1. International Health Partnership: a welcome initiative
   Page 801
   The Lancet

2. New guidelines for better asthma control
   Page 802
   The Lancet

3. Eliminating human trafficking: a new arm of HIV prevention
   Page 802
   The Lancet

4. The International Health Partnership
   Pages 803-804
   Douglas Alexander

5. Vascular outcome in type 2 diabetes: an ADVANCE?
   Pages 804-805
   Norman M Kaplan

6. Launching a new movement for mental health
   Page 806
   Richard Horton

7. Most people with mental illness are not treated
   Pages 807-808
   Graham Thornicroft

8. Depression is very disabling
   Pages 808-809
   Gavin Andrews and Nickolai Titov
9. Stigma and mental health
   Pages 810-811
   Norman Sartorius

10. Royal Society of Tropical Medicine and Hygiene: 100 years old
    Pages 811-813
    Brian Greenwood

11. Can drug regimens be adapted to patients, or vice versa?
    Pages 813-814
    Peter AGM De Smet, Cees Kramers and Nicky Britten

12. Earthquake highlights mental-health issues in Peru
    Pages 815-816
    Barbara Fraser

13. Shortage of psychiatrists a problem in Pakistan
    Pages 817-818
    Khabir Ahmad

14. Psychiatry without context: turning sadness into disease
    Pages 819-820
    Arthur Kleinman

15. Teen depression
    Page 820
    Patrick McGorry

16. Vikram Patel: promoting mental health in developing countries
    Page 821
    Stephen Pincock

17. Pax Allan Chingawale
    Page 822
    Sumeet Sodhi

18. Biko to Guantanamo: 30 years of medical involvement in torture
    Page 823
    David J Nicholl, Trefor Jenkins, Steven H Miles, William Hopkins, Adnan Siddiqui and Frank Boulton

19. ICRC and confidentiality
    Pages 823-824
    Yves Daccord
20. Use of evidence in WHO recommendations
   Page 824
   Eamonn MM Quigley, Michael Fried and Justus H Krabshuis

21. Use of evidence in WHO recommendations
   Pages 824-825
   Jonathon Simon and Donald M Thea

22. Use of evidence in WHO recommendations
   Page 825
   Herbert B Peterson, Kathryn M Curtis, Anna Glasier and Kerstin Hagenfeldt

23. Use of evidence in WHO recommendations
   Pages 825-826
   Julian Kelly, Harry Campbell, Naor Bar-Zeev, Giorgio Tamburlini and Trevor Duke

24. Use of evidence in WHO recommendations
   Page 826
   Julian H Elliott and Russell L Gruen

25. Use of evidence in WHO recommendations – Authors' reply
   Pages 826-827
   Andrew D Oxman, John N Lavis and Atle Fretheim

26. Right to health care for vulnerable migrants
   Page 827
   Ike Anya

27. Right to health care for vulnerable migrants
   Pages 827-828
   Antonio Virgilio, Loris Defilippi, Konstantinos Moschochoritis and Raffaella Ravinetto

28. Right to health care for vulnerable migrants
   Page 828
   Joost J den Otter and Daniel Tavenier

29. Department of Error
   Page 828
30. Effects of a fixed combination of perindopril and indapamide on macrovascular and microvascular outcomes in patients with type 2 diabetes mellitus (the ADVANCE trial): a randomised controlled trial
   Pages 829-840
   Anushka Patel

31. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys
   Pages 841-850

32. Depression, chronic diseases, and decrements in health: results from the World Health Surveys
   Pages 851-858
   Saba Moussavi, Somnath Chatterji, Emese Verdes, Ajay Tandon, Vikram Patel and Bedirhan Ustun

33. No health without mental health
   Pages 859-877
   Martin Prince, Vikram Patel, Shekhar Saxena, Mario Maj, Joanna Maselko, Michael R Phillips and Atif Rahman

34. Resources for mental health: scarcity, inequity, and inefficiency
   Pages 878-889
   Shekhar Saxena, Graham Thornicroft, Martin Knapp and Harvey Whiteford

35. Human papillomavirus and cervical cancer
   Pages 890-907
   Mark Schiffman, Philip E Castle, Jose Jeronimo, Ana C Rodriguez and Sholom Wacholder

36. An unexpected cause of fever and seizures
   Page 908
   Dhanpat Kumar Kochar, Deepak Pakalapati, Sanjay Kumar Kochar, Parmendra Sirohi, Mahesh Pal Khatri, Abhishek Kochar and Ashis Das
International Health Partnership: a welcome initiative

On Sept 5, the UK Prime Minister Gordon Brown launched the International Health Partnership—a global “compact” for achieving the Health Millennium Development Goals—at a prestigious gathering at number 10 Downing Street.

The support for this initiative is impressive: Ministers of Health from several developing countries in Africa and Asia, Ministers and Prime Ministers from donor countries, heads of, or high ranking officials from, international institutions such as WHO, The Global Fund, and the World Bank, and other major donors, such as the Bill and Melinda Gates Foundation, were all signatories to this partnership. But what does the International Health Partnership (IHP) mean for people living in poorer countries?

The IHP is an agreement between donors and developing countries. Global and country level partnerships will set out a process of mutual responsibility and accountability for the development and implementation of the national health plans of developing countries. The overall aim of the IHP is to improve the coverage and use of health services—whether through public or private channels, or through non-governmental organisations—in order to deliver improved health outcomes, especially for the health-related MDGs, and other international commitments such as universal access to antiretroviral therapy. The IHP does not provide any new funding.

In an accompanying Comment, Douglas Alexander, the UK Secretary of State for International Development, explains that the three specific aims of the IHP are to better coordinate donor aid, to widen the focus of donor aid so that it also includes health systems, and to help support the national health plans of developing countries.

Donors who have signed up to the partnership will be expected to better coordinate external support to help develop and implement comprehensive national health plans, provide aid in ways that strengthen health systems, and provide more long-term, flexible support though national systems. In return, partner developing countries will invest further in their own health systems, strengthen planning and accountability mechanisms, and also better link external support to improvements in health outcomes.

This concept may sound quite familiar. Over 2 years ago, through the Paris Declaration on Aid Effectiveness, over one hundred Ministers and heads of agencies committed their countries and organisations to improve aid harmonisation and to better monitor indicators of aid effectiveness. The IHP is putting this Paris Declaration into practice in the health sector. The signatories to the partnership will evaluate progress on an annual basis.

The implementation of the IHP indicates that donors and developing countries have done much more than merely listen to the long list of complaints charged at one another. For example, developing countries have repeatedly stated that strengthening health systems is the most important factor in delivering sustainable results. Donors have often voiced concerns about the capacity of national governments to implement their health plans. The IHP should help to address these criticisms. But to do this most effectively, it is imperative that there is an equal share of power, responsibility, and accountability between all parties.

The language of the IHP says all the right things. National governments are to lead in formulating their own health plans, ending the reign of donor dictatorship. International and bilateral donors are to work to ensure that disease specific approaches and mechanisms to achieve broad health system strengthening are mutually reinforcing, rather than opposing each other. Donors will also test and evaluate ways to link their support to achieving results at a country level, including strengthening health systems. And since future funding is likely to be linked to performance, research into finding robust outcome measures to comprehensively evaluate the performance of health systems should go straight to the top of the priority list.

It is also encouraging that the IHP recognises the crucial role of civil society in helping governments to formulate national health plans, and also in holding governments and donors to account. This, of course, relies on there being strong, knowledgeable, and inclusive civil society groups within partnership countries, which have the freedom and confidence to speak out without reproach.

There is much promise in the International Health Partnership. Its future success or failure will depend on the details of putting the welcome words of this initiative into practice. ■ The Lancet
New guidelines for better asthma control

Last week, the US National Asthma Education and Prevention Program issued the first comprehensive update of its clinical guidelines for the diagnosis and management of asthma in a decade. The 500-page document is rigorous and evidence-based. It integrates the latest scientific evidence into the four essential components of asthma care: assessment and monitoring, patients’ education, control of factors contributing to asthma severity, and drug treatment.

There is increasing evidence that asthma is a heterogeneous disease and differs among individuals and by age. The new focus on children in the guidelines is especially welcome because most experts believe that children are under-medicated with inhaled steroids. These guidelines point out the importance of full control of symptoms in children and the fact that despite potential side-effects, the treatment benefits far outweigh the risks. Additionally, there are now separate treatment recommendations for children aged 0–4 years, 5–11 years, and 12 years and older. The 5–11 year age group was added (earlier guidelines combined this group with adults) because of new evidence suggesting that children might respond differently from adults to asthma drugs.

The guidelines place a strong emphasis on monitoring asthma control. The new approach focuses on two related yet distinct aspects of the disease: the level of daily impairment that a patient is experiencing and the patient’s future risk for exacerbations, loss of lung function, and drug side-effects. This new distinction is important because it addresses the fact that some patients can have few day-to-day symptoms, yet still be at high risk of frequent exacerbations.

The importance of teaching patients to self-monitor and manage their asthma, and for physicians to do a more thorough assessment of their patient’s condition is highlighted. With appropriate medical care, well-informed and empowered patients can control their asthma and live full normal active lives. These guidelines will be invaluable for clinicians and patients alike.

Eliminating human trafficking: a new arm of HIV prevention

Two recent reports highlight an issue that HIV prevention programmes can no longer afford to ignore. On Aug 22, the UN Development Programme (UNDP) released Human trafficking and HIV: exploring vulnerabilities and responses in south Asia. And, in the Aug 1 issue of the Journal of the American Medical Association, an article shows that repatriated Nepalese girls, who had been trafficked for sex, had a high prevalence of HIV. Overall, 38% of the 257 girls in the study tested positive for HIV. This figure increased to 60% when the analysis was confined to the 33 girls trafficked before the age of 15 years.

Similarly, the UNDP report shows that people who are trafficked are highly vulnerable to the risk of HIV infection. The report explores the link between trafficking and HIV in six countries of south Asia: Afghanistan, Bangladesh, India, Nepal, Pakistan, and Sri Lanka. In all the countries surveyed, the researchers found that there was little effort to address HIV and trafficking in an integrated way.

This situation is due in part to the fact that the link between the two issues is fairly new, complex, and under-researched. However, as the UNDP report shows, human trafficking and HIV infection share many common underlying factors, such as poverty and gender discrimination. And, although the percentage of HIV transmissions that can be attributed to trafficking is unknown, the coercion of thousands of women and girls into unprotected sex with multiple partners is highly likely to be a contributor to the spread of HIV/AIDS.

People who are trafficked for purposes other than sex might also find themselves in situations that increase their vulnerability to HIV infection. However, a sensitive approach is needed when highlighting the link between human trafficking and risk of HIV, to avoid increasing the stigmatisation and discrimination that many trafficked people already face.

Despite being illegal in every country of the world, human trafficking flourishes. Governments must clamp down on this affront to dignity, human rights, and health. Implementing HIV programmes, which include both the prevention of trafficking and the rescue of people who have been trafficked, would be a good start.
The International Health Partnership

In 2000, world leaders declared their commitment to achieve the Millennium Development Goals (MDGs) by 2015, with the aim of radically transforming the lives of the world’s poorest people for the better. Yet halfway to that date, the three health goals are furthest off track. On September 5, the UK Prime Minister, Gordon Brown, and I were joined by representatives from the leading international health agencies from developed and developing countries. Together, we launched the International Health Partnership (IHP), an initiative intended to accelerate progress on health internationally.

Lancet readers will doubtless be familiar with some of the statistics on international health, but they bear repeating. Malaria claims a million lives each year, tuberculosis nearly 2 million and HIV/AIDS 3 million. More than 500 000 women still die each year from treatable complications of pregnancy and childbirth. In July, the UN Secretary-General, Ban Ki-moon, said the world needs to take “urgent and concerted action” to deliver the MDGs. This appeal was echoed by Gordon Brown who called for an international campaign to deliver the MDGs. As part of this international effort, the launch of the IHP shows our intention to deliver the health MDGs. The IHP will do three things: provide better coordination among donors; focus on improving health systems as a whole; and develop and support countries’ own health plans.

First, health is a crowded field. There are over 40 bilateral donors, 26 UN agencies, 20 global and regional funds, and 90 global-health initiatives. In Mozambique alone, there are 26 health donors, in Cambodia 22, in Zambia at least 15. One developing country has an entire floor of the Ministry of Health simply to house the donors. To make best use of our efforts, the IHP commits us to work more closely together, particularly on the ground. This commitment means we, the donors, will wherever possible jointly appraise national health-plans. And we will jointly agree with our developing-country partners how to finance the national health-plans and provide technical assistance. In Ethiopia, for example, five bilateral donors contribute to a Health Pooled Fund used by the Ministry of Health to cover the costs of technical assistance, operations research, study tours, and joint review meetings. By increasing our coordination on the ground, we will reduce the administrative burden on overstretched ministries of health and increase our collective impact.

Second, in the past 5 years we have seen aid for health more than double, but donors have often focused too much of the funding on specific diseases. There has been considerable success in this approach. The Global Fund for AIDS, Tuberculosis and Malaria, and the US President’s Emergency Plan for AIDS Relief (PEPFAR) have played a huge role in getting more than a million people on antiretroviral treatment in sub-Saharan Africa, compared with 100 000 just 3 years ago. Yet Cambodia’s success in tackling HIV/AIDS is a result of an expanded health-care system, not simply funding medicines or condoms. And in Zambia, my office supports the Government’s Retention Scheme to ensure health workers are paid additional incentives to work in the most remote areas. There is likely to be a shortage of 4 million health workers by 2015. Tackling the problem of weak health systems must be an immediate priority if we are to build on the medical successes we have had so far.

Third, the IHP partners have committed to help support developing countries’ health-plans. These
Comment

Vascular outcome in type 2 diabetes: an ADVANCE?

In today’s *Lancet*, the ADVANCE investigators report the results of a large randomised trial of perindopril and indapamide in patients with type 2 diabetes and at least one additional risk factor. Eligible patients were assigned the angiotensin-converting enzyme (ACE) inhibitor perindopril with the diuretic indapamide, or placebo. All other drugs already being taken (except for another ACE inhibitor or thiazide diuretic) were continued; other antihypertensive drugs could be added by the physician. Just over 11 000 patients were randomised. In a mean 4.3 years of follow-up, those taking perindopril and indapamide had lower blood pressure (5.6 mm Hg systolic, 2.2 mm Hg diastolic) than did those on placebo, and a 9% reduction in relative risk of major macrovascular or microvascular events (placebo 16.8%; intervention 15.5%; absolute risk reduction 1.3%). The overall relative risk of death was reduced by 14% in the treatment group. The investigators make much of their results, recommending that this fixed-dose combination of perindopril and indapamide be considered for all patients with type 2 diabetes irrespective of their existing blood pressure, stating: “If the benefits seen in ADVANCE were applied to just half the population with diabetes worldwide, more than a million deaths would be avoided over 5 years.”

Before examination of the validity of this claim, the study methods warrant scrutiny. First, the rarity of cough (intervention 3.3%; placebo 1.3%) as a reason for discontinuation seems surprising. However, 43% of participants were already taking an ACE inhibitor or thiazide diuretic) were continued; other antihypertensive drugs could be added by the physician. Just over 11 000 patients were randomised. In a mean 4-3 years of follow-up, those taking perindopril and indapamide had lower blood pressure (5.6 mm Hg systolic, 2.2 mm Hg diastolic) than did those on placebo, and a 9% reduction in relative risk of major macrovascular or microvascular events (placebo 16.8%; intervention 15.5%; absolute risk reduction 1.3%). The overall relative risk of death was reduced by 14% in the treatment group. The investigators make much of their results, recommending that this fixed-dose combination of perindopril and indapamide be considered for all patients with type 2 diabetes irrespective of their existing blood pressure, stating: “If the benefits seen in ADVANCE were applied to just half the population with diabetes worldwide, more than a million deaths would be avoided over 5 years.”

Before examination of the validity of this claim, the study methods warrant scrutiny. First, the rarity of cough (intervention 3.3%; placebo 1.3%) as a reason for discontinuation seems surprising. However, 43% of participants were already taking an ACE inhibitor before the trial and 1.8% of eligible patients dropped out because of cough in the 6 weeks before randomisation. Therefore, most of the 10–15% of patients who started taking an ACE inhibitor and who would be expected to develop a cough were already excluded. Such infrequency of cough would not be seen if ACE inhibitors were started in ACE-inhibitor-naive patients.

Second, according to the design, all patients (except for those who dropped out) who were allocated to active treatment should be taking perindopril at the end of follow-up. However, 55% of those assigned to placebo were also taking perindopril at the end of follow-up.

