</td>
<td>101/712 (14%)</td>
</tr>
<tr>
<td>First physical findings at arrival of EMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaspng breathing</td>
<td>50/439 (11%)</td>
<td>73/712 (10%)</td>
</tr>
<tr>
<td>Pupilometer (mm*)</td>
<td>5.0 (5.0–6.0)</td>
<td>5.0 (5.0–6.0)</td>
</tr>
<tr>
<td>Initial cardiac rhythm</td>
<td></td>
<td>0.9116</td>
</tr>
<tr>
<td>VF/pulseless VT</td>
<td>124/439 (28%)</td>
<td>205/712 (29%)</td>
</tr>
<tr>
<td>PEA</td>
<td>94/439 (21%)</td>
<td>145/712 (20%)</td>
</tr>
<tr>
<td>Asystole</td>
<td>221/439 (50%)</td>
<td>362/712 (51%)</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defibrillatory shock</td>
<td>373/439 (33%)</td>
<td>267/712 (38%)</td>
</tr>
<tr>
<td>Tracheal intubation</td>
<td>474/439 (97%)</td>
<td>695/712 (98%)</td>
</tr>
<tr>
<td>Epinephrine</td>
<td>397/439 (90%)</td>
<td>650/712 (91%)</td>
</tr>
<tr>
<td>Hypothermia</td>
<td>40/439 (2%)</td>
<td>18/712 (3%)</td>
</tr>
<tr>
<td>Medical history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td>52/429 (12%)</td>
<td>70/689 (10%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>60/429 (14%)</td>
<td>118/689 (17%)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>12/429 (3%)</td>
<td>23/689 (3%)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>52/429 (12%)</td>
<td>81/689 (12%)</td>
</tr>
<tr>
<td>Performer of bystander resuscitation</td>
<td>100/439 (22%)</td>
<td>300/712 (49%)</td>
</tr>
<tr>
<td>Lay person with CPR training</td>
<td>70/439 (16%)</td>
<td>128/712 (18%)</td>
</tr>
<tr>
<td>Lay person with dispatcher-assisted resuscitation</td>
<td>139/439 (32%)</td>
<td>113/712 (19%)</td>
</tr>
<tr>
<td>Lay person with no training or assisted resuscitation</td>
<td>133/439 (30%)</td>
<td>101/712 (14%)</td>
</tr>
<tr>
<td>Time (min)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>From collapse to first bystander resuscitation attempt</td>
<td>4.0 (2.0–5.0)</td>
<td>4.0 (2.0–5.0)</td>
</tr>
<tr>
<td>From first bystander resuscitation attempt to first AED analysis</td>
<td>9.0 (8.0–11.0)</td>
<td>9.0 (7.0–11.0)</td>
</tr>
<tr>
<td>From first AED analysis to departure from scene</td>
<td>14.0 (10.0–18.0)</td>
<td>14.0 (11.0–18.0)</td>
</tr>
<tr>
<td>From departure to arrival at hospital</td>
<td>10.0 (6.0–13.0)</td>
<td>10.0 (6.0–13.0)</td>
</tr>
</tbody>
</table>

Data are median (IQR) or numerator/total number (%). Calculations based on available data. AED=automated external defibrillator. CPR=cardiopulmonary resuscitation. EMS=emergency medical services. PEA=pulseless electrical activity. VF=ventricular fibrillation. VT=ventricular tachycardia.* The pupil was measured by EMS workers using one patient’s checklist card with pupillometer rulers, and was recorded for 421 patients in the cardiac-only resuscitation group, and for 659 in the conventional CPR group. †The time was recorded for 414 in the cardiac-only resuscitation group, and for 527 in the conventional CPR group. §The time was recorded for 436 in the conventional CPR group and for 689 in the cardiac-only resuscitation group. ¶The time was recorded for 323 patients in the cardiac-only resuscitation group, and for 1120 patients for the no bystander resuscitation group on the basis of a two-sided alpha value of 0·05 and a beta error of 0·10. Baseline characteristics were compared by use of the chi² test for categorical variables and the Mann-Whitney U test for continuous variables, as appropriate. Odds ratios and their 95% CI were calculated for the study endpoints. A multiple logistic-regression analysis was done for independent predictors of resuscitation, including age, cause of cardiac arrest, technique of bystander resuscitation, resuscitation-related time intervals, and initial recorded cardiac rhythm. The non-linear regression analysis with logarithm was used to illustrate the relation between favourable neurological outcome 30 days after cardiac arrest and the time between first bystander resuscitation attempts and first use of automated external defibrillator (AED) in patients receiving bystander CPR with ventricular fibrillation or ventricular tachycardia as an initial cardiac rhythm.

Role of the funding source
The sponsors had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.

Results
9592 patients received advanced life support by paramedics and were transported to emergency hospitals during the study. Of those, 3464 patients were not eligible. The SOS-KANTO study therefore included 4068 adult patients who had bystander-witnessed cardiac arrest out of hospital; 1151 (28%) received bystander resuscitation, including 439 (11%) who received cardiac-only resuscitation and 712 (18%) who received conventional CPR, and 2917 (72%) did not receive any bystander resuscitation. No patient was lost to follow-up for study endpoints at 30 days after cardiac arrest (figure 1). At baseline, significant differences (table 1) were seen between the any resuscitation group and the no bystander resuscitation group for cause of cardiac arrest, location where the cardiac arrest happened, proportion with gasping breathing and pupillometer reading at arrival of emergency medical services, initial recorded cardiac rhythm, and the need for defibrillation by emergency medical services in the cardiac-only resuscitation group and the conventional CPR group.

Statistical analysis
An estimate of the number of patients needed to test our hypothesis was derived from analyses of three previous large-scale studies in Japan.12–14 The calculation was based on a two-fold improvement in favourable neurological outcomes from a baseline outcome of 1·6, and a 1:1:6 ratio of patients who had bystander-witnessed cardiac arrest and received cardiac-only resuscitation versus conventional CPR versus no resuscitation. The minimum sample size for comparison of favourable neurological outcome at 30 days was estimated to be 400 patients for each bystander resuscitation group, and 1120 patients for the no bystander resuscitation group. The sample size was calculated for the study endpoints. A multiple logistic-regression analysis was done for independent predictors of resuscitation, including age, cause of cardiac arrest, technique of bystander resuscitation, resuscitation-related time intervals, and initial recorded cardiac rhythm.
medical workers. Generally, the two groups that received bystander resuscitation had similar baseline characteristics (table 2). However, higher proportions in the cardiac-only resuscitation group than in the conventional group were male, more patients had cardiac arrest at home, and more were treated by lay people. The group who had any resuscitation attempt had significantly higher frequencies of favourable neurological outcome at 30 days than the no bystander resuscitation group in the whole cohort (5% [57/1151] vs 2% [63/2917]; p<0·0001) and in the subgroups of patients with cardiac causes (7% [53/751] vs 3% [54/1836]; p<0·0001), with apnoea (4% [44/1028] vs 2% [55/2751]; p<0·0001), with ventricular fibrillation or tachycardia as initial cardiac rhythm (14% [47/329] vs 8% [45/549]; p=0·0044), and with both of the times between call to emergency medical services and first AED analysis (for interval ≤10 min; 6% [44/689] vs 3% [13/462]; p=0·0069) and (for interval >10 min; 3% [13/462] vs 1% [11/1116]; p=0·0069). The cardiac-only resuscitation group also had significantly higher frequencies of favourable neurological outcome at 30 days than the no bystander resuscitation group in those categories. Although the frequency of favourable neurological outcome at 30 days did not differ between the cardiac-only resuscitation group and the conventional CPR group for the whole cohort (p=0·1459, table 3), cardiac-only resuscitation resulted in a higher proportion of patients with favourable neurological outcomes than conventional CPR in the subgroups of patients with apnoea (p=0·0195), with ventricular fibrillation or tachycardia as initial cardiac rhythm (p=0·041), and with

<table>
<thead>
<tr>
<th>Findings at arrival of EMS</th>
<th>Cardiac-only resuscitation</th>
<th>Conventional CPR</th>
<th>No bystander resuscitation</th>
<th>Any resuscitation versus no bystander resuscitation</th>
<th>Cardiac-only resuscitation versus no bystander resuscitation</th>
<th>Cardiac-only resuscitation versus conventional CPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cause</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac</td>
<td>26/305 (9%)</td>
<td>27/504 (5%)</td>
<td>54/1836 (3%)</td>
<td>2·3 (1·6–3·4)</td>
<td>3·1 (1·9–5·0)</td>
<td>1·6 (0·9–2·9)</td>
</tr>
<tr>
<td>Non-cardiac</td>
<td>1/134 (0·7%)</td>
<td>3/211 (1%)</td>
<td>9/1081 (0·8%)</td>
<td>1·4 (0·4–4·6)</td>
<td>0·9 (0·1–7·1)</td>
<td>0·5 (0·1–5·1)</td>
</tr>
<tr>
<td>Apnoea (no gasping)</td>
<td>24/389 (6%)</td>
<td>20/639 (3%)</td>
<td>55/2751 (2%)</td>
<td>2·2 (1·5–3·3)</td>
<td>3·2 (2·0–5·3)</td>
<td>2·0 (1·1–3·7)</td>
</tr>
<tr>
<td>Gasping breath</td>
<td>3/50 (6%)</td>
<td>10/73 (14%)</td>
<td>8/166 (5%)</td>
<td>2·3 (0·9–5·8)</td>
<td>1·3 (0·3–4·9)</td>
<td>0·4 (0·1–1·5)</td>
</tr>
<tr>
<td>VF/pulseless VT</td>
<td>24/124 (19%)</td>
<td>23/205 (11%)</td>
<td>45/549 (8%)</td>
<td>1·9 (1·2–2·9)</td>
<td>2·7 (1·6–4·6)</td>
<td>1·9 (1·0–3·5)</td>
</tr>
<tr>
<td>PEA</td>
<td>1/94 (1%)</td>
<td>3/145 (2%)</td>
<td>8/155 (1%)</td>
<td>1·6 (0·5–5·3)</td>
<td>1·0 (0·1–8·2)</td>
<td>0·5 (0·1–5·0)</td>
</tr>
<tr>
<td>Asystole</td>
<td>2/221 (0·9%)</td>
<td>4/362 (1%)</td>
<td>10/1613 (0·6%)</td>
<td>1·7 (0·60–4·6)</td>
<td>1·5 (0·3–6·7)</td>
<td>0·8 (0·1–4·5)</td>
</tr>
<tr>
<td>Time from EMS call to first AED analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10 min</td>
<td>22/286 (8%)</td>
<td>22/403 (5%)</td>
<td>52/1801 (3%)</td>
<td>2·3 (1·5–3·5)</td>
<td>2·8 (1·7–4·7)</td>
<td>1·4 (0·8–2·7)</td>
</tr>
<tr>
<td>&gt;10 min</td>
<td>5/153 (3%)</td>
<td>8/309 (3%)</td>
<td>11/1116 (1%)</td>
<td>2·9 (1·3–6·5)</td>
<td>3·4 (1·1–9·9)</td>
<td>1·3 (0·4–4·0)</td>
</tr>
<tr>
<td>Time from collapse to first bystander resuscitation attempt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤4 min</td>
<td>23/227 (10%)</td>
<td>18/351 (5%)</td>
<td></td>
<td>2·1 (1·1–4·0)</td>
<td>1·6 (0·8–3·3)</td>
<td>2·6 (0·1–10·0)</td>
</tr>
<tr>
<td>&gt;4 min</td>
<td>2/96 (2%)</td>
<td>4/176 (2%)</td>
<td></td>
<td>0·9 (0·2–5·1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performer of CPR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lay person</td>
<td>23/342 (7%)</td>
<td>19/362 (5%)</td>
<td></td>
<td>1·3 (0·7–2·4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Off-duty medical worker</td>
<td>4/97 (4%)</td>
<td>11/350 (3%)</td>
<td></td>
<td>1·3 (0·4–4·3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are numerator/total number or odds ratio (95% CI). AED=automated external defibrillator. CPR=cardiopulmonary resuscitation. EMS=emergency medical services. PEA=pulseless electrical activity. VF=ventricular fibrillation. VT=ventricular tachycardia.

Table 3: Patients with a favourable neurological outcome at 30 days after cardiac arrest

![Figure 2: Adjusted odds ratios for primary endpoint associated with selected factors in patients receiving bystander resuscitation whose bystander resuscitation-related time intervals were known](image-url)
resuscitation starting within 4 min of collapse (p=0·0221). However, the bystander resuscitation groups had similar frequencies of favourable neurological outcome at 30 days in the subgroups of patients who received bystander resuscitation from a lay person or from a medically trained person. Additionally, there was no evidence for any benefit from the addition of mouth-to-mouth ventilation in any subgroup of patients who received bystander resuscitation.

Multiple logistic-regression analysis (figure 2) showed that cardiac-only resuscitation resulted in higher proportions of favourable neurological outcome than conventional CPR (p=0·0144), and other independent predictors of favourable outcome were age, time between first bystander resuscitation attempt and first AED analysis, and ventricular fibrillation or tachycardia as initial cardiac rhythm. When sex and gasping breathing were included in the analysis, the results did not change.

Figure 3 shows the relation between favourable neurological outcome at 30 days and the time between first bystander resuscitation attempt and first AED analysis (ie, the duration of bystander resuscitation) in patients with ventricular fibrillation or tachycardia. Cardiac-only resuscitation resulted in higher proportions of favourable neurological outcome than conventional CPR (22% [22/100] vs 10% [17/166], the odds ratio, 2·5; 95% CI, 1·2–4·9; p=0·0086), and likelihood of a favourable neurological outcome decreased in both bystander resuscitation groups for every 1 min increment in time from first resuscitation attempt to first AED analysis (p=0·0105 for the cardiac-only resuscitation and p=0·0003 for conventional CPR). When the time from collapse to first AED analysis was used instead, the results were unchanged.

The proportions of people who were alive at 30 days differed from those who had favourable neurological outcome at 30 days (figure 4). In all subgroups of patients, the proportions surviving at 30 days showed no differences between the two bystander resuscitation groups, and these two groups also had similar frequencies of survival until hospital admission.

Discussion
This report shows that bystander cardiac-only resuscitation is equivalent or superior to conventional bystander CPR in adult patients with witnessed out-of-hospital cardiac arrest, in terms of neurological benefit. Not only the any resuscitation group, but also the cardiac-only resuscitation group had higher proportions of favourable neurological outcome than the no bystander resuscitation group for the whole cohort, and cardiac-only resuscitation resulted in better outcome than conventional CPR in some important subgroups of patients. These subgroups included patients with apnoea (about 90% of patients in this study), and those with shockable cardiac rhythm or short periods of untreated arrest (CPR that started within 4 min of arrest). These patients had the greatest chance of survival (table 3, figures 3 and 4). After adjustment for resuscitation, cardiac-only resuscitation was an independent predictor of favourable neurological outcome in patients who received bystander resuscitation (figure 2). Furthermore, there was no evidence for any benefit from the addition of mouth-to-mouth ventilation in any subgroup of patients (table 3, figures 2, 3, and 4).

Some experts have expressed concern that the absence of assisted ventilation with chest compressions might result in lower survival and worse neurological outcomes in survivors of cardiac arrest.14 Several clinical studies have
shown that cardiac-only resuscitation is at least as effective as chest compression with mouth-to-mouth ventilation.\textsuperscript{2,3,12} In this study, however, cardiac-only resuscitation resulted in better or similar neurological outcome than conventional CPR. Moreover, the proportion of patients surviving, including those surviving until hospital admission, showed no differences between the cardiac-only resuscitation group and the conventional CPR group in any subgroup of patients. The number of patients who received bystander cardiac-only resuscitation was higher in this study (n=439) than in previous ones (n=273,\textsuperscript{2} 241,\textsuperscript{2} 41)\textsuperscript{1,2,13,14} and that gasping breathing is associated with a better outcome.\textsuperscript{1,2,11,13} In this study, most patients who had bystander resuscitation had an untreated arrest interval of less than 6 min and a duration of bystander resuscitation of less than 12 min, and the two bystander resuscitation groups had similar time intervals and had similar proportions of patients with gasping breath.

Another reason for the efficacy of cardiac-only resuscitation could be that mouth-to-mouth ventilation has several potential disadvantages. These disadvantages include gastric insufflations and importantly, less cycle time spent on effective compressions.\textsuperscript{4,5,15,25,30,31} Time spent on mouth-to-mouth ventilation takes precious time away from chest compressions that support cerebral and coronary perfusion.\textsuperscript{3,4,15,25,30,31} Intrathoracic pressure drops after each pause for mouth-to-mouth ventilation, and several chest compressions have to be done before previous rates of cerebral and coronary perfusion are re-established.\textsuperscript{1,2,6} In this study, the quality of chest compressions might not have been as good in the cardiac-only resuscitation group as in the conventional CPR group, because the proportion of patients treated by people with no first-aid training, with or without dispatcher-assistance, was higher in the cardiac-only group (272/439, 62% vs 234/712, 33%). Also, the proportion treated by medically trained individuals was lower in the cardiac-only group than in the conventional CPR group (97/439, 22% vs 350/712, 49%). However, there would be less interruption of chest compressions in the cardiac-only resuscitation group. We suggest that interruption of chest compressions was the main reason why conventional CPR did not result in better neurological outcome than cardiac-only resuscitation.

There are several limitations to our study. It was neither a randomised controlled trial nor a population-based study. However, the overall survival and patients’ characteristics in this study were similar to those of population-based studies from similar large metropolitan areas—Osaka, Japan,\textsuperscript{13} and New York, USA.\textsuperscript{14} Although the total number of patients was large, there were few patients with arrest caused by asphyxia, drowning, or traumatic brain injury. The quality of bystander resuscitation was not assessed, and resuscitation-related event times were known for only 70% of the study population. Additionally, post-resuscitation care could not be standardised. Recent studies\textsuperscript{15,16,17} have shown that therapeutic hypothermia can result in better outcomes for patients with out-of-hospital ventricular fibrillation. In this study, few patients were treated by induction of hypothermia, and the proportions given this treatment were similar in the two bystander resuscitation groups.

On the basis of these findings, we conclude that bystander cardiac-only resuscitation is the preferred approach to resuscitation for adult patients with witnessed out-of-hospital cardiac arrest especially those with apnoea, shockable cardiac rhythm, or short periods of untreated arrest.

\textbf{SOS-KANTO study group}

Nihon University Surugadai Hospital, Tokyo, Japan (Ken Nagao MD, Kimio Kikushima MD); Teikyo University Hospital, Tokyo, Japan (Tetsuya Sakamoto MD); Kawaguchi Municipal Medical Center, Saitama, Japan (Kazuhide Koseki MD); Toho University Oomori Hospital, Tokyo, Japan (Masaki Igarashi MD); St Luke’s International Hospital, Tokyo, Japan (Shinichi Ishimatsu MD); Saitama Medical School Hospital, Saitama, Japan (Akira Sato MD); Keio University Hospital, Tokyo, Japan (Shingo Hori MD); Showa University Fugisaska Hospital, Tokyo, Japan (Shigeru Kanesaka MD); Tokyo Metropolitan Bokutoh Hospital, Tokyo, Japan (Yuichi Hamabe MD); National Defense Medical College, Saitama, Japan (Daizo Saito MD); and Kunito Chuo Hospital, Chiba, Japan (Shinya Kikutama MD).

\textbf{Contributors}

K Nagao and T Sakamoto as principal investigators, participated in idea formation, study design and completion, data collection, data management, data analyses, interpretation of results and revision of the report, and contributed to the final report. K Nagao obtained funding. K Kikushima, K Koseki, M Igarashi, S Ishimatsu, A Sato, S Hori, S Kanesaka, Y Hamabe, D Saito, and S Kikutama participated in study idea formation, study design and completion, data collection, data management, and interpretation of the results. K Nagao, T Sakamoto, K Kikushima, and D Saito did the statistical analysis. All authors approved the final version.

Conflict of interest statement
We declare that we have no conflict of interest.

Acknowledgments
We thank all the bystanders who undertook basic resuscitation and the paramedics, emergency medical technicians, nurses, and physicians who participated in the SOS-KANTO study. This study was supported by a grant from the Laerdal Foundation of Acute Medicine, Norway and a research grant for Cardiovascular Disease (14c-7) from the Ministry from Health, Labour and Welfare, Japan.

References
34. Sandeen AB, Ewy GA. Cardiopulmonary resuscitation in the real world: when will the guidelines get the message? JAMA 2003; 293: 363–65.
Effect of zinc supplementation on mortality in children aged 1–48 months: a community-based randomised placebo-controlled trial

Sunil Sazawal, Robert E Black, Mahdi Ramsan, Hababu M Chwaya, Arup Dutta, Usha Dhingra, Rebecca J Stoltzfus, Mashavi K Othman, Fatma M Kabole

Summary

Background Studies from Asia have suggested that zinc supplementation can reduce morbidity and mortality in children, but evidence from malarious populations in Africa has been inconsistent. Our aim was to assess the effects of zinc supplementation on overall mortality in children in Pemba, Zanzibar.

Methods We enrolled 42 546 children aged 1–36 months, contributing a total of 56 507 child-years in a randomised, double-blind, placebo-controlled trial in Pemba, Zanzibar. Randomisation was by household. 21 274 children received daily supplementation with zinc 10 mg (5 mg in children younger than 12 months) for mean 484·7 days (SD 306·6). 21 272 received placebo. The primary endpoint was overall mortality, and analysis was by intention to treat. This study is registered as an International Standard Randomised Clinical Trial, number ISRCTN59549825.

Findings Overall, there was a non-significant 7% (95% CI −6% to 19%; p=0·29) reduction in the relative risk of all-cause mortality associated with zinc supplementation.

Interpretation We believe that a meta-analysis of all studies of mortality and morbidity, will help to make evidence-based recommendations for the role of zinc supplementation in public health policy to improve mortality, morbidity, growth, and development in young children.

Introduction Pneumonia, diarrhoea, and malaria account for 45% of the 10·6 million yearly deaths of children younger than 5 years despite some success with preventive and therapeutic interventions.1 2·6 million of these deaths take place in Africa,2 including 90% of the 0·8 million worldwide children malaria deaths every year.3 Identification of low-cost interventions that would reduce such mortality, and help countries achieve the Millennium Development Goal of two-thirds reduction in child mortality,4 is a priority.

Deficiency of a few essential micronutrients is recognised to greatly increase the risk of morbidity and mortality from infectious diseases in developing countries. Evidence for zinc as one such important micronutrient has emerged; in addition to data linking zinc deficiency to growth retardation and impairment of immune function,5 zinc supplementation has shown significant reduction in rates and severity of diarrhoea6 and pneumonia7 (the two main causes of under-5 mortality). Reports from three small trials in Asian populations without malaria noted that zinc supplementation significantly reduced child mortality.8–10 Evidence for the benefit of zinc supplementation on malaria morbidity has been inconsistent.11–13 We therefore undertook a community-based trial to assess the effect of zinc supplementation on mortality in children aged 1–48 months in Pemba, Zanzibar, a place with a high-frequency of malaria transmission.

Methods

Participants The study was undertaken in Pemba, the smaller of the two islands of the Zanzibar archipelago, with a population of about 350 000, most of whom are African, and Shirazi Muslims. A baseline census of this island suggested an infant mortality rate of 89 per 1000 livebirths. Malaria is holoendemic and has year-round transmission that is highest from June to September, after the Spring rainy season (March–May). The frequency of malaria transmission is representative of coastal east Africa, where a yearly inoculation rate of 405 infective bites per person has been described.14 Plasmodium falciparum accounts for nearly all serious clinical malaria.

From January to December, 2001, we mapped and censused the whole island, and assigned unique numbers to all houses; the island was divided into working areas (clusters) and one female community worker was responsible for one cluster. Details of the study design have been previously reported.15 Parents of eligible children, those aged 1–35 months, who were likely to remain resident on the island and did not have severe malnutrition needing rehabilitation (defined as Kwashiorkor, noted by the enrolment worker), were invited to participate in the study. Enrolment was undertaken one district at a time, starting on Jan 1, 2002 and finishing the initial enrolment on June 29, 2002. All children aged 1–35 months were enrolled if parents gave consent. A study worker read the consent statement to the primary
caregiver and then signed the consent form if consent was given. The study was approved by the Johns Hopkins committee on human research, the WHO ethical review committee, and the Zanzibar Research Council.

Procedures

The study was a double-blind randomised trial with four arms: iron, folic acid, and zinc (IFAZ); iron and folic acid (IFA); zinc; and placebo. The primary endpoint was all-cause deaths. On recommendation of the Data and Safety Monitoring Board (DSMB), the IFAZ and IFA groups were stopped on Aug 19, 2003 because of overwhelming evidence of increased hospital admissions and a trend for increased mortality associated with iron supplementation; the results from these groups were subsequently reported.14

Children from the IFAZ and non-zinc IFA groups were switched to the zinc and placebo groups, respectively. The halting of the IFAZ and IFA arms and switching of children to the other groups were done with the help of WHO and the DSMB statistician. The trial initially used 16-letter codes (four-letter codes assigned to the individual supplementation groups) and two-stage blinding. The four-letter codes for every group were known only to WHO and the manufacturer; the pharmacy dispensing the supplements knew only which letter code was assigned to each child, and the study worker and family knew neither. At the time of switch, WHO and the DSMB statistician provided an alternative letter code for all of the redundant eight-letter codes. The zinc and placebo groups continued, on recommendation of the DSMB, until Sept 30, 2005 and we report results for the primary and secondary (cause-specific) mortality outcomes of these groups.

The supplement was a dispersible (dissolves in water or breast-milk in about 20 s) tablet containing 10 mg of elemental zinc sulphate, manufactured by Nutriset, Bierre, France in scientific collaboration with WHO. The zinc-containing tablets and the placebo tablets were provided in blister strips. Tablets of both groups were similar in packaging, appearance, taste, and inactive ingredients. Children aged 12 months or older were given one tablet a day; children aged younger than 12 months were given a half tablet a day. The tablet was dissolved in 5–10 mL of water or breast milk. All children aged 12 months or older were given 200 000 IU of vitamin A every 6 months; children aged 6–11 months were given 100 000 IU. Children received zinc or placebo supplements until they were 48 months of age.

Randomisation was by household, by an allocation sequence (permuted block randomisation with block length of 16) computer-generated by WHO. The blister strips were coded with one of the letter codes, and every child was assigned a letter on enrolment on the basis of randomisation sequence. Labels with the child’s identification were attached by the pharmacy to the supplement strips.

A card with the child’s and household’s identification information and a control number for verification was provided to each child. The family was asked to maintain this card, show it to the community worker during visits, and present it if the child was taken to hospital. Three grades of supervisor, all of whom were mobile on motorbikes, ensured flow of information and supplies to and from the participants. The children were visited every week at home by a community worker who delivered a strip of seven tablets labelled with the child’s

---

Figure 1: Trial profile
identification, obtained the previous week’s strip, and recorded the mother’s report of tablets consumed in the past week. These reports, and the presence of any tablets remaining in the blister pack, were used to record adherence. On such home visits, workers also obtained information about hospital admissions and deaths of study children. Reports of hospital admissions were given to study staff to compare with information from hospital surveillance. Reports of deaths were given to two supervisors who were trained to do post-mortem interviews with the family to establish the cause of death; a standard instrument was adapted for local use, pretested, and administered within 2 months of the child’s death.

All five hospitals on the island were monitored by teams of two hospital staff and two full-time supervisors employed by the project to be present in the paediatric ward for 18 hours a day. For all admissions to the paediatric ward, standard study procedures included confirmation of the identity of the child, tracking the child throughout their hospital stay, and recording cause of death, if necessary.

The primary outcome was overall mortality in participating children aged 1–48 months. Deaths were counted if they took place during the supplementation period or within 30 days of stopping supplementation, irrespective of adherence to supplement. Secondary outcomes were age-specific, sex-specific, and cause-specific mortality. On the basis of previous evidence, we analysed deaths in two age groups (children younger than 12 months and those 12 months or older), which were selected before starting the study and specified in the protocol. Cause of death was ascertained by post-mortem interview and hospital records. Three teams that consisted of two physicians and a medical assistant, independently assigned one primary and two secondary causes of death. Any disagreements were resolved by discussion. These teams did not include investigators and were masked to supplement allocation.

The sample size of the original trial (with iron supplementation) was designed to detect a reduction in overall mortality of 20% with 90% power, a 5% two-sided type I error, 10% loss to follow-up, a design effect of 1.05, and the assumption of no interaction of iron and zinc affecting mortality (baseline mortality rate 15.3/1000) in children aged 1–48 months. To meet these criteria, we needed a sample size of about 15000 person-years in each group for the marginal comparisons of the zinc arms versus the non-zinc arms. At the second DSMB meeting, 14 months into recruitment, we noted that the mortality rate was substantially lower than originally expected. We then recalculated the sample size estimates, which were about 30000 person-years in marginal comparisons for 20% reduction and 90% power. After the IFA-containing arms were stopped, marginal comparisons were not valid, so recruitment and follow-up for the effects of the zinc supplement were continued to accumulate the required

<table>
<thead>
<tr>
<th>Zinc group</th>
<th>Placebo group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total enrolment (n)</strong></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>10,681</td>
</tr>
<tr>
<td>Girls</td>
<td>10,593</td>
</tr>
<tr>
<td><strong>Age at enrolment (months)</strong></td>
<td></td>
</tr>
<tr>
<td>0–11</td>
<td>6668</td>
</tr>
<tr>
<td>≥12</td>
<td>4013</td>
</tr>
<tr>
<td><strong>Adherence</strong></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>80%</td>
</tr>
<tr>
<td><strong>Education (literate)</strong></td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>3812 (36%)</td>
</tr>
<tr>
<td>Mother</td>
<td>4556 (43%)</td>
</tr>
<tr>
<td><strong>History of sleeping under bed net</strong></td>
<td></td>
</tr>
<tr>
<td>Always</td>
<td>2907 (27%)</td>
</tr>
<tr>
<td>Sometimes</td>
<td>1211 (11%)</td>
</tr>
<tr>
<td><strong>Anthropometric status</strong></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>216 (64%)</td>
</tr>
<tr>
<td>Wasted</td>
<td>19 (6%)</td>
</tr>
<tr>
<td>Stunted</td>
<td>91 (27%)</td>
</tr>
<tr>
<td>Wasted and stunted</td>
<td>11 (3%)</td>
</tr>
<tr>
<td>Haemoglobin* (g/L)</td>
<td>96.5 (12.9)</td>
</tr>
<tr>
<td>Zinc protoporphyrin* (μmol/M of haem)</td>
<td>146.0 (103.4)</td>
</tr>
<tr>
<td>Plasma zinc concentration* (μmol/L)</td>
<td>11.99 (5.0)</td>
</tr>
<tr>
<td>Serum ferritin concentration* (μg/L)</td>
<td>55.39 (77.6)</td>
</tr>
</tbody>
</table>

Data are n (%) or mean SD. Percentages do not necessarily add up to 100% because of rounding. *Restricted to a subgroup of substudy children (n=1066). A person was designated as literate if they had had some schooling and could read and write.

Table 1: Baseline characteristics of children assigned zinc and placebo
sample size of 45,000 child years of follow-up for the original zinc and placebo arms.

A substudy of 4000 children, assessed in more detail than the others, consisted of a random sample of clusters in all four districts selected after the census. Detailed methods have been reported previously.14 Supplementation and adverse event surveillance was the same as in the main trial. Children in the substudy visited a study clinic for baseline assessment, including physical examination, height and weight measurements, and parental interview by a physician or clinician and trained health worker; a 3 mL venous blood sample was obtained from the child for comprehensive haematological analysis, erythrocyte zinc protoporphyrin measurement, and malaria parasite count.

Statistical analysis
A data entry system enabled all information gathered in the community to be entered by the end of the next day.14 The system had many range and logic checks, and any errors were corrected by field or hospital staff on a daily basis. For all outcomes, double data entry was used. We did an intention-to-treat analysis. For children who migrated away from the trial location, withdrew, or died, data were included until they left the study. Person-time analysis was done with the time the child was followed-up as denominator. For the effect on total mortality and cause-specific mortality, we used Anderson Gill time-to-event survival methods in Cox regression with robust estimation of SE to account for multiple events per household (SAS version 9.0, STATA version 8.2). In these analyses, a relative rate of less than 1 was consistent with protection in the zinc group compared with the placebo group. To assess the effect of duration of supplementation on intervention effects, Nelson-Aalen cumulative hazard estimates were calculated and graphically presented with STATA version 8.2.

We assessed the interaction of the effect of Z with age and sex by Mantel-Cox comparison between subgroups, estimating $\chi^2$ for unequal relative rates (effect modification) and its $p$ value (STMC procedure in STATA version 8.2).

We did analyses of both primary cause of death (established by verbal post-mortem assessment) and any cause of death (more than one cause could have contributed to the death). Since the results of these two analyses were very similar, only those for the primary cause of death are presented.

This study is registered as an International Standard Randomised Clinical Trial, number ISRCTN59549825.

Role of the funding source
The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Who coordinated the preparation and delivery of supplements and organisation of DSMB meetings.