The third issue relates to the validity of the claim that better outcomes could be attributed solely to the intake of perindopril and indapamide, which lowered
blood pressure. The physicians who monitored patients were free to adjust antihypertensive and diuretic regimens, and obviously did so. By the end of the trial, more participants assigned to placebo were taking an angiotensin-receptor blocker or a β blocker, a calcium antagonist, a thiazide or other diuretic, or other blood-pressure-lowering drug than were those in the intervention group. That the physicians giving all of these extra drugs did not reduce the blood pressure of patients allocated placebo (not known to the physician) as much as they lowered the blood pressure of those allocated to the intervention in ADVANCE is strange. Perindopril and indapamide are surely not that much stronger than other antihypertensives in view of the almost equal efficacy of all antihypertensive agents in moderate doses.³

Even if issues with the methods can be addressed, do the results warrant the conclusion that a fixed combination of perindopril and indapamide would be of so much value that it should be considered for all patients with type 2 diabetes? Certainly not if the so-called Polypill⁴ becomes available, because that pill contains a generic ACE inhibitor, a diuretic (although in a lower dose than in ADVANCE), a statin, and aspirin. The cost of 30 tablets of trade-name perindopril in Dallas, TX, USA, is US$61; the cost of 30 tablets of generic lisinopril is $4 at many pharmacies. Therefore, giving all patients with diabetes a trial with a generic ACE inhibitor might be logical so that the 10–15% of patients who cough could be excluded from future intake of the Polypill. Thereafter a reconstituted Polypill (without the β blocker but with a calcium antagonist) could be provided to all patients with type 2 diabetes, and to those recommended by Wald and Law—all people older than 55 years and everyone with existing cardiovascular disease.⁴

The fixed combination of perindopril and indapamide could be the best possible protector against hypertension-related consequences for patients with type 2 diabetes, but I believe that other drugs—if they lower blood pressure as much and do not have metabolic side-effects—would be as protective as this combination treatment. As has been said many times before by many experts: in most circumstances, lowering the blood pressure is what counts, not the way by which it is lowered.⁵

Norman M Kaplan
Department of Internal Medicine, Hypertension Division, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
Norman.Kaplan@UTSouthwestern.edu
I declare that I have no conflict of interest.

Launching a new movement for mental health

Despite the great attention western countries pay to the mind and human consciousness in philosophy and the arts, disturbances of mental health remain not only neglected but also deeply stigmatised across our societies. Viewed through a global lens, this marginalisation is only amplified still further. Yet the fragile—and utterly fragmented for the most disadvantaged—state of mental health services in many countries is not for the want of trying.

In 2001, for example, WHO devoted its World Health Report to mental health, with the optimistic message “new understanding, new hope”. Gro Harlem Brundtland wrote: "As the world’s leading public health agency, WHO has one, and only one option—to ensure that ours will be the last generation that allows shame and stigma to rule over science and reason.” As Director-General, she set a deservedly high standard for WHO and others to follow and be judged by.

Since then, WHO has continued to publish reports on mental health. But somehow the agency, through its leadership and partnerships, has been unable to convert fine words into tangible actions at country level. Partly this is because WHO has not backed its words with resources. And partly it is because WHO’s leadership has failed to build a sustainable mechanism across global and country institutions to hold itself and others accountable for its recommendations. This paralysis is surprising. Many low-income countries and civil society groups are crying out for help.

WHO is not the only institution with a responsibility to strengthen mental health services. The World Bank, country donors (such as the USA, UK, and European Union), foundations (such as the Gates and Rockefeller Foundations), research funding bodies (eg, the US National Institutes of Health), and professional associations all share a duty to make mental health a central theme of their strategies and financial flows. For the most part, these organisations have done far too little, if anything at all. In the past, The Lancet has tried to draw attention to mental health services in particular countries. With a Series of papers launched today from an internationally diverse Lancet Global Mental Health Group, to whom we owe a deep debt of thanks, together with a call to action and a commitment to track and monitor progress across a range of mental health indicators in the run up to a global summit on mental health in 2009, we aim to change this culture of lost opportunity.

The key messages from our Series are clear. First, mental health is a neglected aspect of human well-being, which is intimately connected with many other conditions of global health importance. Second, resources for mental health are inadequate, insufficient, and inequitably distributed. Third, there is already a strong evidence base on which to scale up mental health services. Fourth, most low-income and middle-income countries currently devote far too few resources to mental health. Fifth, there are critical lessons to learn from past successes and failures—for political leadership and priority setting, for increasing financial support, for decentralising mental health services, for integrating mental health into primary care, for increasing health workers trained in mental health, and for strengthening public health perspectives in mental health. Finally, any call to action demands a clear set of indicators to measure progress at country level.

During the next 2 years, The Lancet will make mental health one of its campaign focal points. We urge partners to join the broad new social movement we are launching to strengthen mental health.

Richard Horton
The Lancet, London NW1 7BY, UK

Most people with mental illness are not treated

In today’s *Lancet*, Philip Wang and colleagues present data for worldwide use of mental health services by people with anxiety, mood, and substance disorders. Their study shows startling findings about the treatment of people with mental illness worldwide: few are treated and most are neglected. Wang and colleagues have completed the largest international psychiatric epidemiological study so far: almost 85 000 members of the general population were personally interviewed in 17 countries. However, in another 11 countries, data gathering is incomplete. Furthermore, a potential weakness of the study is the definition of minimally acceptable treatment: for example, that psychotherapy should last at least eight sessions. This number may be reasonable in high-resource countries, but there is no evidence base for it in low-resource countries.

However, Wang and colleagues’ study overcomes its limitations by the disturbing clarity of the findings. Put simply: the degree of neglect of people with mental disorders worldwide is of such a size that any methodological quibbles are minor. The first stark finding is that although up to 30% of the population worldwide is expected to have clear-cut mental illness every year, in every country studied at least two-thirds of people who are mentally ill receive no treatment. This undertreatment occurs even in countries with the best resources: in the USA, 31% of the population are affected by mental illness every year, but 67% of these individuals are not treated. Moreover, in Europe mental illness affects 27% of people every year,7 74% of whom receive no treatment. By comparison, only 8% of people with diabetes mellitus in Europe receive no care.4

Furthermore, WHO states that the quarter of the world’s population with more common forms of mental illness (usually mixtures of anxiety and depression) should be treated in primary care, and that specialist psychiatric services should be reserved for patients who are more severely ill.5 However, real-life practice does not follow this rational template. For example, in Colombia, Mexico, and Israel more people with mental illness are treated by specialist services than by primary care. By contrast, in Germany, Italy, Japan, Spain, and the USA the numbers of people treated by these two sectors are much the same.6

In view of such substantial undertreatment of people with mental illness, can we at least be sure that health care is not squandered? Not at all. In all countries studied by Wang and colleagues, many people who did not have a diagnosable mental illness nevertheless received treatment. Almost unbelievably, of those treated, there were more with no disorder than with severe disorders for every country assessed. In the USA, every year a staggering 10% of the population is treated despite not being mentally ill. Wang does not give a financial value for this level of waste; I look forward to these data in a subsequent study.

Is it true that the people most affected by mental disorders are seen at least twice? Once more we are disappointed. Wang and colleagues show that for those who are treated but are not mentally ill, at least three-quarters are followed up—little different than for the most-severe patients. Indeed, in most of the countries they studied, more than half of the most severely ill patients received no treatment.

Previous studies have suggested a treatment gap. A WHO review of 37 studies worldwide showed that the proportion of people who are untreated for particular disorders is: schizophrenia 32%, depression 56%, dysthymia 56%, bipolar disorder 50%, panic disorder 56%, generalised anxiety disorder 58%, obsessive compulsive disorder 57%, and alcohol abuse and dependence 78%.3,7 A study of people with depression in Russia reported that only 3% of people with depression in St Petersburg were treated.5 Wang and colleagues...
present strong evidence that the scale of the treatment gap is greater than previously understood.

What do these new data mean? They are a sobering global account of how many people who are mentally ill are not treated, and how many countries have no effective treatment triage to separate individuals who are well from those with moderate or severe disorders. Treatment coverage is poor, and treatment targeting is weak. At the core of Wang and colleagues’ study, however, lie several uncomfortable questions. Why do we invest so little in our mental health care? To what extent is the underuse of services due to most people with mental illness actively avoiding help? Why have we allowed this global and gross neglect to persist and be denied for so long? Specific actions are needed now to redress this silent scandal.

Graham Thornicroft
Health Service and Population Research Department, Institute of Psychiatry, King’s College London, London SE5 8AF, UK
g.thornicroft@iop.kcl.ac.uk
I declare that I have no conflict of interest.

Depression is very disabling

“Depression produces the greatest decrement in health compared with the chronic diseases of angina, arthritis, asthma, and diabetes”, say Saba Moussavi and colleagues in today’s Lancet. From a study of more than 240 000 people in 60 countries, Moussavi also reports that respondents with angina, arthritis, asthma, or diabetes had an increased risk of depression compared with that expected in the general population. When depression was comorbid with any of these diseases, the health score was worse than with any other pair of these chronic physical diseases; diabetes with depression was the most disabling combination. This study is a head-to-head comparison of mean disability associated with five chronic diseases, as assessed in the community by use of the same method. That depression is the most disabling is a strong finding.

The first Global Burden of Disease report had a similar conclusion. Non-fatal burden for a disease was estimated from prevalence and expert judgments about average disability level or disability weight. Many people were surprised when depression was ranked as the highest determinant of disability worldwide. Subsequent burden of disease studies confirmed the finding in developed countries, such as Australia, and worldwide (table). Surveys of mental health have defined the prevalence of depression in different countries, but estimates of average level of disability, or disability weight, continue to depend on expert opinion. Moussavi and colleagues’ data lend support to the finding that depression is associated with high levels of self-reported disability.

What is known about self-reported disability? Moussavi and colleagues developed a new measure for estimation of an individual’s health-state or relative disability. For analyses of data for depression, they took care to ensure that their results were not confounded by symptoms of depression. The questions to participants covered multiple health domains to measure disability—an “umbrella term for impairments, activity limitations, or participation restrictions”. Self-reported disability in depression and other chronic diseases is reliable, and is paralleled by objective measures of performance, such as work attendance and productivity. A plaintive response set in depression could not account for the findings. Could different diagnostic thresholds account for the findings? The ICD-10 (International Classification
of Diseases) diagnostic criteria for a depressive episode need four symptoms or more for the diagnosis to be made. Moussavi and colleagues used ICD-10 criteria, and thus people with subthreshold levels of depressive symptoms were not included. Identification of the four physical diseases varied, but questions were calibrated against people who had been treated for the disorder and whose diagnosis had been validated by diagnostic tests. Thus people with mild cases of angina, arthritis, asthma, or diabetes who had not sought help would not have been included. Different standards for diagnosis seem unlikely to account for the findings.

Depression is a chronic disorder that remits and recurs; the physical disorders assessed in Moussavi and colleagues’ study do not. Could a remitting disorder generate such disability? Depression occurs throughout the lifespan (frequency 3·3%, lifetime risk 30% for men and 40% for women). The prevalence of depression is similar to the chronic physical diseases studied, but the lifetime risk—the number of people who cycle in and out of depression—is five to ten times greater than the lifetime risk of any of the other diseases studied. When depression occurs, it remains sufficiently disabling to outrank the four chronic physical diseases assessed by Moussavi.

Moussavi and colleagues appeal for clinicians to improve their recognition and treatment of depression—a disorder that is simple to recognise and not difficult to treat. Why does the burden persist? If there was a laboratory test to confirm the diagnosis, doctors might be more assertive about insisting that patients adhere to treatment in this episode and the next. In Australia, less than 30% of patients receive good treatment with antidepressants, cognitive behavioural therapy, and proactive maintenance care. By contrast, 80% of patients with arthritis and 90% of patients with asthma receive an acceptable standard of care. Perhaps differential access to treatment is one reason why disability is less with the physical disorders. Treatment for depression should at least be on a par with that for other chronic diseases.

*Gavin Andrews, Nickolai Titov
Clinical Research Unit for Anxiety and Depression, University of New South Wales at St Vincent’s Hospital, Sydney, NSW 2052, Australia
gavina@unsw.edu.au

We declare that we have no conflict of interest.
Stigma and mental health

The stigma attached to mental illness is the main obstacle to the provision of care for people with this disorder. Stigma does not stop at illness: it marks those who are ill, their families across generations, institutions that provide treatment, psychotropic drugs, and mental health workers. Stigma makes community and health decision-makers see people with mental illness with low regard, resulting in reluctance to invest resources into mental health care. Furthermore, stigma leads to discrimination in the provision of services for physical illness in those who are mentally ill, and to low use of diagnostic procedures when they have physical illness.

Stigma of mental illness can be defined as the negative attitude (based on prejudice and misinformation) that is triggered by a marker of illness—eg, odd behaviour or mention of psychiatric treatment in a person’s curriculum vitae. The presence of stigma starts a vicious circle that leads to discrimination in all walks of life, decreasing self-esteem and self-confidence (resulting at least partly from the experience of a person with mental illness), a low treatment effect or high probability of relapse for those in remission, and thus to a reinforcement of the negative attitudes and discrimination. This model of the vicious circle suggests that there can be various strategies for those who wish to fight stigma. We can think of ways of reducing the visibility of markers (eg, by provision of treatment that is not associated with extrapyramidal side-effects), ways of reducing discrimination, interventions that will help raise the self-esteem of people with mental illness, education of families, and more investment in treatment that rapidly reduces the severity of illness or that prevents relapses.

Many people contribute to the development and reinforcement of stigma. Health-care workers commonly use words that are stigmatising—eg, speaking of schizophrenics, or use of pejorative terms for mental illness instead of speaking of the person who has the illness. Medical personnel may refuse to treat physical illness or injury in those with mental illness. Psychiatrists and mental health personnel are no exception in this general unawareness of how their behaviour contributes to stigma. People who have mental illness and who have lost much of their self-confidence over time contribute to the image of the person with mental illness who does not try to contribute to their health and sustenance. Governments make statements or take action that reinforce prejudice—eg, by proposing sterilisation of people with mental illness or retardation without their consent, or by avoiding parity in reimbursement for treatment of mental illness.

However, there is good news. The global programme against stigma and discrimination because of schizophrenia was successfully implemented in more than 20 countries worldwide. Furthermore, local initiatives have successfully removed or reduced stigma. The European Union’s recent consultation about mental health promotion identified the fight against stigma as an important area of work for European countries. WHO has highlighted the need to combat stigma and to foresee appropriate measures in national mental health policies.

Moreover, experience over the past two decades has identified the main parts of successful action against stigma. These are: consultation of people with mental illness and their families about targets for action against stigma and their involvement in relevant programmes; conception of the fight against stigma as a long-term endeavour that is incorporated into health and other social services; involvement of all stakeholders in the programme, including government, health-service personnel, and the media; and a focus on specific problems that result from stigma (eg, discrimination.
against people with mental illness) rather than generic approaches to change people’s attitudes.

Stigma attached to mental illness is the main obstacle to the success of programmes to improve mental health. It is fortunate that determined action can remove stigma, to a large extent.

Norman Sartorius
Association for the Improvement of Mental Health Programmes, 1209 Geneva, Switzerland
sartorius@normansartorius.com

I declare that I have no conflict of interest.

1 Fang H, Rizzo JA. Do psychiatrists have less access to medical services for their patients? J Ment Health Policy Econ 2007; 10: 63–71.

Royal Society of Tropical Medicine and Hygiene: 100 years old

The Royal Society of Tropical Medicine and Hygiene (RSTMH) was established at a meeting of “Medical Men and others interested in Tropical Medicine” in the Colonial Office on Jan 4, 1907. Sir Patrick Manson, generally regarded as the father of tropical medicine, was elected President and Ronald Ross Vice-President—an auspicious start (figures 1 and 2). The first ordinary meeting of the Society followed on June 26, 1907. The American Society of Tropical Medicine had been established 2 years previously and several European societies were established at about the same time—the classic age of tropical medicine when the causes of most major tropical infectious diseases were established.

The initial objectives of the RSTMH were broad: “to promote and advance the study, control and prevention of disease in man and other animals in warm climates, to facilitate discussion and the exchange of information among those who are interested in tropical diseases, and generally to promote the work of those interested in these objectives”. With replacement of warm climates by developing countries, these remain the Society’s objectives. In its first 50 years, the Society’s main role was to provide support to the predominantly European clinicians and scientists in the colonies. However, in recent years the Society has gained an increasing number of fellows from developing countries. The Society currently has a fellowship of about 1500 residents in 105 countries.

In its early years, the RSTMH achieved its goals through the sponsorship of meetings that were published in the Transactions of the Royal Society of Tropical Medicine and Hygiene. For 100 years, the Transactions has had a key educational role for clinicians working in the tropics. As late as the 1970s, when I was based in northern Nigeria with no communication apart from mail, the Transactions arriving by sea mail after about 6 weeks was a key source of information.

During the first 30–40 years after the independence of European colonies, the rich countries of the northern hemisphere showed little interest in the health problems of the developing world, and many of the research
institutes established during the colonial period and universities established around the time of independence decreased. During this period, the RSTMH was one of the few organisations that helped to sustain an interest in the health problems of the developing world. During the past 5–10 years, there has been a massive increase in support for international health from both public and private donors, and sums of money have become available for control of major infectious diseases and for research into improved methods of treatment and prevention, on a scale that was unimaginable 10 years ago. How should societies of tropical medicine respond to this different environment? It could be argued that, with a focus on meetings in the UK and publication of a fairly expensive hard-copy journal, the RSTMH has served its time and should be left to expire. However, such an outcome would seem bizarre when there is increasing interest in international health and an increasing number of people working in this specialty. A more constructive approach should be for the Society to find new ways to contribute. Thus, during the period leading up to its centenary, the Society has reviewed its activities and how these might be expanded to meet new challenges. Fellows of the Society and other interested parties have been consulted, with views solicited at meetings in The Gambia, Tanzania, Peru, Malaysia, and Australia. An initial debate was whether the name of the Society should be changed to one better suited to 2007. Some felt that the term hygiene, used initially for what would now be considered public health, was anachronistic. However, most felt that, although changing what the Society does is crucial, changing the brand name may not be wise.

The RSTMH will continue to do the things that it already does well. It will support high-quality meetings in the UK and continue to support the professional development of UK clinicians engaged in tropical medicine and infectious diseases. This support is especially important for those working for a period in the developing world whose needs are not always taken into account. However, the Society must also expand its scope of activities if it is to continue to make a major contribution to international health. It must reach out to students and young graduates in the UK who are interested in international health, and encourage them to become involved in the Society’s activities. This expansion will be done by reducing subscriptions and by sponsoring meetings of special interest to these groups. The Society should also find ways of ensuring that fellows outside the UK obtain some benefit from membership. Improvements in communications across much of the developing world make that easier. The Society now maintains regular electronic communication with nearly all its fellows and, in 2007, Councillors were elected electronically. Council members outside the UK can now contribute actively to the Society and attend meetings of Council via telecommunication.

The advent of free-access electronic journals, including an increasing number relevant to international health, poses a challenge to the traditional hard-copy journals of tropical medicine, such as the Transactions. In addition to the free-access journals, many journals not generally available freely online are now available without cost in many developing countries through schemes such as the Health InterNetwork Access to Research Initiative. Thus, for many, joining the RSTMH to obtain its journal is no longer an attractive option, which has potentially serious consequences for the Society by reducing the size of the fellowship and its budget. However,
Can drug regimens be adapted to patients, or vice versa?

Drug non-adherence can lead to increased morbidity and mortality, reduced quality of life, and wasted health-care resources. According to a recent systematic review by Sunil Kripalani and colleagues,1 some types of interventions can improve drug adherence but reported effect sizes for clinical outcomes were highly variable. These authors recommend that research should seek to develop innovative ways rather than persist with existing approaches. We offer a new piece for this puzzle, which starts from the recognition that drug non-adherence is often intentional.

Qualitative research shows that many people change their regimens because they have concerns about their drugs.2 Besides stopping medications altogether, they may start to take drugs symptomatically or strategically, adjust doses to reduce unwanted consequences, or make regimens more socially acceptable. Such modifications show a desire to keep drug use to a minimum, which is sometimes also evident from supplementation or replacement of a treatment with non-drug measures or unconventional remedies.2 A professional response to such findings has been the notion of concordance, in which prescribers no longer use contacts with patients merely to give instructions but as opportunities for reaching agreements with them. We propose that prescribers do not only need good communication skills for pursuing concordance. They could also benefit from evidence about the potential advantages and paradoxically, increasing free electronic access to the Transactions across the developing world makes it potentially a more valuable means of disseminating information than ever before, with downloads of about 10 000 articles every month. The Society is, therefore, making strenuous efforts to improve its journal and to respond to a rapidly changing environment. In future, the journal will cover articles on a wide range of topics relevant to international health, such as health policy and implementation research, and will have a more interesting format, including editorials, leading articles, reviews, and correspondence, as well as peer-reviewed research papers. The Transactions is supplemented by an electronic bulletin which provides news and views from the fellowship. Expansion of communications across the developing world opens up many new educational possibilities. In collaboration with the Royal College of Paediatrics and the Wellcome Trust, the Society has compiled a compendium of teaching materials available electronically. Development of imaginative ways to use the web to help those working in the developing world will be an important part of the Society’s future.

The Society already supports several modest scholarships and prizes, funded by endowments. Thus, the Council decided that the best way to recognise its centenary was to establish a scholarship fund for developing-country researchers and health-care workers that will support them in acquiring new skills through short-term fellowships in a centre of excellence and by providing mentorship. Fund-raising for this programme is in progress and a longer-term fund-raising programme may be needed if the Society is to achieve its full potential.

The highlight of the Society’s centenary will be a 3-day meeting: “One hundred years of tropical medicine—meeting the millennium goals” on Sept 12–14, at the Queen Elizabeth Conference Centre in London. This meeting will be followed in subsequent years by an annual residential meeting of the Society. These meetings, held in the UK or elsewhere, may be organised in collaboration with other UK groups working in international health and with overseas societies of tropical medicine. It is hoped that these meetings will become an important focus for international health activities in the UK and will be of particular value to young scientists entering this field.

The potential for the RSTMH to make as big a contribution to tropical medicine and international health in its second hundred years as it did in the first is huge, but finding the best ways of doing this will need imagination, hard work, and financial support.

Brian Greenwood
Department of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK
brian.greenwood@lshtm.ac.uk
I thank colleagues on the RSTMH Executive Committee for helpful comments on this text, which reflects personal views and is not an official position of the Society. I am currently President of the Society.
risks of supporting and guiding individual patients in their pursuit of dose minimisation instead of always attempting to enforce compliance.

Scientific support for this idea can be found in the pharmacological fact that minimum effective doses can vary with interindividual differences in bodyweight, organ functions, pharmacogenetic properties, and so on. Dose recommendations in package inserts and textbooks have often been derived from pivotal clinical trials that were not designed to establish minimum effective doses in individual patients. For example, the British National Formulary and Martindale The Extra Pharmacopoeia recommend intramuscular doses of 50 mg per week for sodium aurothiomalate, whereas case series suggest that much lower regimens of intramuscular gold may be sufficient in individual patients without causing the same side-effects as high doses can.7,8 This example shows that data about the possible outcomes of dose minimisation should be disseminated more systematically than textbooks have been doing so far. For instance, searching published work for details about simvastatin provides evidence which suggests that some patients with hypercholesterolaemia may already control LDL-cholesterol by taking less than the recommended dose of at least 10 mg daily.5,6 That taking statins (including simvastatin) less often than every day does not necessarily jeopardise control of LDL-cholesterol (but may increase the risk of non-adherence),9 that a combination of cholesterol-lowering foods may sometimes be an alternative to lovastatin (which is closely related to simvastatin),9 and that the addition of a herbal remedy (psyllium) to 10 mg simvastatin may lower LDL-C by as much as does 20 mg simvastatin alone.10

Prescribers should consider some important caveats, before they offer their support and guidance to patients who are interested in pursuing dose minimisation. First, it must be possible for prescribers to evaluate whether dose minimisation leads to underdosing or not. This evaluation is feasible when a drug is taken purely for symptomatic relief or when an intermediate outcome can be easily and reliably monitored (eg, as with antihypertensive treatment). When such measurements are impossible or impractical (eg, with antiplatelet therapy), professional support of dose minimisation becomes inappropriate. Second, the prescriber should take into account that a risk of temporary ineffectiveness will entail such serious health hazards in some patients that this risk becomes unacceptably high. Third, an individualised approach will need extra time and money in terms of doctor’s visits and laboratory measurements. However, these costs should be weighed against the overwhelming clinical and economic consequences of current non-adherence. Another potentially important benefit of our proposal is that patients may feel taken more seriously by their prescriber and may find the standard dose recommendation more acceptable, after an attempt at individual dose reduction has shown that this standard dose is what they actually need.

We do not advocate large-scale implementation of our proposition. Instead, we call for clinical practice studies to find out how patients with an inclination to dose minimisation can be identified in daily practice and to assess whether professional support and guidance will help such patients to improve their benefit-to-risk ratios and to accept their drug therapies. Such studies might also investigate whether self-monitoring at home (eg, of blood pressure) has to play an additional part in this approach.

*Peter A G M De Smet, Cees Kramers, Nicky Britten
Department of Clinical Pharmacy, University Medical Centre St Radboud, 6525GA Nijmegen, Netherlands (PAGMDes); Department of of Pharmacology and Toxicology, Radboud University Nijmegen, Nijmegen, Netherlands (CK); and Peninsula Medical School, Universities of Plymouth and Exeter, St Luke’s Campus, Exeter, Devon, UK (NB)
pdesmet@winap.nl

We declare that we have no conflict of interest.

4 Cheatum DE. How low can you go? Use of very low dosage of gold in patients with mucocutaneous reactions. J Rheumatol 1996; 23: 944
When psychologist Edgar Bellido arrived in Pisco after a magnitude 8.0 earthquake virtually levelled the southern Peruvian city on Aug 15, 2007, he spent 2 days without eating or sleeping. But he had only one request “I asked them to please find me a computer”.

Bellido, a member of the Peruvian Ministry of Health’s national mental-health technical team, was in Pisco to head the mental-health response to the disaster. Setting up shop in the hospital’s only undamaged wing, he set to work gathering data about the scope of the crisis, as well as counselling the rescue teams searching for victims under the rubble.

Now he is planning for the long term. Some 25 tent cities housing at least 20 000 people have sprung up, and this number could swell to 50. Now that the immediate crisis is past, people are seeing “their entire lives being swept away by the machines that are clearing the debris”, he said.

Unlike a hurricane, an earthquake gives people no time to prepare. In and around coastal Pisco, some 30 000 families saw their homes collapse around them in 2 minutes when the earthquake struck.

The disaster killed over 500 people, injured 2000, and created a mental-health emergency. Rescue crews needed psychological support, patients who were receiving treatment lost their medications, and people with stress or anxiety saw their symptoms magnified.

Psychologists from public and private agencies headed to the scene, but there was an initial lack of coordination, said Bellido. Their work was further complicated by lack of access to patients’ records because of damage to the hospital, according to Luis Encinas, emergency coordinator for Médecins Sans Frontières. The situation highlights the need for a backup system for patients’ records in case originals are damaged or destroyed in disasters, says Encinas.

However, in Pisco, the community organisation system—typical of Peru’s low-income urban neighbourhoods—was an asset, as community leaders quickly identified people who needed medications or special care, Encinas said. That spontaneous organisation “greatly facilitated the delivery of aid” and with specific training, he believes, such leaders could become the front line in a medical emergency response.

Although it may be too early to identify all the lessons to be learned from this disaster, some have emerged (panel). One of the biggest challenges now is providing long-term mental-health assistance in the country, where community-based mental-health care is still in a fledgling state.

“The treatment gap” is not peculiar to Peru. Despite advances in mental-health care throughout Latin America, overall, at least half the people who need care go untreated because of lack of access to services, stigma, discrimination, economic problems, or lack of awareness, according to Jorge Rodriguez, who heads the Pan American Health Organization’s mental-health unit.

Until recently in Peru, as in most of Latin America, the emphasis was on mental illness rather than mental health, and care revolved around psychiatric hospitals. This focus changed with a national mental-health plan that was launched in 2006, but implementation requires changing the mindset of patients, their families, health-care providers, and politicians.

In one sense, the lack of emphasis on mental-health care in developing countries such as Peru is understandable. “While in Europe, there is concern about more subtle aspects, we have had to be concerned..."
about children who are dying from malnutrition and lack of antibiotics”, said Jaime Saavedra, who heads the research unit at the Honorio Delgado-Hideyo Noguchi National Institute of Mental Health in Lima, the Peruvian capital.

But, as well as affecting physical conditions, poverty also has an effect on mental health, and vice versa, according to a study done this year in Lima. The study found that depression was twice as high in poor people as in the well-off. The same was true of mental-health indicators such as child abuse, domestic problems, and suicide attempts.

This situation creates a vicious circle, according to Saavedra, who was the lead researcher for the study. Poverty brings a higher frequency of mental-health problems, and those problems make it more difficult for a person to hold a job, perpetuating the poverty.

This problem suggests that anti-poverty programmes must take mental health into consideration, whereas mental-health programmes should include elements such as occupational training.

Peru’s national mental-health plan points in this direction. It calls for mental health to be seen as a fundamental right and for universal access to community-based mental-health care.

Lozada’s goal is to create a system of community-based mental-health care in which primary-care facilities would resolve some 80% of cases. Local or regional public hospitals would set aside about 5% of their beds for people requiring admission to hospital, and specialised facilities would be used as a last resort.

To implement such a model, primary-care physicians would need continuing education and the university curriculum must be refashioned to ensure that doctors and nurses can manage and refer people with mental-health problems appropriately.

Human resources must also be decentralised. With a population of more than 23 million people, Peru has only about 500 psychiatrists, and at least half are in Lima, where a third of the population lives.

But financing these changes could be problematic. “The budget for mental health actions is zero”, said Lozada, who would like to see a separate mental-health office established in the health ministry. Although about US$2 million has been cobbled together for mental-health services from donations and funds from international agencies, “the big problem is sustainability”, he said. Local governments should pick up at least part of the tab for setting up community-based programmes, but “municipalities do not want to invest”, said Lozada.

However, Lozada is realistic about his limitations. “I cannot pretend to solve in a few months a problem that has existed for years”, he said a week and a half after the earthquake, as he prepared to go to Pisco to meet with local government officials about organising a community-based mental-health response.

Hard lessons from disasters, such as the recent earthquake, are helping Peru to cast mental health into a new light, but there is still a long road ahead.

Barbara Fraser

Panel: Early mental-health lessons from the Peru earthquake

Before an emergency:

- Planning is crucial and should include detailed information about human resources, steps for mobilising them quickly, and guidelines for different scenarios.
- If other countries send mental-health teams, there must be a clear chain of command and criteria for determining whether they are culturally sensitive and where they will work.
- A backup system for patients’ records is needed in case originals are damaged or destroyed.
- Emergency kits containing a month’s worth of essential medications for people with chronic illnesses can be prepared in advance.
- A system for delivering medications to patients in a disaster situation is needed.
- Planners must remember that local service providers may also be affected.

When disaster strikes:

- Immediate support should be provided to rescuers as well as the general public.
- Those responsible for organising services need accurate, updated information about the number of affected people, their ages and location, and shelters and outside aid organisations that are assisting, so as to plan a targeted response.
- Initial psychological support for people who have lost homes or family members should be done in group sessions broken down by age.

Looking to the future:

- Some people will need long-term assistance to recover or to deal with additional stresses, such as those related to overcrowding in the tent cities where earthquake victims will live until their homes are rebuilt.
- Community leaders could be trained to assist in emergency medical situations.
Last month marked the 60th anniversary of Pakistan’s creation and independence from British colonial rulers. But some Pakistani doctors could find little to celebrate about when it came to talking about health in the country. “Over the last 60 years we have failed to establish a proper health system. All health initiatives including those addressing mental health needs are undermined by massive corruption, poor governance, and gross mismanagement”, says Murad Moosa Khan of the Aga Khan University, Karachi, Pakistan.

The situation is especially worrying for mental-health advocates because suicide rates have surged in Pakistan in recent years—from a few hundred pre-1990s to almost 7000 last year. And Khan says that “this is probably an underestimated figure given the legal, socio-cultural, and religious sanctions against suicide in Pakistan”.

Suicide bombings and robberies are also on the rise. Psychiatrists believe these violent episodes create a huge mental-health problem for the population. Data from population-based studies indicate that a third of Pakistan’s population have anxiety and depression.

Although cost-effective treatments exist for these disorders, Pakistan does not have a critical mass of trained mental-health professionals, such as psychiatrists, psychologists, nurses, and paramedics to deal with them. There are only 250 psychiatrists for a large population of 160 million, but less than half of them have proper postgraduate qualifications in psychiatry.

“Psychiatry is not a certifying examination in the medical schools, with the result that generations of Pakistani physicians have gone through medical training with little or no exposure to mental-health issues”, according to Khan.

One way to develop the critical mass would be to include mental health as a separate subject in the curriculum in medical, nursing, and paramedic schools. Many Pakistani health experts are also calling for the need to increase training positions in psychiatry, introduce a certifying examination for psychiatry, and establish a national community mental-health programme that is low cost, culturally relevant, and evidence based.

But Khan believes that WHO-led efforts to integrate mental health into primary health will not work in Pakistan because public funded primary health care is fractured, ineffective, and under-resourced. In addition, much of the primary health care is provided by illegal, unqualified primary-care providers working in the private sector. Unless the quality and the spread of state-run primary health care is improved, some experts think that efforts to introduce mental-health services as part of the existing system would be futile.

The need for mental-health services is clear in a country that has not only been hit by several natural disasters during the past couple of years but has also seen a deterioration of law and order since the Sept 11, 2001, terrorist attacks in the USA.

In North West Frontier Province (NWFP), a conflict between the pro-Taliban local tribal communities and Pakistan’s military in the neglected tribal areas is taking a heavy toll on mental health. The proportion of civilian deaths is rising. Because tribal areas do not have mental-health services, people with mental illnesses use the few services that exist in the large cities of NWFP, especially Peshawar.

“A significant proportion of our patients are from tribal areas, but many of them cannot get the care they need because of the ongoing conflict”, explains Saeed Farooq, from the Lady Reading Hospital in Peshawar. He says that people in the tribal areas have high rates of conflict-related mental trauma.

The mental-health situation is complicated by the existence of several natural disasters that have taken their toll on mental health in Pakistan. But a shortage of mental-health professionals and lack of access to psychological services means the needs of the Pakistani population are not being met. Khabir Ahmad reports.
Frontier Crimes Regulation (FCR)—a draconian law introduced by the British colonial rulers in 1901 to govern the tribal areas and which violates even the most basic human rights guaranteed in Pakistan’s Constitution and international human rights law. For example, when a person in the tribal area commits an offence, a whole tribe or area is punished. Often punishments include large fines, house destruction, and even imprisonments of the entire tribe.

Around 70% of Pakistan’s population lives in rural areas where there are hardly any mental-health services. The worst affected has been Baluchistan, which is geographically the largest of the four provinces of the country and which contains 40% of Pakistan’s natural gas reserves.

Almost untouched by development, the province has some of worst health indicators in South Asia. Pakistan’s military is fighting against the inhabitants of the province who have been demanding a fair share in the development and gas profits.

Decades of poverty, discrimination, and droughts have badly affected the mental health and psychosocial wellbeing of people living in the province. The continuing violence, and recent floods and cyclones, have only worsened this situation. Since late June this year, an estimated 2·5 million people in Baluchistan and Sindh have been affected by floods and cyclones. In August, the Pakistan’s National Disaster Management Authority, which has come under harsh criticism for lacking the capacity to manage or implement disasters operations, said that the total death toll from floods since late-June has reached 417—203 in Baluchistan and 214 in Sindh.