<table>
<thead>
<tr>
<th></th>
<th>Zinc group</th>
<th>Placebo group</th>
<th>Relative risk</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaths</td>
<td>401</td>
<td>433</td>
<td>1.42</td>
<td>0.93</td>
<td>0.294</td>
</tr>
<tr>
<td>Rate per 100 child-years</td>
<td>1.42</td>
<td>1.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (months)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–11</td>
<td>204</td>
<td>192</td>
<td>1.06</td>
<td>0.87–1.29</td>
<td>0.566</td>
</tr>
<tr>
<td>≥12</td>
<td>197</td>
<td>241</td>
<td>0.82</td>
<td>0.68–1.00</td>
<td>0.045</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>125</td>
<td>218</td>
<td>0.81</td>
<td>0.66–0.99</td>
<td>0.04</td>
</tr>
<tr>
<td>Girls</td>
<td>226</td>
<td>215</td>
<td>1.05</td>
<td>0.87–1.26</td>
<td>0.605</td>
</tr>
<tr>
<td>Sex and age (months)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys 0–11</td>
<td>86</td>
<td>91</td>
<td>0.95</td>
<td>0.71–1.28</td>
<td>0.751</td>
</tr>
<tr>
<td>≥12</td>
<td>89</td>
<td>127</td>
<td>0.71</td>
<td>0.54–0.93</td>
<td>0.013</td>
</tr>
<tr>
<td>Girls 0–11</td>
<td>118</td>
<td>101</td>
<td>1.15</td>
<td>0.88–1.30</td>
<td>0.296</td>
</tr>
<tr>
<td>≥12</td>
<td>108</td>
<td>114</td>
<td>0.95</td>
<td>0.73–1.24</td>
<td>0.722</td>
</tr>
</tbody>
</table>

Table 2: Effect of zinc supplementation on overall mortality and by sex and age groups
Results

Of 33 617 households contacted, 91.9% consented to participate, leading to a total enrolment of 60 225 children; 42 546 children were enrolled in the two arms included in this analysis (zinc and placebo), contributing a total follow-up of 56 507 child-years (figure 1). The mean duration of supplementation was 484.7 days (SD 306.6; 10 and 90 percentile=110 and 952 days, respectively). The two groups were similar for all measured baseline characteristics (table 1), including use of bednets, transfusions in previous 6 months (107 for zinc, and 128 for placebo) and reported hospital admissions in the previous 6 months. Adherence for consumption of supplements during follow-up (proportion of follow-up days supplement was consumed) was similar between the two arms. To account for age and sex interactions of treatment effects, we also assessed the baseline features by sex (table 1) and age (data not shown), which revealed comparability at baseline by age and sex subgroups.

The mortality rate in the placebo group was 15.3 per 1000 child-years at risk. Of 834 deaths in study children, 155 (19%) took place in hospital. Overall, there was no significant 7% (95% CI −6% to 19%) reduction in the relative risk of all-cause mortality associated with zinc supplementation (table 2). However, this risk seemed to differ marginally between children younger than 12 months (−6% [−29% to 13%]) and children aged 12–48 months (18% [0% to 32%]; p=0.071). The effect on mortality also marginally differed by sex, being −5% (95% CI −6% to 19%) for girls and 19% (1% to 34%) for boys; p=0.036. We calculated Nelson-Aalen cumulative hazard estimates for effect of duration of supplementation on intervention effects overall, for age younger than 12 months and 12–48 months (figure 2). In infants, data were similar in those aged less than 6 months and those aged 6–12 months (data not presented).

Overall, there were non-significant trends for lower mortality due to malaria, other infections, and diarrhoea in the zinc group than in the placebo group (table 3). Cause-specific mortality analysis in the subgroups (table 3) suggested that, in the subgroups that had a significant change in total mortality, the effect was not restricted to a specific cause (figure 3), but contributed mainly by malaria, other infections, and to lesser extent diarrhoea. Of 84 deaths attributed to other infections, 47 were due to pneumonia and 18 to sepsis, meningitis, measles, or pertussis; 19 deaths were attributable to indeterminate febrile illness.

Discussion

In our study, zinc supplementation did not result in a significant reduction in overall mortality in children aged 1–48 months in a population with high malaria transmission. However, there was a suggestion that the effect varied by age, with no effect on mortality in infants, and a marginally significant 18% reduction of mortality in children 12–48 months of age (p=0.045). This effect was mainly a consequence of fewer deaths from malaria and other infections. Any effect on mortality in this trial was in addition to a possible effect of vitamin A supplementation that both groups received according to recommendations. That the verbal autopsy had low specificity for identification of malaria-related deaths needs to be borne in mind. However, the common error is misclassification between malaria and other febrile infections. Since both types of illness were counted, misclassification would probably not affect overall conclusions. We powered our study on the basis of earlier preliminary results for a 20% reduction in mortality, and therefore, we did not have sufficient power...
to detect a smaller, but still clinically important, say 10%, relative reduction in total mortality.

The effect of zinc supplementation on prevention of morbidity from diarrhoea and pneumonia in developing country settings has previously been shown. Many trials of daily or weekly zinc supplementation in low income settings have been reported and nearly all noted a reduced frequency of diarrhoea and pneumonia when these outcomes were assessed. Three trials of zinc from Africa with diarrhoea or pneumonia outcomes, or both, showed significant effects on diarrhoea, and one that assessed pneumonia in children with protein energy malnutrition reported a beneficial effect.

Trials on prevention of malaria morbidity have produced conflicting results. A trial in Papua New Guinea reported that zinc supplementation significantly reduced the frequency of clinic visits for malaria associated with *P. falciparum* and noted greatest effectiveness of zinc in ill children with high-density parasitaemia. An earlier study in the Gambia had also reported a non-significant reduction in malaria clinic visits in the zinc-supplemented group compared with the placebo group (*p* = 0.07). In Uganda, zinc-supplemented children had non-significantly fewer infections (82% of which were malaria) than children given placebo. In Burkina Faso, zinc supplementation had no effect on childhood malaria morbidity diagnosed at home visits.

During the past 5 years, three studies of zinc supplementation, which were not powered by design to assess mortality, did find significant effects on mortality. A study in full term, small for gestational age children in India noted a 68% relative reduction in mortality compared with control and two studies from Bangladesh, one with zinc as treatment for diarrhoea and the other with weekly zinc supplementation recorded 50% and 80% relative reductions in mortality, respectively. Two other small trials had results suggestive of mortality reduction. A study in Brazil in low birthweight children showed a 50% relative mortality reduction. In the Burkina Faso trial, although there was no effect on malaria morbidity, a 59% lower relative mortality in the zinc group was seen.

After iron, zinc is the second most abundant trace element in the body and it mediates many different physiological functions. It is a necessary component of many metallo-proteins, including those important for DNA replication and cell division, and is crucial for maintenance of immunological integrity. In studies of zinc deficiency, the production of tumour necrosis factor-α, interferon-γ and interleukin-2 by peripheral blood mononuclear cells, which are all Th1 products, are decreased, whereas products of Th2 cells (interleukin-4, interleukin-6 and interleukin-10) are unaffected. These changes, however, are easily reversed with zinc supplementation.

Because of its role in maintenance of cell integrity and immunity, zinc is thought to play an important part in the prevention of infectious diseases. The effect of zinc on diarrhoea might also be related to its role in water and electrolyte transport, intestinal permeability, enzyme functions of enterocytes, enhanced intestinal tissue repair, or enhanced local immunity restricting bacterial overgrowth and causing early pathogen clearance. More than one of these mechanisms could be implicated.

The interaction between zinc effects and age is consistent with findings of no significant effect on diarrhoea morbidity in children aged 6–11 months in the previous pooled analysis of trials. Other studies assessing the effect of zinc in children aged 0–6 months have also recorded no effect both in prevention and therapy of diarrhoea. Our study consisted of a large number of children, so insufficient power to detect an effect in infants unlikely to explain our results. There are several possible explanations for the absence of effects of zinc supplementation in children younger than 12 months. Infants might have acquired adequate zinc in utero (there is preferential zinc shunting across the placenta) and are able to obtain adequate zinc from breast milk, even when maternal stores are suboptimum. Since breastfeeding rates were 95% or greater until the infant was aged 12 months in the study population, we cannot compare the effects of zinc supplements between breastfed and non-breastfed infants. Alternatively, the absence of effect in this age group might be related to the low 5 mg dose used, compared with 10–20 mg given in previous studies, in which effects on morbidity or mortality were seen. During infancy, the cellular immune system matures, with a shift from Th2 predominant immunity at birth to predominantly Th1 immunity by age 2 years. Effects of zinc might be mediated through improvement in immunity by stimulating optimum Th1/Th2 balance and this effect could be restricted in infants because of intrinsic limitations in the capacity of infants to produce interferon and other Th1 (interleukin-2, interleukin-12) interleukins. As the maturation of the immune system is at least partly driven by the exposure of children to microorganisms, variation in response to zinc supplements in infants in different populations might be expected. Our findings of no effect in infants need further investigation with existing datasets, and in subsequent studies because they could have important implications for targeting of children who would benefit from additional zinc.

The apparent differences in the effect of zinc supplementation by sex, with a benefit in boys and no benefit in girls, could possibly be due to chance, but are consistent with previous studies recording substantial differences in effects of zinc supplementation on growth, diarrhoea morbidity and respiratory morbidity in boys. Nutritional and immunological differences might affect responses to infections and survival. Estimated zinc requirements for infant growth are higher for boys than for girls, which has been suggested as a possible reason for greater effect in boys. However, plasma zinc concentrations have not shown a large differential between boys and girls. Sex differences in effects on mortality have been recorded for vitamin A supplementation, iodine supplementation and antibiotic use.
and are also related to non-specific mortality effects of measles, hepatitis B, and DTP vaccinations. Little is known about possible mechanisms for the biological basis of these effects. Animal models of malaria, respiratory syncytial virus and influenza suggest that Th1-mediated immunity is protective against severe disease, whereas Th2-mediated immunity increases susceptibility to disease. Such animal models also suggest that female mice might have a stronger Th2 profile than male mice. Antibody-dependent cellular cytotoxicity antibody titres have been noted to be lower in young females during acute infection and after vaccination. These findings suggest that sex differences in effects of zinc supplementation could be mediated through differential effects on the immune system.

Thus, the results of this large community-based placebo-controlled zinc supplementation trial suggest that in settings with high-frequency malaria transmission typical of sub-Saharan Africa, zinc supplementation did not have any effect on mortality in infants, but there was a suggestion of reduced mortality in children older than 1 year. Feasible and sustainable methods of enhancing the bioavailable intake of dietary zinc need assessment. We also need to know whether a higher dose would have a different effect in infants, and to elucidate the mechanisms of the effects of zinc and any differences between boys and girls. Our results suggest a need for meta-analysis of all available studies both for mortality and morbidity to make evidence-based recommendation for public health policy to improve mortality, morbidity, growth, and development.

Contributors
S Sazawal and R E Black coordinated the trial, and made primary contributions to development, rationale, design, execution, analysis, and writing. R J Stollfus contributed to study design. A Dutta contributed to implementation of the trial, design of surveillance systems, quality control, and with U Dhiranga was responsible for programming, data management, and analysis. M Ramsan, H M Chwaya, Mashavi K Othman, and with U Dhingra was responsible for data collection, instruments, surveillance systems, and clinical care of patients.

Conflict of interest statement:
We declare that we have no conflict of interest.

Acknowledgments
This study was supported by a grant from WHO. Department of Child and Adolescent Health and Development with funds from United Nations Foundation, from the Family Health and Child Survival and Global Research Activity, Cooperative Agreements between Johns Hopkins Department of International Health and the United States Agency for International Development, and from the Bill & Melinda Gates Foundation through its support for micronutrient research at the Johns Hopkins Bloomberg School of Public Health. We thank the participating children, their parents, and the study team, including field workers, supervisors, physicians, data managers and other support staff for their contributions and support; the Ministry of Health and Social Welfare of the Government of Zanzibar; Tanzania for guidance and infrastructure support at every stage of the study; the Public Health Laboratory (Ivo de Carnari) as collaborating institution for contributions to management, community relations and other support. We also thank Olivier Fontaine for contributions on behalf of WHO in study design, development, supply of supplements, active monitoring, coordination of the DSMB, and interpretation of results. We would also like to thank the Data and Safety Monitoring Board members Michael Hambridge, William Blackwelder, and Zul Premji.

References
46 Moran TM, Park H, Fernandez-Sesma A, Schulman JL. Th2 responses to inactivated influenza virus can be converted to Th1 responses and facilitate recovery from heterosubtypic virus infection. J Infect Dis 1999; 180: 579–85.
Bipolar disorder (recurrent depressive and hypomanic episodes) and related disorders (united in the bipolar spectrum) are understudied, despite a prevalence of about 5% in the community and about 50% in depressed outpatients. The apparent increase in prevalence of the bipolar spectrum is related to several changes in diagnostic criteria, including improved probing for history of hypomania (focused more on overactivity than on mood change), lower minimum duration of hypomania, and inclusion of unipolar depressions with bipolar signs (eg, family history of bipolar disorder, mixed depression). Prevalence of mixed depression, a combination of depression and manic or hypomanic symptoms, is high in patients with bipolar disorders. Controlled studies are needed to investigate treatment of mixed depression; antidepressants can worsen manic and hypomanic symptoms, and mood stabilising agents might be necessary.

Bipolar disorder is defined by recurrent periods of highs and lows of mood, thinking, and activity, of varying severity, duration, and frequency, including combinations of manic or hypomanic and depressive symptoms in the same episode (the mixed states). This Seminar focuses more on bipolar II disorder than on bipolar I disorder, because the former is much understudied, and because recent advances in this area are mainly to do with bipolar II and related disorders (the bipolar spectrum).

Kraepelin’s concept of mood disorders

Kraepelin’s description of manic-depressive insanity in 1913 comprised all currently defined mood disorders, without categorical boundaries between subtypes. Manic-depressive insanity was described in terms of excitement or inhibition (increase or decrease) of three basic domains: mood, thinking, and activity. Of these three, all the domains could move in the same direction, leading to episodes of manic, hypomanic, or depressive states; or the domains could move in opposite directions, leading to the mixed states, which were defined as combinations of manic or hypomanic and depressive symptoms in the same episode. Mania could be diagnosed even if mood was decreased, if thinking and activity were increased. The main components of the concept of manic-depressive insanity were: the recurrent, episodic course, including episodes of opposite polarity (ie, recurrent mania or hypomania and depression) and of one polarity (eg, recurrent depression); mixture of opposite-polarity symptoms in the same episode (the mixed states); mood temperaments (persistent traits); family history of mood disorder; and young age at onset (table 1). Kraepelin’s no-priority approach to the basic symptoms of mood disorders is the opposite of the approach of the current most commonly used diagnostic system, the diagnostic and statistical manual of mental disorders, fourth edition, text revision (DSM-IV-TR). This manual sets mood change—elevated or irritable mood in mania/hypomania, low mood in depression—as the core feature (criterion A) of mood disorders. However, no sound evidence supports this approach. In the WHO International Classification Diseases (ICD)-10 Classification of Mental and Behavioural Disorders, overactivity is given the same priority as mood change. Recently, overactivity (increased goal-directed activity, listed in DSM-IV-TR as criterion B symptom of mania or hypomania) has been upgraded to the same level of priority as mood change in recent studies.

Kraepelin described mania as a psychotic, excited state necessitating hospital admission. Hypomania was described as having different levels of severity, and often did not require hospitalisation. Disordered thinking in hypomania included creative thinking (more goal-directed, structured planning, more brilliant ideas), crowded thoughts (non-stop thinking), racing thoughts (speedy thinking), and flight of ideas (quick, disconnected thinking). In overactivity, functioning was frequently increased, but could also be decreased (although less so than in mania), according to the degree of thinking disorder. Kraepelin, who studied patients in hospital, could not see the milder hypomanic states frequently found in the community. These milder hypomanic states, which often improve functioning, were described by Hecker in outpatient private practice as cyclothymia (similar to the DSM-IV-TR definition of bipolar 2 disorder), and by Falret as circular insanity.
Kraepelin’s description of mixed states supported his unitary concept of mood disorders: if symptoms of opposite polarity could be present in the same episode, then mania or hypomania and depression could not be distinct, categorical disorders. The mixed states described by Kraepelin were mainly seen in inpatients, and Hecker described milder depressive and hypomanic mixed states in outpatients. Kraepelin’s description of manic-depressive insanity included the fundamental states—temperaments (persistent traits) defined by the same symptoms of the manic, hypomanic, and depressive states, but of mild severity. A grading and overlap existed between manic, hypomanic, and depressive states and temperaments. The cyclothymic temperament (defined by frequent upswings and downswings of mood, thinking, and behaviour) was, according to Kraepelin, the main root of manic-depressive insanity.

### DSM-IV-TR classification of mood disorders

In DSM-IV-TR, mood disorders are categorically divided into bipolar disorders and depressive disorders (tables 2, 3, and 4). The basic feature distinguishing these groups is mania and hypomania, which is present in bipolar disorders, but absent in depressive disorders. According to Kraepelin, the recurrent course, rather than polarity, was probably the core feature of manic-depressive insanity. The symptoms of mania and hypomania are similar in DSM-IV-TR, except for increased severity of symptoms and presence of psychosis in mania (table 4).

This list of symptoms, the prioritisation of mood change, and the minimum duration of episodes specified, are consensus-based, not evidence-based. The criterion separating mania and hypomania—ie, marked impairment of functioning in mania—is an unclear boundary that can lead to misclassification. Recent findings have challenged these criteria and suggested different groupings of bipolar disorders and depressive disorders, according to the continuum or spectrum concept.

The distinction between bipolar I and II disorders in DSM-IV-TR is mainly based on findings that followed the criteria of Robins and Guze for diagnostic validity (clinical picture, delimitation, course including diagnostic stability, family history, biology).

### Clinical picture and delimitation

Many differences exist between mania, which defines bipolar I disorder, and hypomania, which defines bipolar II disorder. Psychosis often occurs in mania, but never in hypomania. Mania markedly impairs functioning, whereas hypomania often improves functioning, or, less commonly, causes mild impairment. Mania often requires hospital admission, whereas hypomania, by definition, never does. Mania usually has a longer duration than hypomania (months rather than weeks).

Thought disorders, all of which involve mental overactivity, differ between mania, which is associated with flight of ideas (speedy, disconnected thinking plus non-stop talking), and hypomania, in which the severity of thinking disorder can vary between creative thinking, crowded thoughts, racing thoughts, and infrequently, flight of ideas. The marked difference in functioning between mania and hypomania is partly related to differences in thinking: hypomania often includes creative and quick thinking that leads to increased goal-directed activities, or behavioural overactivity; whereas mania is associated with speedy, disconnected ideas, leading to aimless overactivity.

Critical evaluation of actions also differs: hypomania can include moderate risk-taking behaviour; mania includes severe risk-taking behaviour. That hypomania is not simply a mild version of mania is shown by differences in symptom structure. Mania has many defining factors or dimensions of symptoms, whereas hypomania has two to three factors, according to the sample studied (one factor dominated by elevated mood, one dominated by mental and behavioural overactivity, and one dominated by irritable risky behaviour). Differences between the sexes are also noted: in bipolar I disorder, female and male patients are equally common, whereas more female than male patients are seen in bipolar II disorder.

<table>
<thead>
<tr>
<th>Mood</th>
<th>Thinking</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mania/hypomania</td>
<td>Increased</td>
<td>Increased</td>
</tr>
<tr>
<td>Depression</td>
<td>Decreased</td>
<td>Increased</td>
</tr>
<tr>
<td>Mixed states (manic and depressive): concurrent, opposite polarity symptoms</td>
<td>Increased or decreased</td>
<td>Increased or decreased</td>
</tr>
</tbody>
</table>

Table 1: Kraepelin’s manic, hypomanic, depressive, and mixed states (episodes)

<table>
<thead>
<tr>
<th>Major depressive disorder</th>
<th>Dysthymic disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>No history of mania or hypomania</td>
<td>+</td>
</tr>
<tr>
<td>MDE</td>
<td>+</td>
</tr>
<tr>
<td>Depressive symptoms not meeting MDE criteria</td>
<td>+/-</td>
</tr>
<tr>
<td>Duration</td>
<td>At least 2 weeks</td>
</tr>
</tbody>
</table>

MDE=major depressive episode. +=present. –=absent.

Table 2: DSM-IV-TR classification of depressive disorders

<table>
<thead>
<tr>
<th>Bipolar I disorder</th>
<th>Bipolar II disorder</th>
<th>Cyclothymic disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mania and MDE</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>Hypomania and MDE</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Hypomanic and MDE</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>symptoms, frequently switching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>At least 1 week for mania, 2 weeks for MDE</td>
<td>At least 4 days for hypomania, 2 weeks for MDE</td>
</tr>
</tbody>
</table>

MDE=major depressive episode. +=present. –=absent.

Table 3: DSM-IV-TR classification of bipolar disorders
Course

Long-term follow-up studies have shown high diagnostic stability of bipolar I and II disorders—ie, patients who present with one type rarely go on to develop the other.9–22 The course differs: bipolar I disorder is more likely than bipolar II to include manic or hypomanic episodes and symptoms, whereas bipolar II disorder is more likely to include depressive episodes and symptoms.15,21

Family history

The type of bipolar disorder is likely to run in families—eg, bipolar I is more common than bipolar II in relatives of probands with bipolar I disorder, and vice versa.15–27

Biology

Pharmacological treatment differs somewhat between bipolar I and II disorders. In bipolar I disorder, complex drug combinations are needed to reach some stability, and antidepressants can induce or facilitate switching from depression to mania or hypomania. In bipolar II disorder, the risk of such switching caused by antidepressants is lower, and mood-stabilising agents may be less necessary—depression is more common than hypomania in this condition.28–30 However, drug treatment of bipolar II disorder is much understudied compared with that of bipolar I.15,31

Despite these differences, several lines of evidence (some of which result from different interpretations of the findings mentioned above) question the division between bipolar I and II disorders (panel 1). The findings that support a distinction between the two seem to be stronger than those that suggest a grading or continuum. The findings of Ghaemi and Baldessarini38 and Goodwin and colleagues39 seem to support a distinction between bipolar I and II disorders, and to support a mixed concept of the spectrum, including distinct disorders, with bipolar I disorder and major depressive disorder as extreme poles and, in the middle, a series of overlapping disorders including bipolar II and major depressive disorder plus bipolar signs. This concept of spectrum is partly different from Angst’s and Akiskal and colleagues’ continuum or spectrum concept, which includes partial overlap of mood disorders (including depressive disorders).

The reshaping of bipolar II disorder

The blurring of the boundaries between bipolar disorders and depressive disorders is mainly based on changes in the definition of bipolar II disorder, and on the re-emergence of the concept of mixed depression.31,32 DSM-IV-TR criteria for bipolar II disorder require the presence of recurrent depressive and hypomanic episodes (tables 2 and 3).

DSM-IV-TR reports a lifetime community prevalence of 0·5–4·8% for bipolar II disorder, 0·3–4·8% for bipolar II disorder, 0·5–6·3% for cyclothymic disorder, and 0·8–13·5% (5% is the most likely value) for bipolar II.28–30 Among depressed outpatients with bipolar II disorder, bipolar spectrum, or major depressive disorder, about 50% have been found to be bipolar.1,40–48

The increase in prevalence of bipolar II disorder, compared with that reported in older studies, is related to changes in diagnosis.28–30

Table 4: Criteria for major depressive episode and mania/hypomania according to DSM-IV-TR

<table>
<thead>
<tr>
<th>Major depressive episode</th>
<th>Mania/hypomania</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of symptoms</td>
<td>At least five*</td>
</tr>
<tr>
<td>Stem criterion (ie, always present)</td>
<td>Low mood or loss of interest*</td>
</tr>
<tr>
<td>Weight/eating</td>
<td>Increased or decreased*</td>
</tr>
<tr>
<td>Sleep</td>
<td>Increased or decreased*</td>
</tr>
<tr>
<td>Psychomotor activity</td>
<td>Increased or decreased*</td>
</tr>
<tr>
<td>Energy</td>
<td>Decreased*</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>Decreased*</td>
</tr>
<tr>
<td>Concentration</td>
<td>Decreased*</td>
</tr>
<tr>
<td>Suicidality</td>
<td>Present*</td>
</tr>
<tr>
<td>Talking</td>
<td>Decreased</td>
</tr>
<tr>
<td>Thinking</td>
<td>Decreased</td>
</tr>
<tr>
<td>Goal-directed activity (overactivity)</td>
<td>Decreased</td>
</tr>
<tr>
<td>Risky activity</td>
<td>Decreased</td>
</tr>
<tr>
<td>Psychotic symptoms</td>
<td>May be present</td>
</tr>
<tr>
<td>Duration</td>
<td>At least 2 weeks*</td>
</tr>
<tr>
<td>Functioning</td>
<td>Impaired*</td>
</tr>
<tr>
<td>Substances, antidepressants, medical disorders-induced</td>
<td>No</td>
</tr>
</tbody>
</table>

*Criteria symptoms.

Panel 1: Lines of evidence supporting a continuum between bipolar I and bipolar II disorder

- Similar age at onset15,54
- Same core feature of overactivity in mania and hypomania15,14
- Bipolar I and II have long-term diagnostic stability, but hypomania is frequently present in the course of bipolar I9–22
- Family history: more likely for relatives to have the same type of disorder as probands than different type (therefore, distinct disorders); but bipolar II not uncommon in relatives of bipolar I probands; bipolar I more common in relatives of bipolar II probands than in controls; major depressive disorder most common mood disorder in relatives of bipolar I and II probands9–22
- Grading of symptoms: syndromal and subsyndromal depressions more common in bipolar II than in bipolar I, syndromal and subsyndromal mania and hypomania more common in bipolar I9–22
- Grading of the switching to mania/hypomania during antidepressants (lower in bipolar II)9–22
- Clinical picture of depression: similar atypical symptoms (eg, oversleeping, overeating) and psychomotor changes15,14,21
- Manic and hypomanic symptoms predicting onset of both bipolar I and II15

Table 2: Characteristics of bipolar I and bipolar II disorder

<table>
<thead>
<tr>
<th>Bipolar I</th>
<th>Bipolar II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manic and hypomanic episodes</td>
<td>Recurrent depressive episodes</td>
</tr>
<tr>
<td>Bipolar I</td>
<td>Bipolar II</td>
</tr>
<tr>
<td>Manic and hypomanic episodes</td>
<td>Recurrent depressive episodes</td>
</tr>
<tr>
<td>Bipolar I</td>
<td>Bipolar II</td>
</tr>
<tr>
<td>Manic and hypomanic episodes</td>
<td>Recurrent depressive episodes</td>
</tr>
<tr>
<td>Bipolar I</td>
<td>Bipolar II</td>
</tr>
<tr>
<td>Manic and hypomanic episodes</td>
<td>Recurrent depressive episodes</td>
</tr>
</tbody>
</table>
to several advances,\textsuperscript{14,19–21} use of semi-structured interviews by trained clinicians, rather than the previous use of fully structured interviews by lay interviewers—the yes or no structured questions, and the absence of clinical evaluation in structured interviewing could lead to underdiagnosis of bipolar II disorder, and misdiagnosis as major depressive disorder; reduction of minimum duration of hypomania required to diagnose bipolar II disorder to 2 days (evidence-based), instead of the 4-day cutoff used by DSM-IV-TR (consensus-based); bypassing the stem question of the structured clinical interview for DSM-IV\textsuperscript{22} to assess history of hypomania—this question requires the patient to remember periods of elevated or irritable mood, but many individuals with bipolar II disorder give a negative response, because they see these periods as normal mood fluctuations, or, if the wording of the question includes “much more than usual”, as a severe disorder, which is often denied; shift in focus of investigation for history of hypomania from mood change to overactivity (increased goal-directed activities), an observable behaviour that is easier to remember than mood change, and frequently aids memories of a concurrent mood change.

Findings of several clinical and psychometric studies\textsuperscript{5–7,13–18,50,55–57} suggest that overactivity is the core feature of hypomania, lending support to its upgrading to the same priority level as mood change in the diagnosis of hypomania. Instead, however, DSM-IV-TR states that elevated mood is the “prototypical” symptom of hypomania, misleading clinicians into believing that hypomania cannot be diagnosed if elevated mood is absent. Overactivity—more than usual working hours, social life, sexual activity, travels, spending, brilliant ideas, and plans—can increase functioning if the concurrent mental overactivity is connected (as often happens), or can reduce functioning if mental overactivity is racing too much, leading to disconnection of ideas, or if it is focused more on risky behaviours. Bipolar II disorder and cyclothymic temperament (defined by high instability of mood, thinking, and behaviour) can occur at the same time;\textsuperscript{45,56} suicidality, substance abuse, and risky behaviours are more common when these conditions occur together, and this combination increases the instability of bipolar II disorder, easily leading to its misdiagnosis as borderline personality disorder, which is highly unstable.\textsuperscript{34,42}

**Bipolar II disorder compared with major depressive disorder**

The basic, distinguishing features of bipolar II disorder compared with major depressive disorder are lower age at onset, more recurrences, more bipolar (type I and II) family history, more atypical depressions (including symptoms such as overeating and oversleeping), and more mixed depression.\textsuperscript{1,2,9,14,45,47,51,54,55} Mixed depression, which is a key component of the continuum concept of mood disorders, might have important effects on treatment for depression.

**Mixed depression (depressive mixed states)**

Sporadic reports on mixed depression before the year 2000\textsuperscript{5} went almost unnoticed. By strictly following the DSM-IV structured clinical interview, mixed depression cannot be diagnosed in bipolar II and major depressive disorders, because this method does not allow assessment of hypomanic symptoms in major depressive episodes. In DSM-IV-TR, only bipolar I mixed state can be diagnosed, which requires the co-occurrence of full mania and major depressive episode. Mixed depression is defined by the combination of depression (major depressive episode) and non-euphoric, usually subsyndromal, manic or hypomanic symptoms (panel 2). It is present in bipolar I and II disorders, and in major depressive disorders.\textsuperscript{6,64–76} It is most studied in bipolar II and major depressive disorders. Intradesippressive manic or hypomanic symptoms may start spontaneously during depression (ie, manic/hypomanic symptoms are an independent episode), or can be related to the cycling between mania or hypomania and depression (a transition state, believed by Kraepelin to be the cause of most mixed states). The manic or hypomanic symptoms of mixed depression can also be induced or worsened by antidepressants.

Mixed depression has a moderate diagnostic stability, which does not support a categorical distinction.\textsuperscript{77} Compared with non-mixed depression, it is more common in bipolar disorders, and is more associated with family history of bipolar disorders, lower age at onset, longer duration, worse outcome, and poorer response to treatments.\textsuperscript{20,42,43,50,57–59} These differences might aid its subtyping.

**Definitions**

Few studies have included samples of untreated mood disorders, allowing the study of spontaneous-onset intradesipressive manic or hypomanic symptoms. Several definitions of mixed depression have been

---

**Panel 2: Suggested diagnostic criteria for mixed depression**

(A) A major depressive episode according to DSM-IV-TR

(B) Two to three or more concurrent DSM-IV-TR manic/hypomanic symptoms

(C) Specifiers

- Transition state during cycling between mania/hypomania and depression
- Independent episode (ie, unrelated to cycling)
- With psychomotor agitation (agitated depression)
- Antidepressant-associated
- In bipolar I disorder
- In bipolar II disorder
- In major depressive disorder
suggested, some of which require a cutoff of at least two to three intradepression manic/hypomanic symptoms, and some of which require specific manic/hypomanic symptoms, such as psychomotor agitation (agitated depression). At present, the definition that requires a cutoff of three symptoms is the most validated.

The diagnostic validity of agitated depression, a diagnosis not included in DSM-III and DSM-IV but frequently used in clinical practice, has been tested. Agitated depression was often found to be mixed and bipolar. When agitated depression was mixed it differed, in terms of bipolar validators, from non-agitated depression, but when it was not mixed it did not differ from non-agitated depression. These findings question the diagnostic validity of the definition and suggest that agitated depression is probably a more mixed variant of mixed depression.

Frequency
Frequency of mixed depression in bipolar I, bipolar II, and major depressive disorders ranges between 20% and 70%, depending on setting, interview methods, instruments used to assess intradepression manic or hypomanic symptoms, prevalence of bipolar disorders compared with major depressive disorder, whether individuals have been treated or not, and definitions used. It is more common in bipolar I and II disorders than in major depressive disorder, but is not uncommon in major depressive disorder (around 30% of outpatients with the condition may have mixed depression).

Manic and hypomanic symptoms and bipolar nature
The most common manic and hypomanic symptoms of mixed depression, which are more severe in bipolar I than in bipolar II and major depressive disorders, are irritability, mental overactivity (flight of ideas, racing thoughts, crowded thoughts, creative thinking), and behavioural overactivity (psychomotor agitation and increased talkative ness), which represent the condition’s “manic foundation”, according to Kraepelin. Manic and hypomanic symptoms listed in DSM-IV-TR include those (such as irritability and psychomotor agitation) that are present in many axis 1 disorders—especially anxiety disorders, which are common in mixed depression—and in axis 2 (personality) disorders. Whether a correlation exists between psychomotor agitation and anxiety is unclear. Irritability, psychomotor agitation, and racing or crowded thoughts are more likely to be bipolar than non-bipolar in samples of individuals with mood disorders. The bipolar nature of these symptoms is supported, in individuals with mood disorders, by a family history of bipolar disorders. Kraepelin included these symptoms among the basic symptoms of the manic and hypomanic states. However, these symptoms are unlikely to have a bipolar nature when present in non-mood axis 1 disorders. The same reasoning could be applied to psychotic symptoms, which are present in many axis 1 disorders (including mood disorders), but are the stem criterion for diagnosis of psychotic disorders.