Despite the high burden of psychological problems, mental health got no mention in the priority areas highlighted in Pakistan’s 2001 National Health Policy. “Only 0·035% of the federal health budget for 2007, or US$82 304, has been allocated to health promotion and the prevention of mental illnesses and non-communicable diseases. This is evidence of the low priority that is accorded to mental health in the country”, says Sania Nishtar of Heartfile, a non-governmental organisation based in Islamabad.

But Malik Husain Mubashar from the University of Health Sciences, Lahore in Pakistan believes that “there is strong policy commitment at the national level to mental health as evidenced by the replacement of the British colonial law, 1912 Lunacy Act, by the Pakistan Mental Health Ordinance in 2001”. However, a large gap exists between the stated policy and services delivered.

To address this problem, Mubashar believes it is now the responsibility of provincial government to provide funds and the mental-health professionals, especially psychiatrists, to provide leadership.

But where will the leadership come from? Several mental-health experts interviewed by The Lancet recommended the establishment of a national professional body similar to the Royal College of Psychiatrists in the UK that should become the most authoritative voice of mental health, not only psychiatry, in Pakistan. The Pakistan Psychiatric Society, they argue, has been ineffective as was witnessed in the mental-health relief efforts during the 2005 earthquake. There was a clear difference in earthquake-related relief efforts, with military mental-health services on one side and civilian on the other.

There were also problems with WHO-led efforts. Haider Naqvi from the Aga Khan University, Karachi, says that “although there were lots of funds available and the intentions were good, there was a lack of capacity to use them and there was little or no co-ordination of the mental health efforts”.

The 7·6-magnitude earthquake resulted in 74 000 deaths and left around 3·5 million people homeless in northwestern Pakistan and Kashmir. Many national and international organisations provided psychological support services and training, but had separate treatment training and teaching protocols. And, according to Farooq, evidence that psychological interventions work in the aftermath of a disaster remains flimsy. “Post-traumatic stress disorder was uncommon, but the areas [affected by the earthquake] had pools of untreated chronic [mental] disorders because of decades of neglect and that is what we need to address first.”

Khabir Ahmad
There are many things in life to make one feel sad. Losses abound. Relationships go to pieces. People get sacked from a decent job. Careers fail. Aesthetic or moral projects are checked. Families fall down the class and status ladder. A myriad of disappointments can demoralise and defeat any of us. And, as we age, we sense death coming.

Researchers and clinicians (even the general public) have come to use the euphemism “stress” to stand for the routine and extraordinary dangers that each of us experience. These run from financial crisis to health catastrophes; from serious accidents to disabling chronic disorders; and, especially among the truly poor, from incidental to structural violence. In most societies the popular culture’s wisdom makes the point that life is difficult, uncertain, and only poorly predicted or controlled. That folk wisdom increasingly flies in the face of hyped claims by experts that we know enough about life to manage “risk”. The actual experiences of people (including the experts) in the real world belie this utopian distortion of our existential condition: what the great American philosopher-psychologist-physician William James called “genuine reality”.

Allan Horwitz and Jerome Wakefield’s important book, The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder, is part of a gathering blowback against the pathologicalisation and medicalisation of the ordinary human condition of sadness after loss. More specifically, these senior social scientists, whose research careers have been devoted to studying mental illness and psychiatry, respond to the inflated claims (and findings) of psychiatric epidemiologists that we are living through an epidemic of depressive disorder. Although they do not underestimate the social forces driving the political economy of pharmaceuticals and the global cultural changes that underpin the overconsumption of these substances, Horwitz and Wakefield choose to focus on professional diagnosis as the main culprit.

Their argument runs like this. For thousands of years of recorded medical history, it is well documented that physicians understood that symptoms “Does anyone really know what the range of duration of normal sadness is among the 90% of the world’s population that lives outside EuroAmerica?” cannot be interpreted outside of the actual context of the patient’s life. No master clinician, in the past, would confuse depressive disorder with normal grief, unless the symptoms of grief lasted such a disproportionately long time and were dangerously dysfunctional to the patient and his or her world as to indicate pathology. Horwitz and Wakefield suggest that the same professional common sense informed the diagnostic systems from Hippocrates and Galen to pre-1980 medicine in the West for other losses from jobs and status to lovers.

Then came the cultural revolution of the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III, 1980). To improve clinical reliability, DSM-III simply added up symptoms. With the exception of bereavement—which the latest version (DSM-IV) grudgingly regards as normally lasting 2 months—DSM-III recognised no contextual events, besides other diseases, that might qualify the depressive syndrome as a normal response to serious life events. A modern Romeo might experience sadness after the break-up of a consuming love affair and would have several weeks or a month of sadness, sleeplessness, exhaustion, difficulty concentrating on his work, agitation, and lack of interest in eating and other previously valued things. In the DSM-III, the symptom count would easily make the cut off for depressive disorder, never mind the obvious social source of the problem or even the fact that, left to his own devices, our young man might no longer experience symptoms as he got over his loss and found a new love.

So far so good. Most of us doubtless do not want to see sadness transformed into clinical depression. And for good reason. Treating normal sadness has not been shown to be effective, may expose the non-disordered person to serious side-effects, may interfere with cultural and personal meaning-making that is a natural part of being human, and almost certainly will divert resources away from the care of those who are seriously disordered.

Had the authors stopped at this point, they would have been on firm ground and the book would have been much shorter. But Horwitz and Wakefield are not satisfied with writing a powerful critique of the mainstream psychiatric diagnostic programme; they have an alternative programme of their own to advance. They assert that what amounts to most of human travail in the face of loss can be understood as a specific instance of a putative human evolutionary biology of loss. Sadness, they repeat, has an adaptive function—albeit neither they nor anyone else knows for sure what that might be. Thus, all sadness that is the result of troubling life events is normal, if that syndrome of sadness is proportionate in quality, does not last too long, and does not produce dysfunction. How they (or we) are to recognise what is proportionate, not too long, or a dysfunction seems to be based on even thinner evidence than the claims of the ever-expansive
psychiatric epidemiologists whom they so effectively critique. The data on this allegedly universal biology of loss are simply not there. We do not have the cross-cultural evidence to define the characteristics of normal bereavement, let alone sadness in response to other kinds of loss. Recent research on bereavement among Americans finds strong evidence that yearning is as powerful an affective response as sadness. Does yearning carry with it its own, distinctive biology? How does the biology of anxiety or anger interact with that of sadness? Does anyone really know what the range of duration of normal sadness is among the 90% of the world’s population that lives outside EuroAmerica? Is replacing the medicalisation of depression with the biologisation of sadness a useful trade-off? After all, if the various loss responses are due to a universal process of human evolutionary biology, why wouldn’t that process, once known of course, become a new target for pharmacological modification? And what a target that would be! None of us would escape treatment and prevention efforts. Sadly, Horwitz and Wakefield close one door in a powerful critique of medicalising depression, only to open a new door that turns all of sadness and much of life into a potentially modifiable target for the brave new world of human evolutionary biology.

Georges Canguilhem, the influential 20th-century French historian and philosopher of medicine, showed that the relation between pathology and normality was an abiding theme crucial to medicine. This was so not only because history altered categories, including the line between the normal and the abnormal, but also experiences: context and reaction. Norms, for Canguilhem, and for many of his followers in medical anthropology and history, infold into the body creating normality; the body itself is both universal biological processes and culturally particular local biology.

Horwitz and Wakefield do not have much to say about these issues, which if seriously considered would demand qualification and criticism of their evolutionary project. But what they do accomplish in critiquing psychiatric diagnosis of depression is important enough to make much of this book required reading for depression researchers and clinicians.

Arthur Kleinman
kleinman@wjh.harvard.edu

In brief

Book Teen depression

What’s the major health issue affecting young people today? The answer’s simple. Indeed, the story was broken in The Lancet’s recent series on adolescent health: depression and a raft of related mental and substance use disorders. In developed countries, the greatly improved physical health of adolescents and young adults has allowed the spotlight to fall squarely on mental health issues. Monochrome Days tackles this issue weaving together a blend of personal experience, key facts about depression and its management, and self-help tips and guidelines. Despite the fact that up to a quarter of young people will have clinically significant depressive symptoms at some point during the transition to adulthood, and many others will develop other mental or substance use disorders, there is much ignorance and misinformation about how these problems develop and how they can be helped. Mental health literacy, a concept developed by Australian researcher, Anthony Jorm, is central. We now know that levels of knowledge and competence among young people and their families about depression are quite poor, but can be improved through various strategies. Books like Monochrome Days are part of the process, but recent evidence suggests that electronic media and internet-based strategies are likely to be more potent with young people themselves—for example, the Australian Reachout! website (http://www.reachout.com.au). Penetrating youth culture, educational settings, and workplaces directly and via various media is the way to go. Importantly, this book is part of a broader US initiative to target adolescent mental health, and is based on the real life experience of depression of Cait Irwin, a young woman who successfully overcame the illness. The book is clear and accessible, and woven around Cait’s story are facts about depression and its treatment. Although most of this information is useful and contains much wisdom, the tone has a culture-bound feel to it, embedded as it is in middle America.

A key issue that Monochrome Days fails to highlight is the degree of difficulty in seeking help and gaining access to skilled mental health professionals. Little mention of the ubiquitous blind alleys, wrong turns, false starts, and lame therapists. It is quite extraordinary, given the public-health importance of unrecognised and untreated mental and substance use disorders among young people, how low on the health-priority list these feature and how embryonic are mental health literacy strategies for adolescents. A move to polychrome is overdue.

Patrick McGorry
pmcgorry@unimelb.edu.au
Profile

Vikram Patel: promoting mental health in developing countries

In the summer of 1993, Vikram Patel boarded a flight from London’s Heathrow airport bound for Harare, Zimbabwe. A few hours later, stepping into the heat of an African afternoon, he began a 2-year post as a research psychiatrist at the University of Zimbabwe Medical School, an experience that would shape the trajectory of his future career. As Patel and his wife settled into the small flat that they had rented from the university, they began to fall in love with the country. Its physical beauty, the fervour of its political and economic discourse, and the quality of its cultural and academic life were exhilarating. Patel recalls, “There could have been few places on the planet better to live”, he says. “I was completely mesmerised by that experience.”

It was just as bracing in a professional context, he says. The first challenge was simply to find ways of communicating with his patients, most of whom spoke Shona. “This motivated me to go to twice weekly language classes, but clearly 2 years was not enough to master it!” Gradually, Patel found Zimbabwe changing his preconceptions about mental health in the developing world. “I went in there highly sceptical that mental illnesses like depression were valid diagnostic categories of suffering in a place like Zimbabwe”, he says. “I set out to prove myself, but left convinced I had been wrong.” He saw that although patients in Africa used different concepts and symptoms to describe their problems, and often turned to traditional healers for relief, their conditions were comparable to those experienced by people in the developed world. “I was convinced that although social factors were a major cause of depression, depression was a genuine cause of human illness in this very different cultural setting, and one which could respond to clinical and public-health interventions.”

This realisation would prove to be fundamental in honing the preoccupations of Patel’s current work, which has focused on addressing mental disorders in developing countries from a public-health perspective. Another key factor was his own background. Patel was born in 1964 in Bombay, India, into a family that had specific aspirations for his career. “My first career choice was catering, but in those days to be a chef was considered unacceptable for a middle-class Indian boy”, he says with a laugh. “There was absolutely no question in the milieu I grew up in that my choices were medicine, law, or engineering.” He chose the former, and then—much to the disappointment of friends and family—psychiatry. He completed medical school in Bombay, and spent the last year of his internship in the state of Goa. After finishing his degree, Patel won a Rhodes Scholarship to Oxford University in the UK.

At Oxford, the young trainee psychiatrist was taken under the wing of Tony Hope, now the university’s professor of medical ethics. Patel then trained in psychiatry at the Maudsley Hospital in London where he met Anthony Mann who became his supervisor for his PhD. Patel cites Hope and Mann as two of the four most influential people in his professional career, the others being David Mabey and Betty Kirkwood, who have supported his work since he joined the London School of Hygiene and Tropical Medicine in 2000. “What I have learned most from them is the art of mentoring younger scientists and researchers”, he says. “I hope I can be nearly as effective as they were.”

After leaving Zimbabwe, Patel was drawn back to Goa as a researcher, initially with the Institute of Psychiatry, London, and then with the London School of Hygiene and Tropical Medicine, funded by the Wellcome Trust. He now divides his time between London and Goa: teaching; building community mental health institutions; promoting research, clinical care, and training; and raising awareness of the burden of, and treatments for, mental illness in the developing world. It hasn’t been an easy task, he says.

Most of the world’s 400 million people with mental disorders live in poor countries; indeed, mental illness accounts for more than 10% of the total burden of disease in low-income and middle-income countries, Patel says. Many people with mental disorders can be helped with simple and cheap treatments to recover or, at the very least, to vastly improve the quality of their lives. However, their plight receives scant attention and almost no funding from donor agencies that specialise in addressing the problems of developing nations. “The tendency has been to think that if a problem also occurs in developed countries, then it isn’t really a developing country problem”, he says. Even as countries themselves are waking up to the scale of the problem, most aid agencies do not even consider proposals to treat mental health problems. “That’s the amazing mismatch here; there are poor countries that want mental health reform, but no donors willing to help them”, says Patel.

Still, he does see signs of change. At a recent meeting of the International Advisory Panel to the National Rural Health Mission of India, of which Patel is a member, the economist Jeffrey Sachs and the Indian Minister of Health, Anubam Ramadoss, both called for greater importance and resources to address mental health in rural impoverished settings. “There definitely is now an acknowledgment that mental health is a major global health issue, but the changes needed to transform this acknowledgment to resources and action are coming ever so slowly”, Patel says.

Stephen Pincock
stephen.pincock@journalist.co.uk
Pax Allan Chingawale

Community health worker and HIV/AIDS activist. He was born on Aug 8, 1953, in Nyani Village, Malawi, and died after a road collision near the town of Zomba, Malawi, on July 24, 2007, aged 54 years.

Pax Allan Chingawale became an HIV/AIDS activist at the comparatively late age of 48 years, when he discovered he was HIV-positive. A retired government auditor, Chingawale was studying towards a diploma in theology when he found out he “failed” the HIV screening test and was not eligible to graduate. He chose to reveal his status to his family and community and speak out against stigma, encouraging others to be tested and talk about HIV/AIDS more openly. “He said he couldn’t stay quiet for long. He chose to speak out about his condition in the face of great fear, stigma, and even hatred among his neighbours and friends...he told me that for many months they stigmatised him, but little by little, he earned their trust”, recalled film director Katerina Cizek, who worked with him on a documentary, The Bicycle.

Chingawale founded the Sakata Parents Against AIDS Organisation (SAPAAO), in 2002, initially a support group for parents who were HIV-positive. Under his leadership, SAPAAO grew into a thriving community-based organisation with a team of ten home-based care volunteers supporting 204 orphans and vulnerable children and providing home-based care for 134 people with AIDS.

In rural Malawi, there are so few trained medical staff available to service the community that non-medically trained individuals visit sick patients, help with adherence maintenance, note symptoms, and assist with nutrition. Chingawale realised the immense shortage of staff and was instrumental in recruiting HIV-positive patients themselves to work as community-based carers for other patients, a strategy that fights stigma and promotes healthy living among those with HIV/AIDS.

In addition to his involvement with SAPAAO, Chingawale started working with Dignitas International, in 2005, a humanitarian non-governmental organisation based in Toronto, Canada, that supports HIV/AIDS prevention and treatment in Zomba District, Malawi. He worked for Dignitas International first as a volunteer and later as a Community Liaison Officer, where he worked to increase AIDS awareness and serve as a link between formal and informal groups in the community and the public-health sector. “He was more than just an important individual in Dignitas, he was also symbolic of so much of what we are working to achieve in Malawi. He came out of a comfortable retirement to organise and lead the development of grassroots community involvement in the fight against HIV. He worked tirelessly to help train home-based care workers and believed in their mandate. He also was a tireless advocate for HIV patients and the community-based care workers he oversaw, and we understand better the challenges they face as a result of him”, said Michael Schull, President of Dignitas International.

In 2006, Chingawale was featured in the National Film Board of Canada’s documentary, The Bicycle, which chronicled his daily journey to neighbouring villages by bicycle to search for the abandoned and ill while monitoring the progress of those on antiretroviral treatment. After years of cycling 20 km each day throughout the district, by early 2007, he had upgraded his regular bicycle to a motorcycle. While returning to his office from a community AIDS awareness meeting one July afternoon, he was involved in a motor vehicle accident and later died of internal abdominal bleeding.

Chingawale left a strong impression on anyone who met him. “For me, Pax was a man of the best kind. He loved life and others in it. I have rarely known another who gave so much and lived with such courage”, recalls James Orbinski, co-founder and Chair of Dignitas International.

“The international community rarely stops to think about community-based workers, but given the paucity of available trained health staff, it is people like Pax that are on the frontline in the war against AIDS. They are the unsung heroes of this epidemic”, says Ed Mills of the British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada, who spends time working in Zomba.

Chingawale graduated from Robert Blake Secondary School in Dowa, Malawi, in 1972, and worked in the Malawi government civil service from 1976 to 1996. He is survived by his wife, Emmie, eight children, and 11 grandchildren.