Several lines of evidence indicate that mixed depression has a bipolar nature (panel 3).

A categorical definition?
A categorical definition of mixed depression is questionable, following Kendell and Jablensky’s method based on finding a bi-modal distribution, because the distribution of intradepression hypomanic symptoms between bipolar II disorder and major depressive disorder is not bi-modal, and the distribution of lifetime manic or hypomanic symptoms between bipolar I disorder and major depressive disorder is not bi-modal, and the distribution of depressive and manic symptoms of mania is not bi-modal. Furthermore, a proportional relation exists between intradepression hypomanic symptoms and the degree of bipolar family history in bipolar II and major depressive disorders, which should not be found in distinct disorders. For clinicians, however, categories are more useful than dimensions, and a cutoff in the dimensions of mixed depression that defines the need for treatment has to be found. Some investigators have tested definitions of two to three intradepression manic and hypomanic symptoms in pharmacological studies, and found that the presence of these two to three symptoms increased the risk of switching to mania or hypomania in bipolar I and II depression treated by antidepressants.

Panel 3: Lines of evidence supporting the bipolar nature of mixed depression

- More common in bipolar I and II disorders than in MDD
- Family history of bipolar disorders
- Dose-response relationship between intradepression hypomanic symptoms and bipolar family history loading
- “Manic/hypomanic/excitement” factor present in MDD
- Factor structure of intra-depression hypomanic symptoms similar to factor structure of inter-depression hypomania
- Intradepression full hypomania
- Mixed MDD closer to bipolar disorder than to non-mixed MDD on bipolar validators
- MDD shifting to bipolar disorder more likely to have mixed depression
- Mixed depression, compared with non-mixed depression, more likely to switch to mania/hypomania with antidepressants
- Worsening by antidepressants

MDD=major depressive disorder.
Implications

From a clinical practice point of view, mixed depression has systematic implications. Depression should be systematically assessed for concurrent manic or hypomanic symptoms, which, if present, should lead to a careful probing for history of mania or hypomania. Antidepressant treatment of mixed depression should be cautious, because antidepressants alone—ie, not protected by mood stabilising agents—can worsen intradepression manic and hypomanic symptoms such as racing thoughts and psychomotor agitation, sometimes leading to suicidal behaviour. Switching of depression to mania or hypomania while taking antidepressants might be more likely in mixed than in non-mixed depression.46,59,73 The US Food and Drug Administration (FDA), in its recent warning on antidepressants, has listed among possible precursors of occasional suicidality related to antidepressants several features of mixed depression, such as irritability, psychomotor agitation, and bipolarity.96 Mixed depression is associated with a higher risk of suicidality than is non-mixed depression.57,95,102 Intradepression psychomotor agitation and racing or crowded thoughts are independent predictors of suicidal ideas, and mixed depression is often present in individuals before suicide attempts.96–98. Therefore, the suicidality sometimes related to antidepressants is probably not caused by the antidepressants, but by clinicians using only antidepressants to treat mixed depression.95,107,108 In mixed depression, it seems logical first to control the intradepression manic or hypomanic symptoms with mood stabilising agents, and later (or concurrently) to add an antidepressant, if necessary.96 Controlled trials are needed to provide more evidence, but are almost impossible to do in view of the FDA warning. Therefore, retrospective, naturalistic, and re-analysis studies are the only available source of information about treatment of mixed depression. The current, limited evidence seems to lend support to the use of mood stabilisers alone or with antidepressants for treatment of mixed depression.46,97,101,104

Bipolar disorders in children and adolescents

Perhaps the most advanced studies about bipolar disorders are those that have been done in children and adolescents. These findings might increase insight into the onset and mechanisms of the disease, and lead to treatments that prevent development of adult bipolar disorders.

Prevalence of bipolar disorders in children (mainly bipolar II disorder and bipolar spectrum) can be around 6% in community samples, and in 2–15% in clinical samples. Bipolar disorders in children are often mixed, and feature very rapid cycling. Mania and hypomania, which are much more common than depression, are often brief (less than minimum duration specified by DSM-IV-TR) and highly recurrent. Course and outcome of childhood or adolescent bipolar disorders are more severe than in adult-onset bipolar disorders. Likelihood of family history of bipolar disorder is high, and is greater than in adult-onset bipolar disorders. Mixed mania and mixed depression are common. Irritability is not a distinguishing feature of these mixed states, which have a worse course and outcome than in adults—ie, more recurrences, more impairment, and greater severity. Psychosis, suicidality, anxiety disorders, substance abuse, antisocial behaviour, and attention-deficit hyperactivity disorder (ADHD) frequently co-occur in bipolar disorders of this age group. Continuous mood lability, irritability, and temper outbursts are also a common state, not meeting any DSM-IV-TR criteria. Differential diagnosis between bipolar disorders and ADHD is based on classic manic and hypomanic symptoms such as elevated mood, grandiosity, and racing thoughts, and on family history of bipolar disorder, because irritability, hyperactivity, and short attention span are present in both disorders. Misdiagnosis of bipolar disorder as ADHD and major depressive disorder is common, leading to the use of stimulants and antidepressants, which might worsen the course instead of mood stabilising agents.105–107

The continuum or spectrum concept of mood disorders

The DSM-IV-TR classification of mood disorders is based on the concept of unipolar-bipolar division,118 which took the place of Kraepelin’s unitary view in the 1960s. According to the unipolar-bipolar division concept, unipolar depressive disorders and bipolar disorders are distinct, categorical disorders, separated by clear boundaries. The division of mood disorders was based on differences seen with several diagnostic validators. Compared with unipolar depression, bipolar disorder (at that time only bipolar I was diagnosed), was found to have, apart from mania, the following distinguishing features: sex (female as common as male in bipolar I, female more common than male in unipolar depression); age at onset (often younger than 30 years in bipolar I disorder, often older in unipolar depression); number of episodes (greater in bipolar I disorder); clinical picture of depression (psychomotor retardation and hypersomnia in bipolar I disorder, psychomotor agitation and insomnia in unipolar depression); and family history (mania common in relatives of bipolar I disorder probands, uncommon in relatives of unipolar depression probands). An important limitation of this division is that bipolar II disorder, which seems to be the link between bipolar and depressive disorders,107 was not included in this classification. Increasingly, evidence supports instead a return to Kraepelin’s unitary concept of mood disorders. The continuum (also known as spectrum or dimensional) concept of mood disorders includes disorders ranging from bipolar I and II to recurrent major depressive disorder and minor
Research on bipolar spectrum disorders in children and adolescents has added further support to the concept. The continuum of mood disorders is based on several lines of evidence summarised in panel 4.

**Treatment findings**

The treatment of mood disorders is not discussed in detail in this Seminar; the focus will be on the pharmacological and psychotherapeutic evidence supporting and not supporting the continuum concept.

Response to antidepressants and anti-bipolar agents can be seen as a proxy of the biology of mood disorders. A similar response by patients with supposedly different disorders to these drugs might support the continuum concept, whereas differing responses might support the division of mood disorders. Since antidepressants are the gold standard of treatment in depressive disorders, whereas lithium is the gold standard of treatment for bipolar disorders, the basic test is the assessment of the efficacy of lithium in depressive disorders, and of antidepressants in bipolar disorders. Randomised, double-blind, placebo-controlled trials can show the best evidence of efficacy, but these trials are limited by the selection of individuals not representative of usual clinical practice. Naturalistic studies of effectiveness in usual clinical practice are more useful, if designed to reduce biases.

**Antidepressants in bipolar and depressive disorders**

Evidence for the efficacy and effectiveness of antidepressants in acute bipolar depression is mixed.124-128 If antidepressants were as effective in bipolar depression as in major depressive disorder, the concept of a continuum would be supported.

The cycling of depression and mania or hypomania is part of the natural course of bipolar disorders; whether antidepressants do or do not induce switching is uncertain. In depressive disorders, antidepressants do not seem to increase switching more than placebo.129 If antidepressants facilitated switching in bipolar disorders much more than in depressive disorders, the concept of a continuum would not be supported.

Some evidence suggests that antidepressants might increase the number of recurrences (manic and depressive episodes) in bipolar disorders, including rapid cycling, which is defined as four or more episodes per year (mainly depressive, and mainly in bipolar II disorder).127 However, the evidence for this pattern is weak. Kraepelin reported a spontaneous increase of recurrences related to ageing in a subgroup of patients with manic-depressive insanity. Evidence showing that antidepressants increased recurrences in bipolar disorders would contradict the concept of a continuum, since antidepressants can prevent recurrences in major depressive disorder. Prevention of recurrences of bipolar (mainly bipolar II) depression by antidepressants, if any, is weak.124-128,132 Antidepressants are not efficacious in—and can actually worsen—acute mania and hypomania, and can induce, rather than prevent episodes of mania and hypomania.127

**Lithium in bipolar disorders and in depressive disorders**

Table 5 compares the effects of antidepressants and lithium on bipolar disorders and on depressive disorders. In acute bipolar (mainly bipolar I) depression, lithium seems to be effective124,128,132 and no other antidepressant has been shown to be more effective.124,127,132 In acute major depressive disorder, the efficacy of lithium, if any, is lower than in bipolar disorder.127 Cycling and switching can be prevented by lithium.124,128,132 Lithium can prevent recurrences in bipolar disorders (more effective in mania than depression in bipolar I, as effective in depression as in hypomania in bipolar II) and in depressive disorders, and might be as effective as antidepressants in preventing recurrences in MDD.127,128,132 Lithium is effective in treatment of acute mania and hypomania and in prevention of mania and hypomania.124,132 The drug has been surpassed (in clinical use, but not in efficacy) by many modern mood-stabilising agents, such as valproic acid, carbamazepine, neuroleptics (first-generation antipsychotics), and second-generation antipsychotics such as olanzapine, risperidone, ziprasidone, quetiapine, and aripiprazole. These agents are effective in acute mania and hypomania, and in the prevention of mania and hypomania, but less effective or ineffective in prevention of bipolar I depression.128,129 Only quetiapine seems effective in acute bipolar I and II depression,129,130 but more replications in naturalistic settings are needed. Evidence for efficacy of lamotrigine in acute treatment and prevention of depression in bipolar I and II disorders is not strong (negative, unpublished studies.

---

**Panel 4: Lines of evidence supporting continuum of mood disorders**

- Grading of severity, duration, and frequency of mood episodes
- Mixture of depressive and manic/hypomanic symptoms in the same episode (the mixed states)
- Lack of bi-modality in manic/hypomanic symptoms in or between bipolar disorder and MDD
- Correlation between manic/hypomanic symptoms and depressive symptoms in bipolar family history (ie, the higher the number/score of intradepression hypomanic symptoms, the higher the bipolar family history loading)
- No bi-modality of age at onset between bipolar disorder and MDD
- No bi-modality of temperament instability between bipolar disorder and MDD
- Many MDD shifting to bipolar disorder in the long term
- Bipolar signs common in MDD (such as bipolar family history, young age at onset, many recurrences, lifetime and intradepression manic/hypomanic symptoms)
- Manic/hypomanic symptoms predicting onset of bipolar disorder and MDD
- Residual manic symptoms predicting depressive recurrences in bipolar disorder

MDD=major depressive disorder
have been done), and the drug has little or no effect against mania.19,20 It has not been formally tested in major depressive disorder. The combination of olanzapine and fluoxetine seems effective in bipolar I depression,21 but replication studies are needed.

Psychological interventions
Cognitive behavioural therapy is effective in non-psychotic major depressive disorder. Bipolar I and II depression are associated with more behavioural symptoms (eg, loss of energy, hypersomnia) than cognitive symptoms (eg, low self-esteem) compared with major depressive disorder. Variants of cognitive behavioural therapy that are more focused on behaviour seem to be useful in bipolar depression. Interpersonal social rhythm therapy, which is mainly focused on behavioural change, seems as effective as major depressive disorder as in bipolar depression. Efficacy of cognitive behavioural therapy and interpersonal social rhythm therapy is assessed on the basis of effects on symptoms and recurrences.13,24

Summary of treatment findings
The effects of antidepressants and lithium in bipolar disorders and in depressive disorders seem to have more differences than similarities, contradicting the idea of a continuum of mood disorders (table 5). However, treatment of bipolar II disorder is understudied.10 The risk of switching to hypomania by use of antidepressants seems lower in bipolar II than in bipolar I disorder,26–30 and lithium seems effective in preventing recurrences in bipolar II disorder.17 Because bipolar I disorder and major depressive disorder without bipolar signs are the extreme poles of the mood spectrum, differences in treatment responses are more likely to be noted than similarities. However, bipolar II is a disorder positioned between these extremes, and its pharmacotherapy could add important evidence in support of, or contradicting, the continuum concept. Prevention of recurrences in bipolar II disorder and major depressive disorder by lithium seems to be preliminary pharmacological evidence in support of a continuum of mood disorders, but several pharmacological data seem to support the opposite view.

Conclusion
Bipolar disorders, especially bipolar II and the related bipolar spectrum disorders, are much more common that previously reported, and treatment of bipolar II disorder and bipolar spectrum disorders is much understudied. Controlled pharmacological studies are greatly needed, and should stratify samples taking into account the frequently mixed profile of bipolar disorders.

Conflict of interest statement
I have received speakers’ fees from Eli Lilly and Wyeth.

References


81 Benazzi F. Which could be a clinically useful definition of depressive mixed state? Prog Neuropsychopharmacol Biol Psychiatry 2002; 26: 1105–11.


120 Benazzi F. Does age at onset support a dimensional relationship between bipolar II disorder and major depressive disorder? World J Biol Psychiatry 2003; 4: 1–16.

121 Benazzi F. Does temperamental instability support a continuity between bipolar II disorder and major depressive disorder? Eur Psychiatry 2006; 21: 274–79.


126 Moller HJ, Gruenze H, Broich K. Do recent efficacy data on the drug treatment of acute bipolar depression support the position that drugs other than antidepressants are the treatment of choice? A conceptual review. Eur Arch Psychiatry Clin Neurosci 2006; 256: 1–16.


Management of functional somatic syndromes

Peter Henningsen, Stephan Zipfel, Wolfgang Herzog

Although functional somatic syndromes (FSS) show substantial overlap, treatment research is mostly confined to single syndromes, with a lack of valid and generally accepted diagnostic criteria across medical specialties. Here, we review management for the full variety of FSS, drawn from systematic reviews and meta-analyses since 2001, and give recommendations for a stepped care approach that differentiates between uncomplicated and complicated FSS. Non-pharmacological treatments involving active participation of patients, such as exercise and psychotherapy, seem to be more effective than those that involve passive physical measures, including injections and operations. Pharmacological agents with CNS action seem to be more consistently effective than drugs aiming at restoration of peripheral physiological dysfunction. A balance between biomedical, organ-oriented, and cognitive interpersonal approaches is most appropriate at this truly psychosomatic interface. In view of the iatrogenic component in the maintenance of FSS, doctor-centred interventions and close observation of the doctor–patient relationship are of particular importance.

Search strategy and selection criteria

We searched the Cochrane Library, MEDLINE, and PSYCINFO from 2001. We used a string of search terms previously applied in a meta-analysis on FSS and diagnostic analogues and adapted it to the terms of FSS that are indicated in table 1. We included commonly referenced and highly regarded publications from before 2001. We also searched the reference lists of articles identified by this search strategy and selected those we judged relevant. For the section on therapeutic evidence we restrict presentation of search results to systematic reviews and meta-analyses published since 2001. Reviews had to define a comprehensive search strategy and quality evaluation of primary studies and present results in a systematic way. Whereas some of these reviews cover all aspects of treatment of single FSS, others concentrate on particular treatment components for single FSS or across different types of medically unexplained symptoms. We did not include systematic reviews on particular treatments not focused on FSS or diagnostic analogues.
criteria can produce wide variations in prevalence rates, for example, whereas the prevalence for chronic fatigue syndrome in the population has been estimated to be 0·2%, it is 9% for chronic unexplained fatigue. For some syndromes, the published prevalence rates are also influenced by the special interests of supporters of the respective concept of FSS.