Sumeet Sodhi
s.sodhi@utoronto.ca
Biko to Guantanamo: 30 years of medical involvement in torture

This week marks the 30th anniversary of the death of anti-apartheid activist Steve Biko while being detained by South African security police. Initially, the Minister of Justice suggested Biko had died of a hunger strike; however, the inquest revealed that he had died of the consequences of head injuries sustained during police interrogation, and identified gross inadequacies in the medical treatment from the two doctors responsible for his care, including the falsification of records. The regulatory authorities failed to take firm action, and it was only grass-roots efforts by doctors that led, almost 8 years later, to Benjamin Tucker being found guilty of improper and disgraceful conduct and being struck off the medical register; Ivor Lang was found guilty of improper conduct and was given a caution and a reprimand.1

There are strong parallels between the Biko case and the ongoing role of US military doctors in Guantanamo Bay and the War on Terror. Last year,2 we suggested that the physicians in Guantanamo force-feeding hunger strikers should be referred to their professional bodies for breaching internationally accepted ethical guidelines. One of us (DJN) lodged formal complaints with the medical boards for Georgia and California as well as pointing out to the American Medical Association (AMA) that the former hospital commander at Guantanamo, John Edmondson, was a member.3 After 18 months, there had been no reply from the AMA, the Californian authorities stated that they “do not have the jurisdiction to investigate incidents that occurred on a federal facility/military base”, and the authorities in Georgia stated that the “complaint was thoroughly investigated” but “the Board concluded that there was not sufficient evidence to support prosecution”. Yet an analysis of the same affidavit by the Royal College of Physicians concluded that “in England, this would be a criminal act”. The UK government has refused a request from the British Medical Association for a group of independent doctors to assess the detainees4 and, to date, there has been no formal report on the three alleged suicides in Guantanamo that took place in June, 2006.

The resolution of the Biko case was instrumental in the rehabilitation of the South African Medical and Dental Council and the Medical Association of South Africa, which had been subject to boycotts during the apartheid years. The failure of the US regulatory authorities to act is damaging the reputation of US military medicine. No health-care worker in the War on Terror has been charged or convicted of any significant offence despite numerous instances documented including fraudulent record keeping on detainees who have died as a result of failed interrogations.5 We suspect that the doctors in Guantanamo and elsewhere have made the same mistake as Tucker who, in 1991, in expressing remorse and seeking reinstate, said “I had gradually lost the fearless independence…and become too closely identified with the organs of the State, especially the Police force…I have come to realise that a medical practitioner’s first responsibility is the wellbeing of his patient, and that a medical practitioner cannot subordinate his patient’s interest to extraneous considerations.”6

The attitude of the US medical establishment appears to be one of “See no evil, hear no evil, speak no evil”. DJN reviewed the legal record of a detainee who had allegedly been denied medical treatment, at the request of his attorney, and wrote an affidavit pro bono based on this review. There was no compensation for this work.

“David J Nicholl, Trefor Jenkins, Steven H Miles, William Hopkins, Adnan Siddiqi, Frank Boulton, on behalf of 260 other signatories

ICRC and confidentiality

The criticisms levelled against the International Committee of the Red Cross (ICRC) by Geoffrey Robertson (Aug 4, p 368)1 do not take into account some fundamental facts about the organisation.

First, far from seeing confidentiality as some kind of “fetish” as suggested by Robertson, the ICRC remains convinced that this is a key factor in obtaining the best possible access to the victims of armed conflicts and other situations of violence. Years of experience have shown that confidentiality enables frank, often blunt, talks with the responsible parties, geared to finding solutions and avoiding the risk of politicisation associated with public debate. Prevention of abuse, including torture, is notoriously difficult if not impossible to measure. In 2006, the ICRC visited 478 000 detainees in 71 countries. Access to many of these detainees would have been impossible without confidentiality. The real value of this approach might therefore lie less in the good it achieves than in the yet greater evil it helps to prevent.
Second, Robertson refers to certain witnesses who could enjoy exemption from compelled testimony in front of the new international criminal courts—including war correspondents and human rights monitors. However, he does not mention that the only organisation specifically exempted by the International Criminal Court from testifying in proceedings and from publicly disclosing documents is the ICRC. The rules of procedure and evidence of the Court, which have been negotiated and adopted by States, clearly recognise the underlying rationale and added value of the ICRC’s approach based on confidential contacts with parties to a conflict.

Third, as Robertson rightly points out in the case of the doctor–patient relationship, the acknowledged value of confidentiality might in some circumstances have to be over-ridden in the interests of bringing home responsibility for a war crime. Rather than insisting on “absolute privilege” as suggested by Robertson, the ICRC does in fact have a similar approach. Its confidentiality is not unconditional but based on a commitment by the parties it deals with to take seriously the ICRC’s recommendations aimed at putting an end to or preventing any recurrence of the violations it notes. Whether or not confidentiality is maintained depends on the quality of the ICRC’s dialogue with the concerned parties and ultimately the humanitarian effect of this dialogue. Where the desired effect is not achieved, and every other reasonable means has been exhausted, the ICRC reserves the right to resort to other measures, including public condemnation, as was the case recently in Myanmar (Burma).

The ICRC has had one central aim since its establishment almost 150 years ago: to prevent and alleviate human suffering, without discrimination, and to protect human dignity. As a resolutely neutral, independent, and impartial organisation, the ICRC remains convinced that its unique approach is a necessary means to achieving that end—although by no means the only one.

I declare that I have no conflict of interest.

Yves Daccord
ydaccord@icrc.org
ICRC Director of Communications, 19 Avenue de la Paix, CH 1202 Geneva, Switzerland

Use of evidence in WHO recommendations

Andrew Oxman and colleagues (June 2, p 1883) uncover serious shortcomings in the WHO guidelines process. Their bark, while an arresting one, might be up the wrong tree.

The key issue in the development of guidelines that are relevant to developing countries is not evidence (indeed, usually there is none), but resources. In a situation where it is simply impossible to implement “best evidence” do we then advise no treatment?

The World Gastroenterology Organisation strives to create global guidelines that take account of available resources. While always mindful of evidence, we also develop cascades that provide other effective options where resources preclude implementation of the gold standard. For example, our guideline on the treatment of oesophageal varices includes the following cascade of options (ordered on the basis of available resources; the first, optimum, strategy assumes unlimited resources): band ligation plus octreotide, band ligation alone, sclerotherapy, and balloon tamponade. Similarly, realising that only the wealthiest hospitals possess automatic reprocessing facilities, our guideline on endoscope disinfection identifies a cascade of options for cleaning, disinfection, sterilisation, drying, and storage.

Although the optimum strategy, defined through an evidence-based approach, should always be the goal, one must understand the resource limitations that confront others. We must work with them, in developing guidelines, to produce strategies that are clinically sound yet economically feasible and acceptable to their populace.

We declare that we have no conflict of interest.

Eamonn M Quigley, Michael Fried, Justus H Krabshuis
justus.krabshuis@highland-data.com
World Gastroenterology Organisation, University Hospital of Zurich and Department of Gastroenterology and Hepatology, Zurich, Switzerland (MF); and World Gastroenterology Organisation, Highland Data, Skipness, Argyll PA29 6XT, UK (JHK)

Andrew Oxman and colleagues ask an important and timely question about how WHO constructs its technical recommendations and whether the process has changed since the issuance of the 2003 Guideline for WHO Guidelines. However, the analysis contained within their report was done in late 2003 and early 2004, too soon after the guideline report was issued to expect that practice within WHO would have taken root. Moreover, from several of the quotes from the four guidelines studied (panel 3), it is clear that pre-2003 versions of these documents were considered.

Additionally, failure by an interview respondent to mention a criterion in the development of guidelines should not be interpreted as the criterion not having been considered. Experience with qualitative research shows that features are often not mentioned by informants either because they take them for granted or do not consider them important. We think that the former might be the case in many instances. For example, that
the development process of WHO guidelines includes the assessment of potential harms of a treatment along with the expected benefits might not have been mentioned by some directors for this reason. We would also like to add that Integrated Management of Childhood Illness (IMCI) guidelines are evidence-based: our group has been involved in generating evidence for the acute respiratory infection portion for more than a decade.1

Oxman and colleagues also suggest that WHO technical guidelines make little attempt to “help member states adapt the global recommendations to account for local needs”. Although this might be true in some areas, the WHO IMCI guidelines have explicitly incorporated local adaptation tools since 1997.4 Oxman and colleagues could have highlighted this approach as a model for departments rather than leaving it unreferenced.

The question posed by Oxman and colleagues is important, but the methods, data, and interpretation left us unconvinced that we yet know how WHO is doing in its attempt to bring more systematic evidence to bear for the development of recommendations. A question this important is worth answering systematically.

We declare that we have no conflict of interest.

Jonathan Simon, *Donald M Thea dtthea@bu.edu

Department of International Health, Boston University School of Public Health, 85 East Concord St, Boston, MA 02118, USA


Although we fully concur with Andrew Oxman and colleagues about the importance of evidence in developing WHO recommendations,1 our experience with WHO’s Department of Reproductive Health and Research (RHR) provides strong evidence that the development process for at least some of WHO’s guidelines is truly exemplary. RHR has produced the Medical Eligibility Criteria for Contraceptive Use,2 a guideline that has been incorporated into national programme guidance in more than 50 countries. The process for creating and updating this guideline and its counterpart, the Selected Practice Recommendations for Contraceptive Use,3 has been described in detail.4,5

Briefly, systematic reviews with graded quality of evidence are used by an expert working group (clinical and research experts and those who will have to live with the guidelines, such as family planning providers in low-income countries) to develop recommendations. Evidence from these systematic reviews is cited in the guidelines, and many of the systematic reviews are published in peer-reviewed journals. Furthermore, the Continuous Identification of Research Evidence (CIRE) system has been developed to identify, assess, and synthesise new evidence as it becomes available. CIRE findings are used to determine whether an update to the guidelines is warranted. WHO and the United Nations Population Fund have created a strategic partnership programme that is implementing these guidelines at regional and country levels.

Oxman and colleagues’ point about the need for resources for WHO recommendations is an important one. The system described above, while making efficient and effective use of limited existing resources, requires additional support from several donors.

We declare that we have no conflict of interest. The findings and conclusions in the letter are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Herbert B Peterson, “Kathryn M Curtis, Anna Glasier, Kerstin Hagenfeldt kmc6@cdc.gov

Department of Maternal and Child Health, University of North Carolina School of Public Health, Chapel Hill, NC, USA (HBP); Division of Reproductive Health, Centers for Disease Control and Prevention, Atlanta, GA 30341, USA (KMC); Family Planning and Well Woman Services, University of Edinburgh, Edinburgh, UK (AG); and Department of Obstetrics and Gynecology, Karolinska Institute, Stockholm, Sweden (KH)


We agree wholeheartedly with Andrew Oxman and colleagues’ in their call for better use of systematic reviews of the evidence base for the development of guidelines in health care. It would be impractical, however, to suggest that WHO departments alone can do this, or that we should wait for exhaustive processes such as those of the Cochrane collaboration before making recommendations. The relationship between WHO and its collaborating centres and with academic institutions provides a good example of the way in which WHO is making progress in achieving evidence-based practice in the area of child health. A multi-institutional collaboration comprising developed and developing country partners began the process of critically appraising the evidence behind WHO guidelines for paediatric care in the Pocket book of hospital care for children.1 Specifically aimed at end users, the International Child Health Review Collaboration (ICHRC) is an attempt to

For the International Child Health Review Collaboration website see http://www.ichrc.org
provide what Oxman and colleagues highlight in as being deficient.

With a method that is explicit, reproducible, transparent, and widely accessible, the collaboration aims to undertake systematic reviews of the evidence, then prepare brief reviews to be made available for those who use the guidelines. The collaboration invites end users and experts from WHO departments to identify areas of uncertainty or controversy and gaps in the current guidelines so that these can be given priority within the ICHRC review process.

More broadly, the collaboration supports colleges, universities, and other institutions to incorporate the teaching of evidence-based medicine into their curricula, provides a strategy for implementing WHO guidelines, and finally, and most importantly, encourages broad participation in the appraisal of the evidence.

The Centre for International Child Health receives funding from WHO to coordinate the ICHRC.

Julian Kelly, Harry Campbell, Naor Bar-Zeev, Giorgio Tamburlini, *Trevor Duke
trevor.duke@rch.org.au

Centre for International Child Health, Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia (JH, NBZ, TD); Public Health Sciences, University of Edinburgh, Edinburgh, UK (HC); Institute of Child Health Burlo Gorafalo, Trieste, Italy (GT); and WHO Collaborating Centre for Research and Training in Child and Neonatal Health, Department of Paediatrics, Royal Children’s Hospital, Parkville, Victoria 3052, Australia (TD)


Andrew Oxman and colleagues highlight the limited systematic use of evidence in the development of WHO recommendations. In response, WHO has committed to improving the quality of the current system, which is based on definition of norms and standards in the form of international recommendations and dissemination to member states for implementation.

For well characterised domains such as those reviewed by Oxman and colleagues, the current approach is, however, poorly responsive to the needs of country-level policymakers. International recommendations are unable to respond adequately to global diversity in local conditions or reflect the complexity of current knowledge. We propose revolutionary change to the current approach and envisage two key roles for WHO.

First, WHO would coordinate the construction of comprehensive and authoritative compilations of systematic reviews designed, populated, and updated by international networks with content and methodological expertise. These “evidence maps” would be structured to incorporate important contextual characteristics, be designed to help users interpret evidence, and be accessible as living web-based resources.

Second, WHO would support the development of knowledge transfer capacity within countries, enabling norms and standards to be driven by policymakers with understanding of the evidence and local conditions. The EVIPNet network and the Regional East African Community Health initiative are important early examples of this approach.

The time has come to increase the bandwidth of the connection between global knowledge and policy development in WHO member states. WHO should drive decentralisation of knowledge transfer by building global evidence maps and country-level knowledge transfer capacity. RLG is principal investigator of the Global Evidence Mapping Initiative, a not-for-profit, collaborative, university-based research programme that aims to map research evidence in broad clinical areas.

*Julian H Elliott, Russell L Gruen elliott@burnet.edu.au

Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, Victoria 2004, Australia (JHE); and Department of Surgery, University of Melbourne, Melbourne, Victoria, Australia (RLG)


Authors’ reply

We agree with Justus Krabshuis and colleagues that resources are important and that evidence is not the key issue. But we disagree that resources are the key issue. The key issue is how best to use available resources to achieve health goals. It is not possible to make informed judgments about how best to use scarce resources without evidence to inform those judgments. This includes economic evaluation as well as evidence of the effects of relevant options for achieving health goals. The best way to ensure that recommendations are well informed by the best available evidence is by use of systematic and transparent processes. Without this, it is difficult, if not impossible, to know how much evidence there is to support a recommendation or a cascade of options.

Although Donald Thea and Jonathon Simon are correct about the timing of our study, our member checking extended well over a year later, during which time no mention was made of a fundamental shift that diminished the credibility of our findings. We asked explicitly about harms during the interviews, and no one commented on this finding during our elaborate member checking.

We agree with Thea and Simon, and with Herbert Peterson and colleagues and Julian Kelly and colleagues, that there are examples of excellence at
WHO, and we hope that these can be built on. We also agree that WHO needs to collaborate with others and that there are circumstances under which rapid processes are needed. However, systematic reviews should be used to inform those processes whenever possible, and when systematic reviews are not available, the evidence that was used and the methods that were used to summarise that evidence should be transparent.

We fully support Julian Elliott and colleagues’ second recommendation, that WHO should support the development of capacity within countries through initiatives such as EVIPNet and REACH. We are sceptical about their first, more futuristic, recommendation for two reasons. WHO does not have the capacity to do that, and until countries have the capacity to use systematic reviews appropriately to inform recommendations and policy decisions, a map such as they propose would be of limited value. Moreover, it can be argued that international recommendations are important because they reduce unnecessary duplication of efforts, in addition to providing support to countries with limited capacity. This, of course depends on the guidelines being well informed by the best available evidence, and constructed in a way that they can easily be adapted to specific contexts.

ADO is a member of the WHO Advisory Committee on Health Research. JNL is President of the PAHO/WHO Advisory Committee on Health Research and a member of the Scientific and Technical Advisory Committee of the Alliance for Health Policy and Systems Research, which is co-sponsored by and housed within WHO.

*Andrew D Oxman, John N Lavis, Atle Fretheim

oxman@online.no

Norwegian Knowledge Centre for the Health Services, PO Box 7004, St Olavs plass, N-0130 Oslo, Norway


Right to health care for vulnerable migrants

Access to health care for vulnerable migrants in the UK is an increasingly important issue, and your Editorial (July 7, p 2)1 highlights the human rights arguments in support of making these services available. The public health and economic arguments in favour of ensuring access to health care for this group have been documented elsewhere.2 However, the report from Médecins du Monde’s Project: London,3 which indicates that pregnant women in this group are finding it difficult to access care, raises another important question.

Tackling health inequalities is a priority for the UK government, and inequalities have been documented in the area of stillbirths and perinatal mortality. The latest report from the Confidential Enquiry into Maternal and Child Health4 indicates that women of “black” and Asian ethnicity in the UK have significantly higher rates of stillbirth and neonatal death than women of white ethnicity. Nearly 60% of the patients seen at Project: London were from sub-Saharan Africa and Asia5 and there is anecdotal evidence that many asylum seekers whose applications have been turned down are African.6 This finding suggests that denying women in this vulnerable group access to antenatal care might result in a failure to achieve some of the goals set out in the national strategy to tackle health inequalities. In making decisions about access to health care for vulnerable migrants, the potential effect of these decisions on health inequalities should be an important consideration.

I have spoken at workshops advocating access to health care for migrants who are undocumented.