Overlap of single FSS

Many clinicians and researchers who focus on one specific FSS have the impression that it occurs in isolated form. Conceptually, however, the bodily symptoms used as diagnostic criteria (ie, abdominal pain or fatigue) are not specific but overlap for many FSS. Many patients fulfil criteria for more than one syndrome. The extent of this empirical overlap between single FSS is from around 10% in the general population to over 50% in clinic populations. Furthermore, multiple studies have shown high self-reported scores of other unexplained bodily symptoms in different FSS, including pseudo-neurological symptoms and chronic low-back pain. Common characteristics of FSS, in addition to the pattern of bodily complaints, are female preponderance and significant overlap with anxiety and depression, which is higher in FSS than in comparable, organically explained diseases (ie, irritable bowel syndrome vs inflammatory bowel disease or fibromyalgia vs rheumatoid arthritis); however, many cases of FSS also occur without anxiety or depression. This association therefore neither can be seen as an unspecific psychological reaction to the presence of bodily complaints nor as masked or somatised depression or anxiety alone. Further common features of FSS are the substantial effect on quality of life, which is commonly as large as in comparable diseases of clear organic origin, and the response to interventions that primarily influence the function of the CNS by behavioural, psychotherapeutic, or psychopharmacological means.

Taken together, the balancing of research and therapeutic practice between a splitting and a lumping view of FSS is a clinical and conceptual challenge, and this is related to the ongoing debate whether FSS should be classified as a physical or mental disorder, or whether a truly somatic third-way is possible.

Terminology and classification

Currently, no term or classification is fully satisfactory when dealing with the clinical phenomenon of patients reporting persistent bodily complaints for which no clear organic reason can be found. The term medically unexplained symptoms is not adequate because of the difficulties in defining insufficient explanation and also because it implies that explanations that involve psychosocial or cultural factors are not part of medicine. Many of the terms used for single FSS are classified in the different medical sections of International Classification of Diseases 10th Revision (ICD 10). They generally have good acceptance by patients and medical specialists, but they encourage a narrow or “splitting” view by the respective specialist on the current main bodily symptom and symptomatic treatment only; in addition, most of them give no indication of symptom severity. In parallel, there is the possibility to classify organically unexplained bodily symptoms as mental disorders: in chapter V of ICD 10, and in the Diagnostic and Statistical Manual, they are classified as somatoform disorders, with somatisation disorder as the rare, most serious prototype used for patients with multiple symptoms over time. This type of classification encourages a “lumping” perspective (ie, to look at the whole pattern of current and previous bodily symptoms and also at psychological and behavioural characteristics of the patient). However, this perspective historically implied a psychogenic origin of FSS that is not only problematic from a scientific point of view, but also offensive for patients who want to avoid being seen as mentally ill.

Debates about future editions of the classifications for mental disorders range from the contested suggestion to abolish the category somatoform disorders altogether to the suggestion to introduce a category of general-medicine–psychiatry interface disorders. From a

<table>
<thead>
<tr>
<th>Number of reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irritable bowel syndrome (IBS)</td>
</tr>
<tr>
<td>Chronic fatigue syndrome (CFS)</td>
</tr>
<tr>
<td>Fibromyalgia (FMS)</td>
</tr>
<tr>
<td>Multiple chemical sensitivity</td>
</tr>
<tr>
<td>Nonspecific chest pain</td>
</tr>
<tr>
<td>Premenstrual syndrome</td>
</tr>
<tr>
<td>Non-ulcer dyspepsia</td>
</tr>
<tr>
<td>Repetitive strain injury</td>
</tr>
<tr>
<td>Tension headache</td>
</tr>
<tr>
<td>Temporomandibular joint disorder</td>
</tr>
<tr>
<td>Atypical facial pain</td>
</tr>
<tr>
<td>Hyperventilation syndrome</td>
</tr>
<tr>
<td>Globus syndrome</td>
</tr>
<tr>
<td>Sick building syndrome</td>
</tr>
<tr>
<td>Chronic pelvic pain</td>
</tr>
<tr>
<td>Chronic whiplash syndrome</td>
</tr>
<tr>
<td>Chronic lyme disease</td>
</tr>
<tr>
<td>Silicone breast implant effects</td>
</tr>
<tr>
<td>Candidiosis hypersensitivity</td>
</tr>
<tr>
<td>Food “allergy”</td>
</tr>
<tr>
<td>Gulf War syndrome</td>
</tr>
<tr>
<td>Mitral valve prolapse</td>
</tr>
<tr>
<td>Hypoglycaemia</td>
</tr>
<tr>
<td>Chronic low back pain</td>
</tr>
<tr>
<td>Dizziness</td>
</tr>
<tr>
<td>Interstitial cystitis</td>
</tr>
<tr>
<td>Tinnitus</td>
</tr>
<tr>
<td>Pseudoseizures</td>
</tr>
<tr>
<td>Insomnia</td>
</tr>
</tbody>
</table>

Table 1: Number of reviews in which individual FSS are mentioned.

Vol 369 March 17, 2007 947

www.thelancet.com
treatment perspective, the latter suggestion might be a good way to increase awareness for the necessity to use a lumping perspective in many patients with FSS without resorting to the idea of a purely mental or psychogenic disorder.

**Aetiology and pathophysiology**

Management of FSS must be informed by knowledge about the diversity of predisposing, precipitating, and maintaining factors for FSS. In terms of predisposing factors, no clear pattern of genetic influences has been identified, and the heritability of FSS seems to be small.\(^{29,30}\) Childhood adversities are not restricted to sexual or physical abuse.\(^{31}\) Childhood experience of organically unexplained symptoms, parental ill health, and increased parental illness behaviour for bodily symptoms in the child increase the risk of FSS later in life.\(^{32}\) Personality factors, such as cognitive styles and attachment patterns, might affect the maladaptive illness behaviour in patients with FSS,\(^{33,34}\) and in more severe cases, co-morbidity of FSS with personality disorders is high.\(^{35}\)

Precipitating factors for the development of FSS can be organic illness,\(^{36,37}\) accidents,\(^{38}\) stressful life events,\(^{39,40}\) and wider psychosocial mechanisms, such as media coverage of potentially pathogenic environmental health hazards or mass hysteria.\(^{41}\)

Apart from organic comorbidity, maintaining factors have mainly been described on the psychosocial level, such as personality factors that contribute to predisposition, mental comorbidity, a persisting organic illness attribution, and context factors surrounding so-called secondary gain. However, the behaviour of treating physicians also contributes to maintenance and exacerbations of FSS.\(^{42}\) For example, qualitative research in primary care showed that patients with so-called unexplained symptoms often voice psychosocial clues that are not taken up by their doctors; instead, doctors and not patients themselves somatise—ie, commonly initiate further diagnostic testing despite the assumption that the complaint is not explained by structural disease.\(^{42,43}\)

Sociocultural factors clearly affect symptom perception and reporting, and knowledge of explanatory models of bodily distress for patients from different cultural backgrounds is useful in the establishment of a stable doctor–patient relationship.\(^{44}\) Cultural differences also directly affect prevalence of FSS. For instance, German citizens reported up to twice the frequency of back pain compared with those in the UK, and the difference is not attributable to different risk-factor profiles.\(^{45}\)

The model of causation for FSS suggests that the central clinical phenomena, such as experience of chronic bodily symptoms and loss of functioning, are influenced by multiple, but specifiable biological, psychological, interpersonal, and social factors (figure).

Views on pathophysiological changes in FSS reflect the different foci for treatment, with the debates centring around two dichotomies: peripheral changes (eg, immunological, endocrine, muscular, cardiac, and intestinal) versus CNS changes and changes specific for one FSS versus general changes for all FSS.\(^{24,46,47}\) In addition, whether pathophysiological changes are causally relevant or, for instance, a consequence of behavioural changes due to FSS is unclear.\(^{48}\) There is currently a general tendency to view central phenomena, such as experience of chronic bodily symptoms and loss of functioning, as influenced by multiple, but specifiable biological, psychological, interpersonal, and social factors (figure).

Balanced approaches to FSS

There is not one single focus in the management of FSS, and this ambiguity has to be seen as a characteristic feature of these syndromes (panel 1). Some patients (eg, those with uncomplicated irritable bowel syndrome) will benefit most from typical medical interventions that are limited to reassurance and symptomatic relief focussed on gut functioning, but diagnostic elements of the cognitive interpersonal approach have to be integrated to assure that
Evidence for the management of FSS

Description of treatment foci and types

We could not identify trials that tested treatments of groups of patients fulfilling criteria for more than one FSS. Differentiation of the following five treatment foci and types is useful for the organisation of the evidence on the treatment of single FSS and diagnostic analogues, such as somatoform disorders (table 2). Peripheral pharmacotherapy is primarily aimed at peripheral physiological processes (eg, bowel function, muscle tension, inflammation, nociceptive pain, etc). Central pharmacotherapy is primarily aimed at central processes of sensation, cognition, and affect (the distinction of peripheral and central is blurred in some cases, such as with drugs that alter serotonin metabolism). Active behavioural intervention is aimed at change of bodily and interpersonal behaviours, sensations, and cognition and is effected with active participation of patients in treatments, such as exercise and different psychotherapies. Multidisciplinary treatments are also included in this group because active behavioural components are essential parts of them. Passive physical intervention is aimed at passive, non-pharmacological change of peripheral features of the syndrome via physical (including surgical and other skin penetrating) means. For the final treatment group, the rationale is either outside that of those mentioned above (usually seen as part of complementary or alternative medicine) or is not patient-centred but doctor-centred (ie, aimed at doctor’s behaviour via education and training).

Treatment of single FSS and diagnostic analogues

Irritable bowel syndrome

Research on treatment in irritable bowel syndrome is more extensive than for other classic FSS, although, until recently, most research of this subject was thought to be of rather low quality. Current evidence is unequivocal on the value or not for some peripheral pharmacological agents: bulking agents and loperamide seem ineffective, the 5-HT4-agonist tegaserod and the 5-HT3-antagonist alosetron seem effective in selected subgroups (female patients with irritable bowel syndrome dominated by constipation or diarrhoea, respectively). There is moderate evidence for the efficacy of antidepressants and psychotherapy; for the latter, there is not enough evidence to differentiate between different forms. However, there are significant differences in the interpretation of primary studies among reviews. For the interpretation of psychotherapies for irritable bowel syndrome in particular, there is an ongoing debate of whether the application of standard criteria for drug treatments is feasible for the assessment of psychotherapies or not. Differences of opinion in this matter are apparent between gastroenterological and psychological-medicine specialists and even among groups of North American gastroenterologists.

Fibromyalgia

One of the weak spots of treatment research for fibromyalgia syndrome is the heterogeneous nature of outcome measures beyond pain. Attendees of a recent workshop rated global patient-rated improvement, fatigue, health-related quality of life, and multidimensional function as clinically most important outcome measures after pain, whereas manual tender-point examination loses significance in view of the limited validity and usefulness of this criterion. However, there seems to be mounting evidence that drugs with primarily central

---

Panel 1: Different foci in the management of FSS

Focus on patient

Organ-oriented approach
- Current main bodily lead symptoms
- Focus on dysfunction of peripheral organs
- Interventions aimed at peripheral physiology and restoration of organ function

Cognitive interpersonal approach
- Pattern of bodily and mental symptoms over time
- Focus on dysfunction of central processing and context factors
- Interventions aimed at sensations, cognitions, affects, behaviours, and restoration of overall functioning

Focus on doctor
- Early recognition
- Communication skill
- Avoidance of iatrogenic harm

Focus on context factors
- Doctor reimbursement system
- Patient compensation schemes
- Health-care system
- Workplace characteristics
- Cultural beliefs
<table>
<thead>
<tr>
<th>Pharmacotherapy</th>
<th>Non-pharmacological therapy</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peripheral</td>
<td>Psychotherapy and active behavioural therapy</td>
<td>Passive physical intervention</td>
</tr>
<tr>
<td>Central</td>
<td>Passive physical intervention</td>
<td>Chinese herbal medicine</td>
</tr>
</tbody>
</table>

**Irritable bowel syndrome**
- Tegaserod +++ (IBS-C, w)
- Alosetron +++ (IBS-D, w)
- Spasmolytic agents +
- Bulking agents +
- Prokinetic agents –
- Loperamide –
- Tricyclic antidepressants +
- CBT ++
- Hypnotherapy ++
- Psychotherapy (CBT, hypnotherapy, psychodynamic-interpersonal) +
- Acupuncture ++

**Fibromyalgia**
- Corticosteroids –
- NSAID –
- Antidepressants +
- CBT ++
- Exercise +++

**Chronic fatigue syndrome**
- Immunoglobulin –
- Antidepressants +
- CBT ++
- Exercise +++

**Non-ulcer dyspepsia**
- Proton pump inhibitors +
- Helicobacter pylori eradication +
- Prokinetics +
- Antacids –
- Tricyclic antidepressants ++
- Psychological interventions ++
- Adhesiolysis +

**Tension headache**
- Botulinum toxin +
- Antidepressants +
- Behavioural therapy +++
- Spinal manipulation –
- Acupuncture +

**Non-specific chest pain**
- Progestogen –
- Sertraline –
- Counselling and ultrasound scanning –
- Multidisciplinary treatment –
- Uterine nerve ablation –

**Premenstrual syndrome**
- Progesterone/progestogen –
- GnRHα –
- CBT ++
- Exercise therapy ++

**Temporomandibular joint disorder**
- Analgesics –
- Antidepressants +
- CBT for electromagnetic hypersensitivity ++
- Screen filters or shields –
- Selenium for environmental illness –

**Environmental illness or electromagnetic hypersensitivity**
- Nonsteroidal anti-inflammatory drugs +
- Muscle relaxants +
- Tricyclic antidepressants +
- Radiofrequency denervation +
- Acupuncture +

**Chronic low-back pain**
- Diagnostic analogues (MUS, SD, CD) –
- CBT for SD or MUS +
- Hypnosis for CD +
- Consultation letter to doctor for SD or MUS +

---

IBS=irritable bowel syndrome (C=constipation, D=diarrhoea). NSAID=non-steroidal anti-inflammatory drug. SSRI=selective serotonin reuptake inhibitors. SNRI=selective serotonin and norepinephrine reuptake inhibitor. GnRHα=gonadotrophin releasing hormone analogue. CBT=cognitive behavioural therapy. MUS=multiple unexplained physical symptoms. SD=somatoform disorders. CD=conversion disorder. +++=strong evidence; ++=moderate evidence; +weak evidence; –no evidence for efficacy of treatment. This table does not list all treatments with weak or no evidence dealt with in the systematic reviews. For simplicity, strength of evidence for efficacy of a specific treatment type is indicated in four different grades, with the reviews contributing to this summary estimate. 23 Cochrane reviews are included. For such an integration of systematic reviews, which use different criteria and represent different opinions in heterogeneous clinical fields, estimation of effect sizes was not feasible. The table shows general empirical trends in the management of functional somatic syndromes; it is not an adequate basis for individual treatment recommendations. The terms used in the table were taken from the systematic reviews and vary in grade of differentiation (eg, for some FSS the reviews state the evidence for psychotherapy, whereas for others they state the evidence for different forms of psychotherapy separately).

Table 2: Management of FSS—evidence from systematic reviews since 2001
action and active behavioural interventions are effective, whereas peripheral drugs and passive physical interventions show only weak or no evidence for efficacy.

**Chronic fatigue syndrome**
Few studies have been added to the therapy research literature on chronic fatigue syndrome since 2001. The two treatments with best evidence for efficacy since the 1990s are graded exercise and cognitive behavioural therapy, whereas there is little evidence to support the use of antidepressants and different immunological and steroidal drugs. In one systematic review, around 130 different outcome measures were identified in 44 studies, thus limiting the generalisability of the findings.

**Non-ulcer dyspepsia**
The evidence for peripheral drugs is best for proton-pump inhibitors. Evidence for others is weaker, with many contradicting studies and systematic reviews from the past currently distilled into evidence for a consistent but limited effect of proton-pump inhibitors, \( \text{H}_2 \)-receptor agonists, and prokinetics and of *Helicobacter pylori* eradication in infected patients. Antidepressants have not been tested rigorously in this type of functional gastrointestinal disorder. All four studies on psychotherapy showed positive effects for different types of psychotherapy each, but interpretation is limited by small sample sizes and adjustments for differences between intervention and control groups.

**Tension headache**
Behavioural psychotherapies and tricyclic antidepressants seem to have the best treatment effects in patients with tension headache. Patients given acupuncture improved relative to waiting-list controls, but this intervention is probably no more effective than minimum or sham acupuncture. Botulinum toxin, although showing promise in open trials, shows contradictory results in randomised controlled trials.

**Non-specific chest pain**
The only systematic reviews we could identify were of psychological interventions for which there is moderate evidence of efficacy. This syndrome refers to chest pain but no evidence of underlying coronary, gastrointestinal, or other organic pathology, therefore it is more exclusive than the category non-cardiac chest pain. For the latter, there is some indication that typical gastrointestinal drugs, such as proton-pump inhibitors, are effective.

**Chronic pelvic pain**
The term is used differently by gynaecologists and urologists. The former see it as different from FSS with primarily urological symptoms, whereas the latter use it as superordinate category with subcategories like interstitial cystitis and chronic prostatitis. In terms of treatment efficacy, the lack of effect of surgical procedures is as noteworthy as the weak to moderate effect of hormones and counselling with reassuring ultrasound scanning.

**Premenstrual syndrome, temporomandibular disorder, and environmental illness**
The therapeutic evidence as reported in systematic reviews is very limited for these syndromes. For temporomandibular joint disorder, there has recently been an upsurge of interest in the role of somatisation and its combination with organic joint findings.

**Chronic low-back pain**
The evidence for treatments of chronic low-back pain is broader than for most other syndromes, with nine different Cochrane reviews and a large systematic review initiated by the European commission attesting to this. Treatments with active involvement (in the form of psychotherapy or exercise or other) and drugs with a central rather than peripheral mode of action seem to be the most promising.

### Panel 2: Management recommendations for FSS

**Assessment**
- Think of the possibility of FSS in patients with enduring physical symptoms; do not equate them with malingering
- Be attentive to clues of the patient indicating bodily or emotional distress beyond the current lead symptom and outside your specialist field
- Assess functioning, patient expectations, and illness behaviour
- Avoid repetitive investigations only to calm the patient or yourself
- Decide whether patient has uncomplicated single FSS or complicated FSS. Are there bodily or mental symptoms clearly beyond single FSS? Is there excessive loss of functioning? Is there dysfunctional expectations or illness behaviour?

**Differential stepped care**

<table>
<thead>
<tr>
<th>Step 1a: uncomplicated FSS</th>
<th>Step 1b: complicated FSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reassurance with positive explanation of FSS; do not only convey negative test results</td>
<td>Measures as in step 1a for current main symptom</td>
</tr>
<tr>
<td>Symptomatic measures like pain relief</td>
<td>Consider antidepressant treatment</td>
</tr>
<tr>
<td>Advise graded activation or exercise rather than rest</td>
<td>Advise on dysfunctional attributions and illness behaviour and encourage reframing of symptoms within biopsychosocial framework (ie, incorporate both the patients’ beliefs about the organic nature of their symptoms and how these can be affected by a range of psychological and contextual factors)</td>
</tr>
<tr>
<td>If appropriate: appointments at regular intervals rather than patient-initiated</td>
<td>If appropriate: appointments at regular intervals rather than patient-initiated</td>
</tr>
</tbody>
</table>

**Step 2:** If either step 1a or step 1b prove to be insufficient
- Prepare referral to psychotherapist or mental-health specialist with reappointment
- Ensure that traumatic stressors and maintaining context factors, such as litigation, are assessed
- Continue with appointments at regular intervals rather than patient-initiated
- Liaise with psychotherapist or mental-health specialist on further treatment planning and difficulties

**Step 3:** If step 2 proves insufficient and if appropriate in your country
- Multidisciplinary treatment including symptomatic measures, activating physiotherapy, and psychotherapy
Review

Panel 3: Issues for further research on management of FSS

- Establishment of valid and generally accepted diagnostic criteria and procedures across medical specialties
- Extension of good quality evidence to poorly researched syndromes
- Extension of evidence base to doctor-centred and context-centred interventions
- Systematic consideration of overlap cases and psycho-behavioural features for trial design (including testing of effects outside specific focus of treatment)²⁴
- Determination of relevant outcome criteria (symptoms vs functioning)²³⁻²⁵
- Investigation of variability of placebo response²⁶⁻²⁸
- Investigation of differential effects of treatments in syndrome-specific clinics versus generic treatments for all FSS in specialised consultation-liaison clinics
- Investigation of differential treatment effects in primary versus secondary care²⁹

most effective, whereas there is little to no evidence that many different passive physical interventions (including invasive ones) are effective.

Diagnostic analogues

Much less research has been done on the treatment of patients described with the diagnostic analogues of FSS used primarily in psychological medicine, such as somatoform disorders, multiple unexplained physical symptoms, conversion disorder (hypocondriasis is excluded because health anxiety exceeds bodily symptoms).³⁷ Most research into these analogues involves psychotherapy and psychopharmacotherapy (there is no systematic review for the latter in the period covered by this review).

Though limited in extent, studies that describe patient-centred effects of interventions aimed at the doctor highlight the importance of adequate doctor–patient interaction in the treatment of patients with FSS.³⁸ Most of the work with educational programmes aimed at doctors who treat patients with FSS has been done in primary care. The most noteworthy of these educational interventions is the so-called TERM (the extended reattribution and management) model for the assessment and treatment of functional disorders in general practice,³⁹ with first results of a cluster randomised trial showing that it favourably changes the attitudes of general practitioners towards patients with unexplained or functional symptoms.³⁹

Management of functional somatic syndromes as a whole

The effect of non-pharmacological passive treatments, be they invasive or non-invasive, seems to be weaker than the effect of non-pharmacological treatments that involve active patients’ cooperation. The evidence for efficacy of the latter across different FSS, in particular graded exercise and psychotherapy, underlines the importance of common factors in FSS, as the therapeutic rationale of these treatments typically aims at overall function and not at the alleviation of specific symptoms. The effect of pharmacological treatments that primarily aim at peripheral physiological disturbances linked to the different FSS is more variable across different FSS, with best effects shown for functional gastrointestinal syndromes, but little effect in many other FSS. In contrast, the moderate to good effect of antidepressive treatments seems to be distributed more evenly among the different FSS.

Taken together, the current state of evidence appears to support a balanced approach to the management of FSS, with organ-oriented and cognitive interpersonal treatment foci each having their merit in the treatment of single syndromes. The clinical importance of cognitive interpersonal approaches is particularly apparent for the treatment of overlap and other complicated cases, but there is still a lack of evidence to support this claim (panels 2 and 3).

On the basis of the balanced approach model and the evidence presented above, we make several recommendations for diagnosis and treatment of FSS in primary and secondary care (panel 2).¹³ The essence of these recommendations is: to convey to the patient that his symptoms are real, to offer positive advice and treatment and to engage the patient in an active role in alleviating the often chronic symptoms.

Conclusion

The evidence base for the management of FSS has grown considerably over the last years and many patients are served well with the treatments applied in different subspecialties of medicine. However, there is still a predominance of a splitting view on single FSS and a lack of diagnostic criteria and classifications that are valid and generally agreed upon. This not only leads to confusion, but also to the neglect of important therapeutic options for many patients. Stigmatisation of psychosocial factors that are relevant in FSS and the mistaken tendency to view patients with FSS as the worried well further enhance this difficulty.

The future classification of FSS should reflect the need for a balance between biomedical and more integrative approaches. Training of medical students and doctors requires systematic integration of problem-specific communication skills into the training curricula of general medicine and integration of symptom-focussed organ-oriented know-how into the training curricula of psychological medicine. Health care and reimbursement systems have to adapt to the objective needs of doctors and patients at the psychosomatic interface (eg, by testing the value of specialised clinical services for it). Last but not least, practitioners in different specialties of medicine should discover the importance and the rewards of caring for patients with FSS in a balanced way, rather than purely biomedical or psychological approaches.

The management of FSS will most likely continue to be an important health-care focus on all levels of care. Its future outlook will reflect the state of integration, the different fields of general and specialist somatic as well as psychological medicine will achieve. Competent communication, between the fields of medicine as well as between doctors and patients, will be a mainstay in this endeavour.

Contributors
All authors contributed equally to the preparation and writing of this review.

Conflict of interest statement
We declare that we have no conflict of interest.

Acknowledgments
We are grateful to Heribert Sattel for help with the literature search and to Bernd Löwe for valuable support. We are also indebted to Paul Enck and three anonymous reviewers for comments on an earlier draft. Supported in part by DFG (PISO) and BMBF (Funktional).

References
Interpreting health statistics for policymaking: the story behind the headlines

Neff Walker, Jennifer Bryce, Robert E Black

Politicians, policymakers, and public-health professionals make complex decisions on the basis of estimates of disease burden from different sources, many of which are "marketed" by skilled advocates. To help people who rely on such statistics make more informed decisions, we explain how health estimates are developed, and offer basic guidance on how to assess and interpret them. We describe the different levels of estimates used to quantify disease burden and its correlates; understanding how closely linked a type of statistic is to disease and death rates is crucial in designing health policies and programmes. We also suggest questions that people using such statistics should ask and offer tips to help separate advocacy from evidence-based positions. Global health agencies have a key role in communicating robust estimates of disease, as do policymakers at national and subnational levels where key public-health decisions are made. A common framework and standardised methods, building on the work of Child Health Epidemiology Reference Group (CHERG) and others, are urgently needed.

“Hunger kills 6 million children a year”. Such public-health estimates are often headline news and can affect decisions about how resources are allocated to tackle health problems. Many readers, even with access to the scientific report behind the headline, are ill-equipped to assess the quality of the information provided, or to understand how the estimate underlying the headline was arrived at (in this example, undernutrition was just one of many causes that killed half of 11 million children in 2000). Politicians, policymakers, and public-health professionals are increasingly faced with making complex decisions that require the interpretation of estimates of disease burden. The quality of their decisions reflects their ability to assess and interpret this quantitative information, which often comes from different sources and is marketed by skilled advocates.

It is time to revisit basic issues related to interpreting estimates of public-health indices. In the first of the Lancet Health Statistics series, Boerma and Stansfield described the increasing demand for, and complexity of, public-health statistics and called for a rationalisation of measurement strategies. In the next article, Murray discussed the poor availability of health statistics for the Millennium Development Goals health indicators, and recommended explicit and comprehensive data audit trials. In this third article of the series, we discuss different types of global public-health estimates, and offer basic guidance to those who use them in decisionmaking about how to assess and interpret them. The final article will look at enhancing the use of statistics about disease burden from different sources, many of which are “marketed” by skilled advocates. To help people who rely on such statistics make more informed decisions, we explain how health estimates are developed, and offer basic guidance on how to assess and interpret them. We describe the different levels of estimates used to quantify disease burden and its correlates; understanding how closely linked a type of statistic is to disease and death rates is crucial in designing health policies and programmes. We also suggest questions that people using such statistics should ask and offer tips to help separate advocacy from evidence-based positions. Global health agencies have a key role in communicating robust estimates of disease, as do policymakers at national and subnational levels where key public-health decisions are made. A common framework and standardised methods, building on the work of Child Health Epidemiology Reference Group (CHERG) and others, are urgently needed.

Levels of estimation

In public health, different levels of estimate correspond to different relations to a disease or other cause of illness or death. Panel 1 shows the hierarchy of these levels, based on the relation of the indices to mortality or morbidity. Understanding how closely linked a type of statistic is to disease and death rates is crucial in designing health policies and programmes.

Broad determinants of health status

At the broadest level, poverty determines both the risk and adequacy of response to infectious diseases,

Panel 1: Levels of statistics about disease burden

1. Broad determinants of health status, including socioeconomic status
2. Risk factors, such as smoking, environmental exposures, diet, or genetic predispositions
3. Underlying causes, the most recognised of which is undernutrition
4. Direct causes of mortality or morbidity—such as infectious diseases, non-communicable diseases, or injury
5. Indirect causes, whereby one health condition predisposes to another
6. Future disease burden, such as can be anticipated from future infection prevalence
7. Costs and economic consequences of health status, including direct and indirect costs to families for illness care, being orphaned, economic productivity, costs of interventions, costs saved by intervention, and cost-effectiveness
nutritional deficiencies, and other disorders. Although this is so self-evident that estimation of the size of the effect is hardly necessary, it is inappropriate to conclude, as some have, that the way to reduce the disease rates in low-income countries is to simply wait for economic development.\textsuperscript{5,6} The estimation of determinants and causes of disease at lower levels can contribute to the selection of policies and programmes that can reduce disease burden in the shorter term while poverty reduction strategies are underway. If these policies and programmes are designed to reduce inequities within countries, they can have the greatest effect on the poorer segments of the population that have the highest disease burden, so estimation at this level can be useful.

**Risk factors and underlying causes**

Estimates of disease burden caused by risk factors such as tobacco use or sexual behaviour, or underlying causes of disease such as nutritional deficiencies, can be helpful in the identification of preventive interventions with large and long-lasting benefits. On the other hand, inappropriate aggregation of risk factors, for example, attempting to assess the contribution of all “environmental” risk factors to a health condition,\textsuperscript{7} can yield large attributable fractions of global disease burden, but are not useful in prioritising interventions to address the particular environmental exposures (or deficiencies). Traditionally, policymakers have been slow to recognise the importance of risk factors and the cost-effectiveness of preventive interventions despite these having greater benefits than therapeutic or palliative interventions in addressing the direct causes of death. The rapid scale-up of antiretroviral therapy—rather than prevention programmes—for people with HIV/AIDS in developing countries is a good example. Whether the momentum generated by the US President’s Emergency Plan for AIDS Relief (PEPFAR), WHO, and others in promoting prevention will be matched by real investment remains to be seen.

**Direct causes**

The most commonly used estimates of disease burden are measures of death, illness, and disability. Whether these are used as direct or summary measures, such as disability-adjusted life years (DALYs), their contribution is crucial in establishing priorities in disease control. But such statistics can be used in misleading ways. Estimation of the proportions of child deaths from specific diseases at regional and national levels can be important when organising priorities, but can be confusing if used out of context. For example, global advocacy statements that HIV/AIDS is rolling back generations of gains in child survival in Africa\textsuperscript{8} might be true for a few countries, but is misleading because HIV/AIDS is responsible for less than 10% of child deaths in sub-Saharan Africa.\textsuperscript{9} The use of specific target populations as a context for estimates can also be misleading. Statements such as “Measles is the biggest killer among the vaccine-preventable diseases”,\textsuperscript{10,11} for example, is technically correct, but diverts attention from the fact that measles contributes to very few of the total number of infectious disease deaths in children, even in areas where it remains endemic.\textsuperscript{11}

**Indirect causes**

Clinical observations and the results of placebo-controlled trials have shown that some diseases can not only kill directly, but can also have indirect effects through non-fatal illness that increase the risk of death from other infectious diseases. But not all estimates of disease burden include these indirect effects. The indirect contribution of malaria to childhood deaths from other communicable diseases and anaemia, for example, is supported by the finding that some malaria prevention trials have shown a greater reduction in deaths than can be attributed directly to malaria infection.\textsuperscript{12} These results could have been due to misclassification of the cause of death by the verbal autopsy methods used, or because non-fatal malaria predisposes to subsequent pneumonitis or other fatal infections. Quantifying this possible indirect effect is difficult, but efforts to prioritise among disease interventions must take into account the possible indirect effects of many infectious diseases—eg, the contribution of diarrhoea to nutritional deficiencies that place children at higher risk of death from subsequent infections,\textsuperscript{13,14} or the fact that infection with HIV predisposes to other diseases such as tuberculosis.\textsuperscript{15} An awareness of these indirect causes of disease and death, and the hierarchies used to assign a single cause of death, is essential in assessing and interpreting burden estimates correctly.\textsuperscript{16}

**Future disease burden**

Estimates of future disease burden from emerging and potentially epidemic diseases are particularly difficult. Some infections, such as Ebola haemorrhagic fever, receive undue attention in the media and trigger large public-health preparedness efforts, perhaps because of the horrific form of death in many of those infected, despite the very low risk to the rest of the population. Objective estimation of the true risk of this and similar rare diseases would be useful even if imprecise. It can be almost impossible to predict future risk from viruses such as influenza virus H5N1 because an estimate requires speculation about microbial evolution, changes in transmission potential, and human behaviour. Avian influenza has spread from southeast Asia to other parts of the world, killing a small number of people who have had direct contact with infected birds. But if this virus were to mutate into one that could spread from people to people, the numbers could rapidly increase. WHO has estimated that the potential human deaths, if the
virus spreads more widely, range from 2 to 100 million. Other analysts say even the high end of this range is too conservative and that a pandemic could kill more than one billion people. What is the prudent action against a disease that could kill none or just as plausibly kill one billion?

Costs and economic consequences

Finally, estimates of the costs of existing and future disease can be useful in public-health decisionmaking. This is particularly true of costs at the household or community level, such as figures suggesting that the direct and indirect cost of severe illness to families can result in a descent into poverty. At the societal level, these estimates are often subject to large assumptions about the consequences of ill health. For example, iron-deficiency anaemia can slow the body’s development, including that of the brain, which could have important economic consequences for society, yet iron supplementation trials show a benefit of only 1–2 IQ points. The effect of this benefit on productivity in largely agrarian societies is unknown.