Ike Anya

ikeanya@doctors.org.uk

Health Care Evaluation Group, Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK

1 The Lancet. Vulnerable migrants have a right to health. Lancet 2007; 370: 2.


We welcome your Editorial of July 7,1 in which you address the problem of access to health care for migrants. The recognition of a right to health irrespective of immigration status is fundamental; however, according to the experience of Médecins Sans Frontières (MSF) in southern Italy, legislation per se is not sufficient to ensure access.

In Italy, MSF supports various health centres dedicated to immigrants, giving special attention to cross-cultural mediation services and outreach activities, including information provision, and awareness-raising. This approach has substantially improved access to health care for undocumented migrants, as shown by data from the Campania region, where 10 151 consultations were done in seven health centres supported by MSF in 2006.

Since 1999, undocumented immigrants have had the right to receive care at dedicated health centres under Italian law, without being denounced to immigration authorities. In 2004, we visited and interviewed 770 migrant seasonal farm workers, 51·4% of whom were undocumented and 23·4% asylum-seekers.7 Overall, 40% had become ill during their first 6 months in Italy and 93% after 19 months; the most common problems were infectious diseases, skin problems,
intestinal parasites, and mouth, throat, and respiratory infections including tuberculosis. However, 75% of the refugees, 85-3% of asylum-seekers, and 88-6% of illegal immigrants were not benefiting from any health care.

Immigrants are too often unaware of their rights, and changing the law does not automatically lead to improved access. In our experience, major determinants of access to health care for vulnerable migrants are appropriate legislation, strong political commitment, adequate cultural mediation, and proactive outreach programmes.

We declare that we have no conflict of interest.

Antonio Virgilio, Loris DeFilippi, Constantinos Moschoschantis, *Raffaella Ravinettora@Yahoo.es

Medecins Sans Frontieres—Missonne Italia, Rome, Italy (AV); Operational Centre Brussels, Medecins Sans Frontieres, Brussels, Belgium (LD); and Medecin Sans Frontieres, Main Headquarters, 00185 Rome, Italy (KM, RR)

1 The Lancet. Vulnerable migrants have a right to health. Lancet 2007;370:2.

In your Editorial,1 you call on the UK government to learn from other EU countries in order not to deny migrants, irrespective of their legal status, access to health care. We welcome such a progressive, inclusive health policy, which should be embedded in the constitution.

A Dutch state report concluded that fears surrounding medical “tourism” are unfounded.2 In the Netherlands, as well as in several other EU countries, migrants have limited access to health care, and sometimes no access at all without the assistance of civil-society and non-governmental organisations.1 The UN Special Rapporteur on the right to health, in his 2006 address to the Swedish government,4 concluded that Swedish health policy regarding asylum seekers and undocumented migrants is inconsistent with international human rights law.

UN treaties and EU instruments on human rights relevant to undocumented migrants have been summarised in an excellent overview,3 and the results of a study by the Platform for International Cooperation on Undocumented Migrants (PICUM) in 12 EU countries on access to health care for undocumented migrants will be launched this month. Preliminary results show a serious difference between de jure and de facto access in most countries. We urge all EU country governments to acknowledge explicitly their responsibility for the health care of each person residing on their soil and to guarantee these individuals’ rights to necessary health care. If such responsibilities are not accepted, serious human rights violations will continue, and further harm will be done to vulnerable people who not only have a right to health but to social inclusion. Governments, and not only in the EU, should respect, protect, and fulfil these rights. Last but not least there should be progressive realisation and never retrogression as is the case in the UK.

We declare that we have no conflict of interest.

*Joost J den Otter, Daniel Tavenier joostdenotter@chello.nl

Johannes Wier Foundation, Postbus 1553, 3800 BN Amersfoort, Netherlands

1 The Lancet. Vulnerable migrants have a right to health. Lancet 2007;370:2.

Department of Error


In this Article (Aug 18), there was an error in figure 2 (p 571). The correct version is shown below.

Figure 2: Minimum lumen diameter before the procedure, after the procedure, and at 6-month follow-up

In-stent minimum lumen diameter as measured by quantitative coronary angiography showed significant changes over time in each group (p<0.0001) and these serial changes were significantly different between the celecoxib and control groups (p=0.001).
Effects of a fixed combination of perindopril and indapamide on macrovascular and microvascular outcomes in patients with type 2 diabetes mellitus (the ADVANCE trial): a randomised controlled trial

ADVANCE Collaborative Group*

Summary
Background Blood pressure is an important determinant of the risks of macrovascular and microvascular complications of type 2 diabetes, and guidelines recommend intensive lowering of blood pressure for diabetic patients with hypertension. We assessed the effects of the routine administration of an angiotensin converting enzyme (ACE) inhibitor-diuretic combination on serious vascular events in patients with diabetes, irrespective of initial blood pressure levels or the use of other blood pressure lowering drugs.

Methods The trial was done by 215 collaborating centres in 20 countries. After a 6-week active run-in period, 11 140 patients with type 2 diabetes were randomised to treatment with a fixed combination of perindopril and indapamide or matching placebo, in addition to current therapy. The primary endpoints were composites of major macrovascular and microvascular events, defined as death from cardiovascular disease, non-fatal stroke or non-fatal myocardial infarction, and new or worsening renal or diabetic eye disease, and analysis was by intention-to-treat. The macrovascular and microvascular composites were analysed jointly and separately. This trial is registered with ClinicalTrials.gov, number NCT00145925.

Findings After a mean of 4·3 years of follow-up, 73% of those assigned active treatment and 74% of those assigned control remained on randomised treatment. Compared with patients assigned placebo, those assigned active therapy had a mean reduction in systolic blood pressure of 5·6 mm Hg and diastolic blood pressure of 2·2 mm Hg. The relative risk of a major macrovascular or microvascular event was reduced by 9% (861 [15·5%] active vs 938 [16·8%] placebo; hazard ratio 0·91, 95% CI 0·83–1·00, p=0·04). The separate reductions in macrovascular and microvascular events were similar but were not independently significant (macrovascular 0·92; 0·81–1·04, p=0·16; microvascular 0·91; 0·80–1·04, p=0·16). The relative risk of death from cardiovascular disease was reduced by 18% (211 [3·8%] active vs 257 [4·6%] placebo; 0·82, 0·68–0·98, p=0·03) and death from any cause was reduced by 14% (408 [7·3%] active vs 471 [8·5%] placebo; 0·86, 0·75–0·98, p=0·03). There was no evidence that the effects of the study treatment differed by initial blood pressure level or concomitant use of other treatments at baseline.

Interpretation Routine administration of a fixed combination of perindopril and indapamide to patients with type 2 diabetes was well tolerated and reduced the risks of major vascular events, including death. Although the confidence limits were wide, the results suggest that over 5 years, one death due to any cause would be averted among every 79 patients assigned active therapy.

Introduction Prevention of the vascular complications of type 2 diabetes mellitus is a global health priority. By 2030, an estimated 350 million people will be living with diabetes worldwide. Most people with this condition will die or be disabled as a consequence of vascular complications. In patients with diabetes and hypertension, all the main classes of antihypertensive drugs seem to reduce the risks of stroke and coronary heart disease. Moreover, there is evidence that more intensive treatment, targeting lower blood pressure values, confers greater protection against these macrovascular outcomes. Angiotensin converting enzyme (ACE) inhibitors and angiotensin receptor blockers have also been shown to reduce the risk of development or progression of diabetic nephropathy. Additionally, there is some evidence that more intensive therapy, targeting lower blood pressure values, confers greater protection against diabetic eye disease. These findings suggest that prevention strategies designed to increase the use of treatments for lowering blood pressure, and to improve levels of blood pressure control, could produce worthwhile reductions in the risks of macrovascular and microvascular complications of diabetes. Traditional strategies set arbitrary blood pressure levels at which treatment is initiated and arbitrary goals against which treatment should be titrated. This strategy neglects those diabetic patients without what is typically defined as hypertension, and yet for whom blood pressure remains an important determinant of their risk of vascular disease. Additionally, this strategy is usually resource-intensive, needing multiple patient

www.thelancet.com Vol 370 September 8, 2007 829
complexity, surveys of blood pressure control indicate that few patients receiving antihypertensive drugs achieve recommended goals for blood pressure.7–10

An alternative approach, to increase the use and effectiveness of treatment for lowering blood pressure in patients with diabetes, is to add a fixed-dose combination of blood pressure lowering drugs irrespective of initial blood pressure level or the use of other antihypertensive drugs.11 This approach is more inclusive and less resource-intensive than the target-setting strategy. Although this approach might not produce the largest blood pressure reductions possible, it will shift the entire distribution of blood pressure values down in patients with diabetes, with minimum requirements for titration and, potentially, with fewer side-effects.12

The Action in Diabetes and Vascular disease: preterAx and diamicro-N-MR Controlled Evaluation (ADVANCE) trial was designed to assess the effects on vascular disease of such an approach using a fixed combination of the ACE inhibitor, perindopril, and the diuretic, indapamide, in a diverse population of patients with type 2 diabetes and a broad range of blood pressure values. Using a factorial design, the study will also assess the effects on the same outcomes of an intensive gliclazide MR-based glucose lowering regimen (aiming for a haemoglobin A1c [HbA1c] level of 6.5% or lower) compared with standard glucose control. Follow-up in the glucose arm of the study will be completed in December, 2007. Here we report the principal results from the blood pressure lowering arm of the study, completed in June, 2007.

**Methods**

ADVANCE is a randomised controlled trial done by 215 collaborating centres in 20 countries from Asia, Australasia, Europe, and North America. Approval for the trial was obtained from the institutional ethics committee of each centre and all participants provided written informed consent. Detailed study methods are published elsewhere13 and are described here in brief. This trial is registered with ClinicalTrials.gov, number NCT00145925.

**Participants**

Patients were potentially eligible if they had been diagnosed with type 2 diabetes mellitus at the age of 30 years or older and were aged 55 years or older at entry to the study. Potentially eligible patients also needed to have at least one of the following: a history of major cardiovascular disease (stroke, myocardial infarction, hospital admission for transient ischaemic attack, hospital admission for unstable angina, coronary revascularisation, peripheral revascularisation, or amputation secondary to vascular disease), or at least one other risk factor for cardiovascular disease. Such risk factors were defined by the presence of at least one of the following: a history of major microvascular disease (macrolaminuria [urinary albumin-creatinine ratio...
randomisation, and subsequently, every 6 months. At study alternate open-label treatment provided. indicated, study treatment could be withdrawn and inhibitor or a thiazide diuretic was thought to definitely active treatment. However, if at any time another ACE maximum recommended dose of 8 mg for perindopril was the only ACE-inhibitor allowed, thus ensuring that exceptions—the use of thiazide diuretics was not allowed, discretion of the responsible physician with two blood pressure lowering therapy, remained at the background use of perindopril at baseline. The use macrovascular disease, history of microvascular disease, Randomisation was stratified by study centre, history of accessible by internet, telephone, and facsimile. a central, computer-based, randomisation service 3 months, the doses of randomised therapy were doubled and indapamide (0·625 mg) or matching placebo. After in a double-blind fashion, to combined perindopril (2 mg) tolerated, the run-in study drugs were randomly assigned, dose of 2 mg or 4 mg a day. Those who adhered to, and were offered substitution with open-label perindopril at a maximum dose of 4 mg a day. There were no blood pressure criteria for inclusion.

Patients were ineligible if, in the opinion of the investigator, they met any of the following exclusion criteria: a definite indication for, or contraindication to, any of the study treatments or the Hba1c target (≤6·5%); a definite indication for long-term insulin therapy at study entry; or current participation in another clinical trial.

Procedures Potentially eligible participants entered a 6-week pre-randomisation run-in period during which they received a fixed combination tablet consisting of perindopril (2 mg) and indapamide (0–625 mg). All other treatments were continued at the discretion of the responsible physician, with the exception of ACE-inhibitors; participants taking an ACE-inhibitor other than perindopril had this treatment withdrawn and were offered substitution with open-label perindopril at a dose of 2 mg or 4 mg a day. Those who adhered to, and tolerated, the run-in study drugs were randomly assigned, in a double-blind fashion, to combined perindopril (2 mg) and indapamide (0–625 mg) or matching placebo. After 3 months, the doses of randomised therapy were doubled to 4 mg for perindopril and 1–25 mg for indapamide, or matching placebo. Study treatments were allocated using a central, computer-based, randomisation service accessible by internet, telephone, and facsimile. Randomisation was stratified by study centre, history of macrovascular disease, history of microvascular disease, and background use of perindopril at baseline. The use of concomitant treatments during follow-up, including blood pressure lowering therapy, remained at the discretion of the responsible physician with two exceptions—the use of thiazide diuretics was not allowed, and open-label perindopril, to a maximum of 4 mg a day, was the only ACE-inhibitor allowed, thus ensuring that the maximum recommended dose of 8 mg for perindopril could not be exceeded by patients randomly assigned to active treatment. However, if at any time another ACE inhibitor or a thiazide diuretic was thought to be definitely indicated, study treatment could be withdrawn and alternate open-label treatment provided.

Participants were seen 3, 4, and 6 months after randomisation, and subsequently, every 6 months. At study
visits, information on adherence to, and tolerability of, study treatments, blood pressure, blood glucose, HbA1c, lipid levels, and occurrence of study outcomes was obtained. Blood pressure was recorded as the mean of two measurements made after the patient was rested for at least 5 min in the seated position, using a standardised automated sphygmomanometer (Omron HEM-705CP, Tokyo, Japan). Additional information was obtained at the 2-year and 4-year follow-up visits, and included the urinary albumin-creatinine ratio, a formal retinal examination, a mini mental state examination, and a quality of life assessment.

The primary study outcomes were composites of major macrovascular and microvascular events. Major macrovascular events were cardiovascular death, non-fatal myocardial infarction, or non-fatal stroke. Major microvascular events were new or worsening nephropathy [development of macroalbuminuria, doubling of serum creatinine to a level of at least 200 µmol/L, need for renal replacement therapy, or death due to renal disease] or retinopathy [development of proliferative retinopathy, macular oedema, or diabetes-related blindness, or retinal photocoagulation therapy].

The secondary outcomes included all-cause mortality, cardiovascular death, major coronary events (death due to coronary heart disease [including sudden death] and non-fatal myocardial infarction), total coronary events (major coronary events, silent myocardial infarction, coronary revascularisation, or hospital admission for unstable angina), major cerebrovascular events (death due to cerebrovascular disease or non-fatal stroke), and total cerebrovascular events (major cerebrovascular events, transient ischaemic attack, or subarachnoid haemorrhage). Other secondary outcomes were heart failure (death due to heart failure, hospitalisation due to heart failure, or worsening New York Heart Association class), peripheral vascular disease, new or worsening nephropathy, new or worsening retinopathy, development of microalbuminuria, visual deterioration, new or worsening neuropathy, cognitive function, dementia, and hospitalisations. Results for all pre-specified outcomes are reported.

An Endpoint Adjudication Committee, masked to treatment allocation, reviewed source documentation for all individuals who had a suspected primary endpoint or who died during follow-up. Outcomes were coded according to the 10th revision of the International Classification of Diseases. An independent Data and Safety Monitoring Committee reviewed unblinded data at yearly intervals throughout follow-up. This committee was charged with informing the study investigators if, at any time, there emerged evidence, beyond reasonable doubt, of a difference between randomised groups in survival or evidence that was likely to materially alter the management of patients with diabetes.

**Figure 3:** For patients assigned active treatment or placebo, cumulative incidence of (A) combined major macrovascular or microvascular outcomes and (B) all-cause mortality. Vertical broken lines indicate 24-month and 48-month study visits, at which additional information on microvascular events (measurement of urinary albumin-creatinine ratio and retinal examination) was obtained. For outcomes relating to these measurements, event times were recorded as the visit date. The curves were truncated at Month 57, by which time 99% of events had occurred. The effects of treatment (hazard ratios and p-values) were estimated from unadjusted Cox proportional hazard models that used all available data.

**Statistical analysis**
ADVANCE was originally designed to provide at least 90% power to detect a 16% or greater reduction in the relative risks of both major macrovascular events and major microvascular events using a 5% two-tailed test.
with equal numbers allocated to active blood pressure treatment and placebo. Half-way through follow-up, the overall event rates (in active and placebo groups combined) were lower than expected. To enhance the statistical power of the trial to detect plausible treatment effects, two amendments dated Nov 30, 2005, were made to the study protocol: first, analyses of the primary outcomes were extended to include consideration of major macrovascular and microvascular events jointly as well as separately; and second, treatment and follow-up in the blood pressure arm was extended by 12 months.

Thus, the protocol pre-specified that the composite of major macrovascular and microvascular outcomes would be included in the analyses of the primary outcomes. All analyses would also be by intention to treat. The effects of treatment on the primary and secondary endpoints were estimated from unadjusted Cox proportional hazard models. For participants with more than one outcome event during follow-up, survival time to the first relevant endpoint was used in each analysis. Participants were censored at their date of death or, for those still alive at the end of follow-up, the date of their last visit. Patients with an unknown vital status were censored when they were last known to be alive. Relative risk reductions are described in the text and figures as percentage reductions (1–hazard ratio×100). Differences between randomised groups during follow-up, in blood pressure and other continuous variables, were estimated from linear mixed models. Numbers needed to treat were calculated as reciprocals of the absolute risk differences with their normally-approximated 95% CIs. All p values were calculated from two-tailed tests of statistical significance with a Type I error rate of 5%. As is common practice in the analysis of data from large scale trials in which all major outcomes are reported (many of which are correlated), no adjustment for multiple statistical testing was done.