Five questions decisionmakers should ask about estimates

A complete methodological discussion of estimates in public health is beyond the scope of this paper. Instead, we suggest questions to help users of disease burden estimates assess and understand their strengths and limitations.

What metric is used in the statistic?

Absolute numbers, rates, or ratios have different meanings. How an estimate is presented can make an enormous difference in its interpretation. Cumulative measures and lifetime risks are particularly hard to understand. Take the example of HIV/AIDS. Advocates of funding for this disease often quote the cumulative number of global deaths from HIV/AIDS since it was first identified. But, if historical estimates were used for other diseases, the number of HIV/AIDS deaths would be small in comparison. For example, if the same statistical procedures were applied for pneumonia as for HIV/AIDS, the cumulative deaths since 1975 would be about 60 million—almost three times the estimated cumulative deaths from AIDS in the same time period.

Even estimates of prevalence and incidence of a disease can be misleading. For example, the mean survival time of an adult with HIV/AIDS, without treatment, is 9–10 years. HIV prevalence is a period measure that stretches over that time. Early in an epidemic, the number of people living with HIV or AIDS and the adult prevalence rises rapidly, as new people are infected and few die. Later these numbers and the rate of increase change as mortality rates begin to rise. By selectively switching between the rate of increase in new epidemics, the number of people living with infection in an older epidemic, and the cumulative number of deaths, one can focus attention selectively on the statistic that most strongly supports the importance of the disease. All the facts are correct, but their juxtaposition is often orchestrated to make the strongest possible impression of importance on the consumer without valid comparisons with other health conditions.

Policymakers and the public often have difficulty in understanding rates. Ratios tend to be more effective in communicating advocacy messages accurately. Statements like “Maternal mortality is 100-fold higher in many low-income countries than in high-income countries” sends a clear message with respect to inequities, but no information about the absolute magnitude of the problem. Statements more useful to decisionmakers are those that use a standard metric to provide sets of meaningful comparisons. For example, the ratio of inequity between low-income and high-income countries for deaths from severe neonatal infections is far lower, at 11-fold. In absolute numbers, however, two to three times as many lives are lost to neonatal infections each year (1·4 million) in developing countries than to maternal mortality (500 000).

Finally, users of health statistics must distinguish between crude, corrected, and predicted statistics, which is addressed in another article in this Series.

Is the statistic a valid measure of population health?

Users of public-health estimates often assume that a measure is a direct or indirect reflection of disease burden or future health risk. This is not always the case. For example, the 2003 WHO report on tuberculosis presents portraits of two young people—one from Italy and one from Russia—accompanied by the statement “One-third of the world is infected with tuberculosis—that’s almost 2 billion people.” Not stated is that this estimate is based on positive skin tests indicating latent rather than active infections; in the USA, for example, only about 5% of individuals with latent infections will develop active tuberculosis in the first year after infection, and an additional 5% will develop clinical disease later in life.

How good are the input data and estimation methods?

Few public-health decisionmakers are provided with adequate information or equipped with the technical skills needed to assess the quality of an estimate. Not only are most current estimates based on few data, but also these data may not be representative of other populations and are widely variable in quality. There are strong biases toward locations where data collection is possible and funding for field research is available. Many countries have adequate data on key public health topics (eg, HIV estimates based on sentinel surveillance...
sites in 137 out of 157 larger population countries but accessing these data requires a lot of time and resources. For most public-health issues there are few valid country-specific data, and a large proportion of country estimates are generated through extrapolation and modelling techniques.

Not only are adequate epidemiological data scarce, support for new data collection has decreased to nearly zero over the past 20 years in some fields such as child health. Unless major new epidemiological data-gathering efforts are begun soon, future estimates, for example those to assess progress toward the Millennium Development Goals, will be based on fewer data than those available today.

Not only that, but the methods used to arrive at most disease burden estimates had not been available for public or peer review. The table illustrates this point by summarising how well global mortality estimates for specific conditions fare when assessed with five basic quality criteria: (1) whether the estimate was reviewed by an independent technical group; (2) whether the estimation methods and input data have been published in a peer-reviewed scientific journal; (3) whether the data underlying the estimates are available for public review; (4) whether tools and software have been developed and are available for review and application; and (5) whether the estimation process was done at country level, where the validity of assumptions and the appropriateness of input data are likely to be best.

The wide variability in methodological transparency across burden estimates reflects an absence of standard procedures for developing, reporting on, and supporting use of disease burden estimates across the various domains of public health. Until such standards are developed and widely used, the onus of assessing the validity of disease burden estimates rests almost entirely with health policymakers.

Do contextual factors have a role?

Users of quantitative estimates must learn to examine every estimate and to decode the considerable communications artistry that goes into their construction. For example, a common practice is to present an estimate at the global or regional level and then to elaborate on it by giving a specific and often unrepresentative example. HIV/AIDS advocates talking about the effect of AIDS on under-5 mortality often use as examples countries in southern Africa where AIDS accounts for 30–50% of deaths in under-5s. But for sub-Saharan Africa as a whole, AIDS is thought to account for less than 10% of under-5 deaths.

A second way that estimates are “spun” by HIV/AIDS spokespersons and advocates is to use worst-case scenarios in projecting the effect of an epidemic. For example, several public-health leaders say that unless we act quickly, the HIV epidemic in Asia will soon look like the epidemic in sub-Saharan Africa. This projection is apparently based on the assumption that all epidemics have the same trajectory. In fact, evidence available for some time suggests that the characteristics and evolution of the HIV epidemic in India and other Asian countries are fundamentally different from those in most of Africa. Given similarities in measures of socioeconomic development, the future of Asia’s HIV epidemic is more likely to resemble that of South America where the epidemic started much earlier but where adult prevalence rates have hovered around 0·6% for the past 6 years. Using worst-case scenarios to

<table>
<thead>
<tr>
<th>Primary organisations creating estimates and models</th>
<th>Estimate reviewed by independent technical reference group? (external to UN)</th>
<th>Methods and input data published in peer-reviewed scientific journal?</th>
<th>Input database available for public review?</th>
<th>Estimation software or tools available for review and application?</th>
<th>Methods for estimating uncertainty or range for the estimate?</th>
<th>Estimate developed at country level?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuberculosis WHO</td>
<td>One-time panel</td>
<td>1999–200331–33</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cause-specific under-5 mortality (multiple causes) WHO, LSHTM CHERG</td>
<td>200635–36</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Pneumonia WHO</td>
<td>CHERG</td>
<td>200237</td>
<td>Yes</td>
<td>No</td>
<td>Yes38–39</td>
<td>No</td>
</tr>
<tr>
<td>Diarrhoea WHO</td>
<td>CHERG</td>
<td>In preparation</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Malaria CDC, WHO</td>
<td>Malara M&amp;E Reference Group (MERG), CHERG</td>
<td>200637</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Measles WHO</td>
<td>CHERG &amp; EPI/WHO</td>
<td>200338</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Neonatal causes WHO</td>
<td>CHERG</td>
<td>200540–41</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Maternal mortality WHO</td>
<td>Ad hoc panel</td>
<td>1999, 2001</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Nutrition as an underlying cause of child mortality JHU, WHO</td>
<td>CHERG</td>
<td>1993, 200238–39</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

LSHTM=London School of Hygiene and Tropical Medicine. UNFPA=United Nations Population Fund. JHU=Johns Hopkins University.

Table: Basic characteristics of mortality estimates in public health
create a sense of urgency around a disease is not unique to HIV/AIDS, and users of estimates should be alert to this practice.

A third example of the influence of context on estimates stems from the competition for disease recognition, priority in policymaking, and, most importantly, funding. In the absence of clear standards for assessing and reporting disease burden estimates, advocates for specific causes understandably choose methods that have the best chance of keeping a particular disease at the top of the health priority list. The Centers for Disease Control and Prevention has estimated that obesity kills 400,000 each year in the USA. This statement resulted in a re-evaluation of the deaths attributable to tobacco use, driven in part by concern among advocates that resources might be reallocated from tobacco control to obesity prevention. Adapting their estimation methods to match those used for obesity, the estimated number of deaths attributable to tobacco use rose from 435,000 to 640,000. Individuals who compare statistics must attempt to determine whether comparable methods were used for all the estimates, and public-health leaders must advocate for the development and use of normative standards.

How is uncertainty addressed?
Estimates are, by definition, imprecise. Determinants of the extent of imprecision include the volume of data (in theory, more data mean less uncertainty), the number of steps between data gathering and disease burden estimation (more methodological complexity often means more uncertainty), and potential biases in data gathering or analysis. Often, raw data are used directly to calculate the metric in which the estimate is presented. But each step in the process of transforming raw data to final estimates has associated uncertainty, and as assumptions are combined, the error or uncertainty around the estimate of the final index increases. An approach for estimating measles mortality is an example. First, there is uncertainty in the estimate of vaccination coverage, which is based on country-reports. Second, there is an assumption that all unvaccinated children will develop measles if exposed to the pathogen. Finally, a case-fatality rate (CFR) is estimated and applied to predict the number of new infections, resulting in an estimate of deaths due to measles. The uncertainty associated with this estimate is a combination of the errors in vaccination coverage, assumptions about exposure and illness, and the error associated with the estimated CFR. Estimates that do not specify the level of uncertainty, either with ranges or confidence bounds, should always be questioned. Those who generate disease estimates have a responsibility to be transparent about the process used and should provide information about the precision of their estimate, even if this is only a range of uncertainty.

Of course, simply providing a range for an estimate does not guarantee that it is correct. There are better and worse ways to construct and use uncertainty ranges. An article by Jeffrey Sachs in the New England Journal of Medicine, calling for better malaria control is a good example. Sachs begins the article by stating “Malaria currently kills up to 3 million people per year worldwide, most of them children in sub-Saharan Africa”. What he does not mention is that the best estimate for deaths directly caused by malaria is 1·2 million, with 3 million being the upper limit estimated using all worst-case assumptions and including possible indirect causes that are normally ascribed to other diseases. This failure to provide the full range in the estimate, and the selective use of only the upper end of the range are clear signs that the article is trying to exaggerate the importance of the problem.

Policy makers are often asked to make decisions about competing priorities based on estimates (eg, this disease kills more than that disease; the rate of increase of this disease means that it will have a higher burden than another in 5 years so more money should be spent) when in fact the uncertainty around the estimates suggests that there are no meaningful differences among them.

Conclusion
Estimates of disease burden, along with considerations of feasibility and cost, should be central to decisions about public-health interventions. Even if the role of evidence in public-health decision-making is balanced by strong political, social, and other contextual factors, estimates provided to decision-makers should be based on sound and transparent methods, applied in comprehensive and systematic ways to the various levels of disease burden assessment, and to diseases and conditions within levels. Our intention is to promote healthy scepticism of health statistics, not cynicism.

Estimates would be more credible if they came from technical groups that are independent of the organisations that implement programmes and advocate for funds. This effort has begun over the past few years in groups like the Child Health Epidemiology Reference Group (CHERG) and the UNAIDS Reference Group on Estimates, Projections and Modelling, through which independent technical experts advised the WHO, UNICEF, and UNAIDS on estimates of disease burden related to pneumonia, diarrhoea, malaria, measles, neonatal illness, undernutrition, and HIV/AIDS. The recent formation of WHO’s Health Metrics Network is a promising means for improving national capacity in this area, providing more and better building-blocks for future estimates of disease burden.

In panel 2, we match news headlines with the “full story”, reflecting the basic rules of communication we
have proposed for the presentation of public-health estimates. Clearly, the full-story versions are not going to make the front page: they are too long and too nuanced for a news story. Thus, modifying the practices of those who develop estimates and prepare technical reports will not be enough to ensure well-informed policy decisions. Not only that, but news reporters will always be pressured to write attention-grabbing headlines, and not all readers will be able to—or need to—process complex scientific information. What is needed is a combination of effective communication of estimates, combined with efforts to improve the skills of policymakers in interpreting them. Good public-health decisions will always be a shared responsibility between those who generate data and estimates, and those who use them to make decisions. In panel 3, we present some tips to discern real evidence. But this is only a starting point. Academic institutions and those working in professional development fields can go further to help key public-health figures build practical skills in assessing the quality of estimates and in their interpretation.

Our comments on specific examples should not be interpreted as attacks on individuals or organisations that have developed estimates or used them to advocate for specific actions. In fact, the diseases from which we have drawn examples are in general those for which there are efforts underway to educate and advocate about particular diseases or programmes. However, policymakers and health-care professionals are asked to prioritise some disease programmes over others, and need a broader view than that presented by disease-specific groups. Objective and methodologically sound estimation of disease burden should be a priority of UN agencies, such as WHO and UNICEF, who are responsible for global health, in collaboration with bilateral donors and international finance institutions. Such efforts must also be applied at national and subnational levels because this is where priorities need to be set and thus need to be informed by the best evidence. A common framework and standardised methods, building on the work of CHERG and other groups, are urgently needed. Recent efforts by WHO to establish common ground rules for estimation are important steps in the right direction."

Conflict of interest statement
We declare that we have no conflict of interest.

References


58 Osterholm M. In a presentation at “The increasing health risks in our globalized world—a conference for journalists.” Transcript provided by http://kaisernetwork.org.

A 78-year-old man presented to our emergency department in October, 2005, with a 6 h history of severe retrosternal, non-radiating precordial pain, described as “someone gripping the inside of my chest”. The pain was episodic, lasting for seconds, followed by complete relief for 3–5 min. His medical history included ischaemic heart disease, chronic obstructive pulmonary disease, and osteoarthritis. Since coronary-artery grafting 3 years previously he had had no further angina. Medications were isosorbide mononitrate, simvastatin, aspirin, and omeprazole. Clinical examination was unremarkable. ECG showed T-wave inversion in leads aVL, and V2 to V5. Troponin T was normal (<0·03 ng/mL). His full blood count and coagulation screen were within normal limits. Chest radiography was normal. An acute coronary syndrome was diagnosed, and treatment with subcutaneous low-molecular-weight heparin (1 mg/kg twice daily) was given.

The following day he developed vomiting, odynophagia, and dysphagia. Upper gastrointestinal endoscopy showed a bluish mass confined to the posterior wall of the oesophagus that extended from 17 cm to 38 cm. The mass almost completely obstructed the lumen (figure A). No mucosal tear was seen. An oesophageal haematoma was suspected. CT showed a diffuse non-obstructing soft-tissue mass intrinsically related to the oesophagus, thought to be a haematoma (figure B). There was no abnormality extrinsic to the oesophagus. The patient was managed conservatively, and aspirin was stopped. He tolerated a soft diet initially and was discharged when eating a normal diet. Endoscopy and CT were repeated 3 months later and showed complete resolution of the oesophageal abnormalities. Final review was in January, 2006, and the patient was completely asymptomatic.

Spontaneous intramural haematoma of the oesophagus was first reported in 1970 and is an uncommon condition. Chest pain is the most common presenting symptom; it may be described as “tearing” and generally does not radiate. Other symptoms can include odynophagia, dysphagia, and small haematemesis. In a study of 91 case reports, only 35% of patients showed the typical clinical triad of acute retrosternal pain, odynophagia or dysphagia, and haematemesis. Reported risk factors include increasing age, the use of anticoagulants, thrombolytics, or antiplatelet drugs, and bleeding disorders. Physical examination is normal in most cases. Diagnosis can be difficult, because symptoms are similar to those of more common oesophageal conditions or ischaemic heart disease. Anticoagulants may be used inadvertently when spontaneous oesophageal haematoma is mistaken for an acute coronary syndrome. In reported cases in which low-molecular-weight heparin has inadvertently been given, this treatment has been judged to exacerbate bleeding into the oesophageal wall. The characteristic appearance at endoscopy is of a bluish submucosal lesion, commonly on the posterior wall of the oesophagus, terminating abruptly above the squamocolumnar junction. Endoscopy does not seem to increase the risk of complications in this setting.

Prognosis is excellent, with spontaneous resolution of the haematoma in most patients. Complications of spontaneous oesophageal haematoma include oesophageal dissection, rupture, obstruction, and upper gastrointestinal bleeding. Conservative management is recommended, with patients receiving nil by mouth while pain continues, and gradual introduction of a soft diet. Our case clearly illustrates a condition that, while rare, is useful to bring to the attention of a general medical readership. The history of dysphagia and odynophagia provided clues to oesophageal pathology in our patient. Chest pain was initially incorrectly diagnosed as angina despite atypical features. The administration of low-molecular-weight heparin was likely to have been detrimental to this patient. Although rare, dissecting intramural haematoma of the oesophagus is a cause of chest pain that must be considered and distinguished from that of acute coronary syndrome, particularly because the administration of anticoagulants is undesirable.

References