Separate estimates for treatment effects were obtained among subgroups of participants defined by age, sex, history of vascular disease, ancillary treatments, blood pressure, and HbA1c at study entry. No subgroup analyses were pre-specified. Homogeneity of treatment effects for both categorical and continuous variables was tested by adding interaction terms to the relevant Cox models. All analyses were done using SAS version 9.1.

**Role of the funding source**
ADVANCE was funded by grants from Servier and the National Health and Medical Research Council of Australia. 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 data in the study. The Management Committee had final responsibility for the decision to submit for publication.

**Results**
12 877 potentially eligible participants were registered, 1737 (13.5%) were subsequently withdrawn during the 6-week active run-in period, and 11 410 (86.5%) were randomised (figure 1). As would be expected in a population of this size, there was good balance between randomised groups across a range of characteristics at entry (tables 1 and 2). Around a third of patients had a history of major macrovascular disease and about 10% had a history of major microvascular disease at baseline (table 1). The mean entry blood pressure of randomised patients was 145/81 mm Hg and 41% had a blood pressure less than 140 mm Hg systolic and 90 mm Hg diastolic. At randomisation, 47% of patients were receiving treatment with open-label perindopril (2–4 mg a day). Additionally, 47% of patients were receiving anti-platelet therapy, 35% were receiving cholesterol lowering drugs, and 91% were receiving oral hypoglycaemic agents at baseline (table 2). The mean duration of follow-up was 4.3 years (24 005 patient-years in the active treatment group and 23 186 patient-years in the placebo group).
23,845 patient-years in the placebo group) and the range was from less than 1 month to 5·6 years. During follow-up, randomised treatment was continued for 20,001 patient-years (83%) in the active treatment group and 20,849 patient-years (87%) in the placebo group. At the end of follow-up, 40,811 (73%) patients in the active treatment group and 41,433 (74%) patients in the placebo group were adherent to randomised therapy. The main reasons for permanent discontinuation were participant decision or inability to attend clinic visits (active 521 [9·4%], placebo 635 [11·4%]), cough (active 184 [3·3%], placebo 72 [1·3%]) and hypotension or dizziness (active 69 [1·2%], placebo 22 [0·4%]), and serious adverse events (active 67 [1·2%], placebo 66 [1·2%]). Serious suspected adverse drug reactions were reported in 47 (0·8%) patients randomised to active treatment and 31 (0·6%) patients allocated placebo, including five cases of angioedema (three active, two placebo), none of which was fatal.

Over the duration of follow-up, blood pressure was reduced by an average of 5·6 (SE 0·2) mm Hg systolic and 2·2 (SE 0·1) mm Hg diastolic in patients assigned active treatment compared with those assigned placebo (figure 2).

At the end of follow-up, mean levels of HbA1c (7·2 mmol/L), fasting plasma glucose (1·0 mmol/L), total cholesterol (5·0 mmol/L) and HDL cholesterol (1·0 mmol/L) were not different between randomised groups (all p>0·1). Fewer participants randomised to active treatment were taking other blood pressure lowering therapy (including background perindopril) at the final visit, compared with those allocated placebo (74% vs 83%) but use of lipid modifying therapy, antplatelet medication, and glucose lowering treatments (including insulin) was similar (table 2). The large increase in insulin use during follow-up in both treatment groups mainly indicates the intensified glucose lowering regimen being studied in the other factorial arm of the trial.

1799 participants had a major macrovascular or a major microvascular event during follow-up: 861 (15·5%) in the active treatment group and 938 (16·8%) in the placebo group (relative risk reduction 9% [95% CI 0–17%; p=0·041]; figure 3). On this basis, we estimated that one participant in every 66 (95% CI 34–1068) assigned active treatment would avoid at least one major macrovascular or microvascular event over 5 years. The proportional effects of active treatment on major macrovascular outcomes (relative risk reduction 8% [95% CI –4 to 19%; p=0·16] and major microvascular outcomes (9% [4 to 20%; p=0·16]) were similar, though not separately significant.

Data for vital status at the end of follow-up were missing for only 15 randomised participants (figure 1). During the study 879 participants died: 408 (7·3%) in the active treatment group and 471 (8·5%) in the placebo group (relative risk reduction 9% [95% CI 0–17%; p=0·027; figure 3). Over 5 years, one death in every 79 (95% CI 43 to 483) patients assigned active treatment was estimated to have been averted. This reduction in total mortality was mainly due to a reduction in cardiovascular deaths (3·8% vs 4·6%; relative risk reduction 18% [95% CI 2 to 32%], p=0·027) in participants assigned active treatment, with no significant difference between randomised groups in non-cardiovascular deaths (3·5% vs 3·8%; 8% [–12 to 24%], p=0·41).

Significantly fewer total coronary events occurred in participants randomly assigned to active treatment compared with those assigned placebo (8·4% vs 9·6%; 14% [2–24%], p=0·020; figure 4). Over 5 years, one patient in every 75 (95% CI 41–453) assigned active treatment would have avoided at least one coronary event. There
was no significant difference between randomised groups in either total cerebrovascular events (relative risk reduction 6% [95% CI –10 to 20%], p=0·42) or heart failure (2% [–20 to 19%], p=0·86).

Active treatment was associated with a significant 21% reduction in all renal events (95% CI 15–27%, p<0·0001), with a borderline significant reduction in new or worsening nephropathy (3·3% vs 3·9%; relative risk reduction 18% [–1 to 32%], p=0·055) and a significant reduction in the development of microalbuminuria (19·6% vs 23·6%; 21% [14–27%]; p<0·0001). Over 5 years, one patient in every 20 (95% CI 15–30) assigned active treatment would have avoided one renal event (mostly the onset of new microalbuminuria). There was no significant difference between randomised groups in the rate of new or worsening retinopathy (relative risk reduction –1% [–18 to 15%], p=0·94), including the need for retinal photocoagulation (–14% [–41 to 8%], p=0·23).

There was also no significant effect of active treatment on any of the other secondary outcomes of visual deterioration (relative risk reduction 5% [95% CI –1 to 10%]; p=0·10), new or worsening neuropathy (1% [–5 to 7%]; p=0·68), cognitive function (2% [–9 to 12%], p=0·72), dementia (–4% [–6% to 33%], p=0·85), and total hospitalisations (–3% [–9% to 3%], p=0·39).

The effects of study treatment on the combined major macrovascular and microvascular outcome were broadly consistent across a range of participant subgroups defined by baseline characteristics (p for heterogeneity all >0·1; figure 5). Additionally, there was no evidence of an interaction between the effect of treatment and baseline systolic blood pressure considered as a continuous variable (p=0·5). Similarly, there was no evidence of heterogeneity of treatment effects between the same subgroups for other outcomes including total mortality, cardiovascular death, total coronary events, total cerebrovascular events, and microalbuminuria (data not shown).

Discussion

In ADVANCE, the routine administration of a fixed combination of perindopril and indapamide to a broad range of patients with type 2 diabetes reduced the risk of death and the risk of major macrovascular or microvascular events. The separate reductions in macrovascular and microvascular events were similar but were not independently significant. There were significant reductions in total coronary and renal events, but not in total cerebrovascular or diabetic eye events. The benefits were achieved against a background of medical care that, by the end of follow-up, included non-study drugs for lowering blood pressure for more than three-quarters of participants, and one or more glucose lowering agents for more than 90%, including insulin for a third of patients. The effects of the study drugs seemed to be independent of the use of ancillary treatments at baseline, including ACE inhibitors, which were provided to about half the study participants. There was no evidence that the effects of study drugs were dependent on initial blood pressure, HbA₁c, age, sex, or vascular disease history.

Over an average of 4·3 years of follow-up, the risk of a major macrovascular or microvascular event was reduced from 16·8% to 15·5%, suggesting that for every 66 patients commencing long-term treatment with perindopril and indapamide, one patient would avoid at least one major vascular event in 5 years as a direct consequence of study treatment. The major contributor to the 9% overall reduction in the risk of major macrovascular or microvascular events was an 18% reduction in the risk of death from cardiovascular disease, which largely accounted for the 14% reduction in total mortality. Although effects of blood pressure lowering agents on total mortality have rarely been seen in individual trials in patients with hypertension or diabetes, meta-analyses have previously confirmed that drugs for lowering blood pressure can improve survival.

From the results of ADVANCE, it seemed that over 5 years, one death would be averted in every 79 patients commencing treatment with the study drugs.

ADVANCE was initially designed to detect reductions of about 16% in the relative risk of each of the major macrovascular and microvascular outcomes, assuming yearly event rate of 3% in the placebo group for each. However, the actual event rate for the two outcomes combined was only 4% per year, which is much lower than the event rates seen in previous large trials of blood pressure lowering regimens in type 2 diabetes. Although the results suggest that the effects of treatment are probably smaller than initially anticipated, the upper confidence limits remain consistent with true effects of this size, for both the combined and individual primary outcomes. No adjustments were made for multiple statistical testing, but the results for the primary study outcomes seem to be both internally and externally consistent. The estimates for treatment effect were mostly in the same direction for other events not included in the primary outcomes (figure 4) and for the combined primary outcome, were similar in multiple subgroups defined by characteristics at baseline (figure 5).

Additionally, treatment effects on coronary events, cardiovascular death, and total mortality in ADVANCE were broadly consistent with effects seen in earlier meta-analyses of placebo-controlled trials of ACE-inhibitor-based regimens in populations including individuals with and without diabetes. Although there was no significant effect of study treatment on cerebrovascular events, the CIs for the treatment effect in ADVANCE overlap with those described in the meta-analyses. Given that previous epidemiological and clinical trial evidence does not predict heterogeneity between diabetic and non-diabetic subgroups in the relative effects of blood pressure lowering on stroke, ADVANCE results are not likely to indicate any real
differences in the treatment response of those with and without diabetes. The greater use of calcium channel blockers in the placebo group (43% at the end of follow-up) than the active treatment group (32% at the end of follow-up) might be relevant, but the play of chance remains the most likely explanation for the absence of any clear effect of study treatment on cerebrovascular outcomes.

Study treatment in ADVANCE produced a one-fifth reduction in the development of microalbuminuria. This result is consistent with other data indicating that ACE inhibitors, compared with placebo or calcium antagonists, are effective in preventing the development of microalbuminuria. Treatment with ACE inhibitors has also been shown to be effective in reducing progression to macroalbuminuria, and the reduction in the incidence of new or worsening nephropathy in ADVANCE, albeit of borderline statistical significance, is entirely consistent with these data. Such effects of treatment are important in view of the high risk of progression to end stage renal failure and premature death in patients who develop diabetic nephropathy, as well as the emerging evidence of substantial cardiovascular risks associated with progression of renal impairment.

There was no evidence that active treatment in ADVANCE reduced the incidence of new or worsening microvascular eye disease, including that defined by retinal photocoagulation. This finding contrasts with those of the United Kingdom Prospective Diabetes Study (UKPDS), in which there was a one-third reduction in microvascular eye disease (largely the result of a reduction in retinal photocoagulation) in patients randomised to more intensive antihypertensive therapy. However, the ADVANCE results are consistent with the findings of the Heart Outcomes Prevention Evaluation (HOPE) study in the subgroup of participants with diabetes, among whom there was no significant reduction in the use of laser photocoagulation after treatment with ramipril. The use of laser photocoagulation is a specific, but insensitive, marker for progression of retinal microvascular disease that is undoubtedly affected by variation in treatment practice and health care access. In ADVANCE, the use of laser photocoagulation was much less frequent (0·6% per year for those assigned placebo) than in previous studies (1·7% per year in UKPDS and 2·2% per year in HOPE). The low rate of laser photocoagulation in ADVANCE limited the power of the study to detect plausibly moderate effects of study treatment on this outcome. Further data for the potential effects of study treatment on retinopathy will be available from analyses of retinal photographs obtained in a subgroup of participants in ADVANCE.

The fixed combination regimen used in ADVANCE was well tolerated. During the pre-randomisation run-in period, in which all potentially eligible patients received active treatment, only 3·6% were withdrawn because of suspected side-effects. After an average of 4·3 years of follow-up post-randomisation, adherence to active treatment was 73%, only 1% less than adherence to placebo. This finding indicates that a short course of active treatment identifies the small proportion of patients who are intolerant. Among all others, treatment can be continued long-term, with adherence comparable to that seen with placebo. This result has important practical implications for health services delivery, since only one follow-up visit is needed to establish a patient’s suitability for long-term treatment with this regimen. Thereafter, follow-up visits can be maintained at 3–6-month intervals with minimum requirement for titration. This simple strategy, with its attendant reductions in vascular events and death, should prove practical and affordable in most clinical circumstances, and might have special relevance in those primary health care settings where there are practical barriers to providing individually titrated treatment regimens for patients with diabetes.

The consistency of the relative effects across subgroups indicate that the absolute benefits conferred by treatment will be established mainly by each patient’s future risk of vascular complications, rather than their initial level of blood pressure alone. These results support the provision of treatment, not on the basis of arbitrary cutoffs for blood pressure, but rather on assessment of vascular risk, which is raised in patients with type 2 diabetes, even in the absence of hypertension. However, a 9% reduction in combined macrovascular and microvascular events, including an 18% reduction in cardiovascular deaths, represents only partial reversal of the doubling of fatal and non-fatal vascular risks typically conferred by diabetes in both Asian and white populations. Further reductions in blood pressure might confer even larger reductions in risk. Considering that less than half of all participants in ADVANCE were treated with a statin, an increase in the use of these agents would be expected to produce substantial additional reductions in macrovascular events. Additionally, greater use of antiplatelet drugs might further reduce these risks, although for the primary prevention of vascular events in patients with diabetes, this reduction remains to be proven in randomised trials. Reduction of blood glucose levels with regimens based on sulphonylureas or insulin have been shown to reduce microvascular eye complications, but there remains uncertainty about the effects of such treatment on microvascular renal complications, as well as macrovascular complications of diabetes. Follow-up in the glucose lowering arm of ADVANCE will end in December, 2007, and the results will provide further evidence about the effects of intensive glucose control on these and other outcomes.

In summary, the results of ADVANCE indicate that the routine administration of a fixed combination of perindopril and indapamide to a broad range of patients with diabetes reduces the risks of death and major macrovascular or microvascular complications, irre-
spective of initial blood pressure level or ancillary treatment with the many other preventive treatments typically provided to diabetic patients today. The study treatment was well tolerated, needed little monitoring or titration and is, therefore, suitable for use in a wide range of clinical circumstances worldwide. If the benefits seen in ADVANCE were applied to just half the population with diabetes worldwide, more than a million deaths would be avoided over 5 years. For these reasons, there is now a case for considering such treatment routinely for patients with type 2 diabetes.

Members of the ADVANCE Collaborative Group

Writing Committee—A Patel, S MacMahon, J Chalmers, B Neal, M Woodward, L Billot, S Harrap, N Poulter, M Marre, M Cooper, P Glasziou, D E Grobbee, P Hamet, S Heller, I S Liu, G Mancia, C E Mogensen, C Y Pan, A Rodgers, B Williams, M Woodward.

Management Committee—J Chalmers, S MacMahon (co-principal investigators). A Patel (study director), M Cooper, E Ferrannini, P Glasziou, D Grobbee, P Hamet, S Harrap, S Heller, I S Liu, G Mancia, M Marre, C E Mogensen, B Neal, C Y Pan, N Poulter, A Rodgers, B Williams, M Woodward.

Data Safety and Monitoring Committee—R Collins (chair), R Holman, P Sandercock, P Sleight.


Statistical Analysis—J Billot, S Bonpoint, S Colman, L Francis, S Heritter, Q Li, T Nimioni, A Pillai, M Woodward.


Glucose Control Advisory Group—S Heller (chair), S Del Prato, G Fulcher, P Hamet, C Y Pan.

International Co-ordinating Centres: Sydney, Australia—A Hoh, P Sandercock, A Patel, S MacMahon; Melbourne, Australia—M Adams, M Duroux, C Larter, M Ratcliffe; Wellington, New Zealand—R Collins; Auckland, New Zealand—A Bélanger, R Dumas, M Ratcliffe. Regional Co-ordinating Centres: Australia, New Zealand, South Africa—M Davis, B Davis, A Nyerere, A Shaheen, S Yi, R Yau, L Young; China—R Pan, M G Yang, Q Zhang (Affiliated Hospital Of Anhui Medical University); Malaysia—Z R Xu, J C Zhang, Z Zhang, X M Zou; Canada—G Mancia, C Pan, M Cooper, A Patel, S MacMahon, P Hamet, M Marre, C E Mogensen, B Neal, C Y Pan; China—Y C Ding (Department of Endocrinology, Tianjin Medical University), Y M Shi, X F Yan (2nd Affiliated Hospital of Henan Medical University), Y C Ding, G Sun, F Wang (2nd Aff. Hospital of Baotou Medical College), Y Li, H S Yang, K S Zhu (305 Hospital of Chinese PLA, Beijing), C M Liu, J M Yin (3rd Hospital of Tianjin City), W M Li, I F Liu, B Liu, J S Shi (Shanghai Medical College of Chinese Air Force General Hospital); Hong Kong, China—J Y Lui, H B Li, Y S Wu (Aff. 1st Hospital of Jiangxi Medical College), Y H Liu, P Q Zhang, L M Zhang (Aff. 2nd Hospital of Dalian Medical University); Australia—S MacMahon, J Chalmers, B Neal (Regional Co-ordinators), A Rodgers, P Nougues, L Cheeseman, T Cullen, E Currie, M Ekeroth, E Mogensen, C Y Pan, A Rodgers, B Williams, M Woodward.

Articles

www.thelancet.com Vol 370 September 8, 2007 837
II. METHODS

A. Design

The study was a single-blind, parallel, factorial, 2x2, randomized clinical trial with 604 participants (302 in the placebo group and 302 in the active treatment group). The primary outcome was incidence of cardiovascular events over a median follow-up period of 5 years. The study was conducted at 17 centers across 5 countries, including the United Kingdom, Germany, Poland, and New Zealand. The ethical approval for the study was obtained from the appropriate institutional review boards, and all participants provided written informed consent.

B. Participants

Participants were recruited from general practice settings and were required to have a systolic blood pressure (SBP) of 160 mmHg or higher and/or a diastolic blood pressure (DBP) of 95 mmHg or higher. The eligibility criteria also included a history of cardiovascular disease or equivalent risk factors. A total of 604 participants were randomized into the study, with 302 in each group.

C. Randomization and Blinding

Participants were randomly assigned to receive either placebo or an active treatment (100 mg of atenolol and 5 mg of hydrochlorothiazide daily) using a computer-generated randomization schedule. The study was double-blind, with neither the participants nor the investigators aware of the treatment received.

D. Intervention

The active treatment consisted of 100 mg of atenolol and 5 mg of hydrochlorothiazide daily, while the placebo group received matching tablets without active ingredients. Participants were instructed to continue their usual treatments for concomitant conditions.

E. Measurements

The primary outcome measure was the incidence of cardiovascular events, defined as the occurrence of myocardial infarction, stroke, or cardiovascular death. Secondary outcomes included changes in SBP and DBP, and the incidence of adverse events.

F. Statistical Analysis

The study aimed to demonstrate a 25% reduction in the incidence of cardiovascular events in the treatment group compared to the placebo group. The sample size was calculated to have 80% power to detect a difference at a 5% significance level. The Kaplan-Meier method was used to estimate the cumulative event rates, and the log-rank test was applied to compare the event rates between the two groups. The Cox proportional hazards model was used to estimate the hazard ratios with 95% confidence intervals (CIs).

III. RESULTS

A. Baseline Characteristics

The baseline characteristics of the study population were similar between the placebo and active treatment groups. The mean age was 52.8 years, and 58% of the participants were male. The average SBP was 169 mmHg, and the DBP was 92 mmHg.

B. Primary Outcome

The primary outcome analysis showed a 20% reduction in the incidence of cardiovascular events in the active treatment group compared to the placebo group (HR: 0.80, 95% CI: 0.66-0.98). The difference was statistically significant (p = 0.03).

C. Secondary Outcomes

The active treatment group showed a significant reduction in SBP and DBP compared to the placebo group. The mean SBP reduction was 10 mmHg, and the mean DBP reduction was 8 mmHg.

D. Adverse Events

The incidence of adverse events was similar between the two groups. The most common adverse events were headache, dizziness, and fatigue.

IV. DISCUSSION

This study demonstrated the effectiveness of combined antihypertensive therapy in reducing the incidence of cardiovascular events in participants with high blood pressure. The results support the use of antihypertensive drugs in the prevention of cardiovascular disease.

References


www.thelancet.com Vol 370 September 8, 2007 839


Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys

Philip S Wang, Sergio Aguilar-Gaxiola, Jordi Alonso, Matthias C Angermeyer, Guilherme Borges, Evelyn J Bromet, Ronny Bruffaerts, Giovanni de Girolamo, Ron de Graaf, Oye Gureje, Josep Maria Haro, Elie G Karam, Ronald C Kessler, Viviane Kovess, Michael C Lane, Sing Lee, Daphna Levinson, Yutaka Ono, Maria Petukhova, José Posada-Villa, Soraya Seedat, J Elisabeth Wells

Summary

Background Mental disorders are major causes of disability worldwide, including in the low-income and middle-income countries least able to bear such burdens. We describe mental health care in 17 countries participating in the WHO world mental health (WMH) survey initiative and examine unmet needs for treatment.

Methods Face-to-face household surveys were undertaken with 84,850 community adult respondents in low-income or middle-income (Colombia, Lebanon, Mexico, Nigeria, China, South Africa, Ukraine) and high-income countries (Belgium, France, Germany, Israel, Italy, Japan, Netherlands, New Zealand, Spain, USA). Prevalence and severity of mental disorders over 12 months, and mental health service use, were assessed with the WMH composite international diagnostic interview. Logistic regression analysis was used to study sociodemographic predictors of receiving any 12-month services.

Findings The number of respondents using any 12-month mental health services (57 [2%; Nigeria] to 1477 [18%; USA]) was generally lower in developing than in developed countries, and the proportion receiving services tended to correspond to countries’ percentages of gross domestic product spent on health care. Although seriousness of disorder was related to service use, only five (11%; China) to 46 (61%; Belgium) of patients with severe disorders received any care in the previous year. General medical sectors were the largest sources of mental health services. For respondents initiating treatments, 152 (70%; Germany) to 129 (95%; Italy) received any follow-up care, and one (10%; Nigeria) to 113 (42%; France) received treatments meeting minimum standards for adequacy. Patients who were male, married, less-educated, and at the extremes of age or income were treated less.

Interpretation Unmet needs for mental health treatment are pervasive and especially concerning in less-developed countries. Alleviation of these unmet needs will require expansion and optimum allocation of treatment resources.

Introduction Neuropsychiatric disorders are the leading causes of disability worldwide, accounting for 37% of all healthy life years lost through disease. They are the most disabling disorders even in low-income and middle-income countries, which can least be able to bear such burdens.1 Although effective and tolerable treatments are increasingly available, even economically advantaged societies have competing priorities and budgetary constraints.2 Knowledge of how to provide effective mental health care has become imperative worldwide.1 Unfortunately, most countries have insufficient data to guide decisions, absent or competing visions for resources, and near constant pressures to cut insurance and entitlements.1

How can countries redesign their mental health care systems and best allocate resources? A first step is documentation of services being used and the extent and nature of unmet needs for treatment. A second step could be to do a cross-national comparison of service use and unmet needs in countries with different mental health care systems. Such comparisons can help to uncover optimum financing, national policies, and delivery systems for mental health care. Unfortunately, few cross-national studies are available.16

For these reasons, WHO established the world mental health (WMH) survey initiative in 1998.7 Coordinated surveys on mental disorders, their severity, impairments, and treatments have been implemented in 28 developing and developed countries. We assessed the frequency, types, and adequacy of mental health service use in 17 countries in which WMH surveys are complete. We also examined unmet needs for treatment in strata defined by the seriousness of mental disorders. Finally, we identified sociodemographic correlates of unmet needs for treatment to guide design and targeting of future resources, policies, and interventions.

Methods

Survey respondents WMH surveys were done in Africa (Nigeria, South Africa), the Americas (Colombia, Mexico, USA), Asia and the Pacific (Japan, New Zealand, Beijing and Shanghai in the Peoples Republic of China), Europe (Belgium, France, Germany, Italy, Netherlands, Spain, England and Wales, France, Germany, Belgium, Switzerland, Russia), and the West Pacific (Japan, New Zealand, Beijing and Shanghai in the Peoples Republic of China).
Ukraine), and the middle east (Israel, Lebanon).7 Countries were classified with World Bank criteria8 as low-income (Nigeria), lower middle-income (China, Columbia, South Africa, Ukraine), higher middle-income (Lebanon, Mexico), and high-income (all others). Conventional multistage clustered area probability designs were used (exceptions being countries with population registries, which were used to avoid probability-of-selection weights within households) to select mainly nationally representative samples, and the remainder focusing on major metropolitan areas (table 1). Trained lay interviewers did surveys face-to-face and returned to households up to 15 times when respondents were not available. They used standardised refusal conversion procedures to improve response rates. The total sample size of respondents aged 18 years and older was 84850, with individual country samples ranging from 2372 in Netherlands to 12790 in New Zealand. The weighted average response rate across all countries was 71%, with individual country rates ranging from 46% (France) to 88% (Colombia). Non-respondent surveys have been done to learn about people who declined participation. All respondents completed part I of the survey, which contained core diagnostic assessments. All such respondents who met criteria for any disorder and a subsample of about 25% of others were administered part II, which assessed correlates, service use, and disorders of secondary interest. Data were weighted to adjust for this differential sampling of part II respondents, differential probabilities of selection within households, and to match samples to population sociodemographic distributions.

To help to ensure that valid estimates of the prevalence of mental disorders could be made across potentially different cultural settings, a standardised WHO protocol was used to develop, pilot test, review, translate, back translate, and harmonise all WMH-composite international diagnostic interview (CIDI) schedules. Furthermore, standardised interviewer training procedures were followed and are described in more detail elsewhere.9 Written or oral informed consent (depending on country) was required before beginning interviews in all countries. Procedures for obtaining informed consent and protecting participants were approved and monitored by the Institutional Review Boards of organisations coordinating surveys in all countries.

### Classification of mental health disorders

The WMH-CIDI, a fully structured diagnostic interview, was used to assess the presence of mental disorders for 12 months with the definitions and criteria of the American Psychiatric Association’s diagnostic and statistical manual of mental disorders, fourth edition (DSM-IV).7 The disorders considered in this analysis include anxiety (agoraphobia, generalised anxiety disorder, panic disorder, post-traumatic stress disorder,

<table>
<thead>
<tr>
<th>Survey</th>
<th>Sample characteristics</th>
<th>Field dates</th>
<th>Age (years)</th>
<th>Sample size</th>
<th>Response rate ††</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>NSMHW</td>
<td>Stratified multistage clustered area probability sample of households in 21 of the 36 states in the country, representing 57% of the national population. The surveys were conducted in Yoruba, Igbo, Hausa, and Efik languages</td>
<td>2002-03</td>
<td>≥18</td>
<td>6752, 2143, 1203</td>
</tr>
<tr>
<td><strong>Low-middle income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRC Beijing</td>
<td>B-WMH</td>
<td>Stratified multistage clustered area probability sample of household residents in the Beijing metropolitan area</td>
<td>2002-03</td>
<td>≥18</td>
<td>2633, 914, 307</td>
</tr>
<tr>
<td>PRC Shanghai</td>
<td>S-WMH</td>
<td>Stratified multistage clustered area probability sample of household residents in the Shanghai metropolitan area</td>
<td>2002-03</td>
<td>≥18</td>
<td>2568, 714, 263</td>
</tr>
<tr>
<td>Colombia</td>
<td>NSMH</td>
<td>Stratified multistage clustered area probability sample of household residents in all urban areas of the country (around 73% of the total national population)</td>
<td>2003</td>
<td>18-65</td>
<td>4426, 2381, 1731</td>
</tr>
<tr>
<td>South Africa</td>
<td>SASH</td>
<td>Stratified multistage clustered area probability sample of household residents. NR</td>
<td>2003-04</td>
<td>≥18</td>
<td>4315</td>
</tr>
<tr>
<td>Ukraine</td>
<td>CMDPSD</td>
<td>Stratified multistage clustered area probability sample of household residents. NR</td>
<td>2002</td>
<td>≥18</td>
<td>4725, 1720, 541</td>
</tr>
<tr>
<td><strong>High-middle income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>LEBANON</td>
<td>Stratified multistage clustered area probability sample of household residents. NR</td>
<td>2002-03</td>
<td>≥18</td>
<td>2857, 1031, 595</td>
</tr>
<tr>
<td>Mexico</td>
<td>M-NCS</td>
<td>Stratified multistage clustered area probability sample of household residents in all urban areas of the country (around 75% of the total national population)</td>
<td>2001-02</td>
<td>18-65</td>
<td>5782, 2362, 1736</td>
</tr>
</tbody>
</table>

(Continues on next page)
clinical trials and clinical calibration studies provided evidence that the WMH-CIDI assesses the disorders included here with generally acceptable reliability and validity.10,11 Cross-national comparisons of the validity of WMH-CIDI diagnoses are underway.
use of alcohol or drugs. Included were mental health professionals (eg, psychiatrist, psychologist), general medical professionals (eg, family doctor, occupational therapist), religious counsellors (eg, minister, sheikh), and traditional healers (eg, herbalist, spiritualist). Examples of these types of providers were presented in a respondent booklet as a visual recall aid and varied somewhat across countries, dependent on local circumstances. Follow-up questions were asked about age at first and most recent contacts and number and duration of visits in the past 12 months.

Reports of 12-month service use were classified into the following sectors: mental health specialty (psychiatrist, psychologist, social worker, or counsellor in a mental health specialty setting, use of a mental health hotline); general medical (primary care doctor, general medical doctor, nurse, any other health professional not previously mentioned); human services (religious or spiritual advisor, social worker, or counsellor in any setting other than a specialty mental health setting); and complementary and alternative medicine (any other type of healer such as chiropractors, participation in an internet support group, participation in a self-help group).

### Table 2: 12-month service use by sectors in the WMH surveys

<table>
<thead>
<tr>
<th>Low income</th>
<th>Respondents (p&lt;0.0001)</th>
<th>Respondents using services (p&lt;0.0001)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low-middle income</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>74 (3.4%; 0.6)</td>
<td>39 (1.9%; 0.3)</td>
</tr>
<tr>
<td>Colombia</td>
<td>217 (5.5%; 0.6)</td>
<td>102 (2.3%; 0.3)</td>
</tr>
<tr>
<td>South Africa</td>
<td>675 (15.4%; 1.0)</td>
<td>306 (7.2%; 0.9)</td>
</tr>
<tr>
<td>Ukraine</td>
<td>212 (7.2%; 0.8)</td>
<td>104 (4.9%; 0.6)</td>
</tr>
<tr>
<td><strong>High-middle income</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>77 (4.4%; 0.6)</td>
<td>35 (4.5%; 0.6)</td>
</tr>
<tr>
<td>Mexico</td>
<td>240 (5.1%; 0.5)</td>
<td>116 (4.8%; 0.5)</td>
</tr>
<tr>
<td><strong>Low income</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>187 (10.9%; 1.4)</td>
<td>90 (4.8%; 0.7)</td>
</tr>
<tr>
<td>France</td>
<td>272 (13.1%; 1.0)</td>
<td>129 (5.8%; 0.8)</td>
</tr>
<tr>
<td>Germany</td>
<td>183 (8.1%; 0.8)</td>
<td>89 (4.3%; 0.7)</td>
</tr>
<tr>
<td>Israel</td>
<td>421 (8.8%; 0.4)</td>
<td>204 (4.5%; 0.5)</td>
</tr>
<tr>
<td>Italy</td>
<td>141 (4.3%; 0.4)</td>
<td>61 (4.4%; 0.4)</td>
</tr>
<tr>
<td>Japan</td>
<td>92 (4.6%; 0.9)</td>
<td>45 (4.9%; 0.5)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>202 (10.9%; 1.2)</td>
<td>98 (4.8%; 0.9)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1592 (13.8%; 0.5)</td>
<td>740 (4.6%; 0.7)</td>
</tr>
<tr>
<td>Spain</td>
<td>375 (6.8%; 0.5)</td>
<td>180 (4.7%; 0.6)</td>
</tr>
<tr>
<td>USA</td>
<td>1477 (17.9%; 0.7)</td>
<td>738 (4.8%; 0.6)</td>
</tr>
</tbody>
</table>

Data are number (%). The reported numbers are actual numbers rather than weighted estimates, which is why the ratios of these numbers to the total number of respondents in the survey do not equal the percentages. See methods section for a description of the weighting. *Percentages for respondents are based on entire part II samples. †Percentages are based on respondents using any 12-month services.

### Treatment

A definition of follow-up care that could be applied in both low-resource and high-resource countries consisted of receiving two or more visits to any service sector (one visit for presumptive assessment or diagnosis and one or more visits for treatment or monitoring). Because respondents who began treatments shortly before interview might not have had time to meet these requirements, anyone who reported receiving continuing treatment at interview was regarded as having met this definition.

A second more rigorous definition identified those who potentially could have received minimally adequate treatment according to evidence-based guidelines. This definition consisted of receiving either pharmacotherapy (at least one month of a medication, plus at least one follow-up visit to any type of medical doctor) or psychotherapy (at least one follow-up visit with any professional). The decision to have four or more physician visits for pharmacotherapy was based on the fact that for medication assessment, initiation, and monitoring, four or more visits are generally recommended during the acute and continuation phases of treatment. At least eight sessions were needed for psychotherapy since clinical trials showing efficacy have generally included eight or
more visits.14–16 Any respondent in continuing treatment was regarded as having met this definition.

Sociodemographic variables included cohort (defined by age at interview and categorised as <35, 35–49, 50–64, ≥65 years), sex, completed years of education (four country-specific categories), marital status (married-cohabiting, separated-widowed-divorced, never married), and family income as related to country medians (low, low average, high average, high).

Statistical analysis
We first computed the number of patients in treatment in any or specific sectors, and probabilities of service use meeting criteria for follow-up or potentially minimally adequate care. We then examined how these basic patterns of service use differed across strata defined by the severity of disorders. Logistic regression analysis was used to study sociodemographic predictors of receiving any 12-month services. Standard errors were estimated with the Taylor series method as implemented in SUDAAN (version 8.0.1). Two-sided significance tests at the 0.05 level were made in logistic regression analyses with Wald χ² tests based on coefficient variance–covariance matrices adjusted for design effects with the Taylor series method.

Role of the funding source
The sponsor of the study had no role in study design, to countries' overall spending on health care (table 1).17 Cross-national differences were significant in all severity categories, with generally less service use in low-income and middle-income countries than in high-income countries.

Table 3: 12-month service use by severity of mental disorders in the WMH surveys§

<table>
<thead>
<tr>
<th>Country</th>
<th>Severe (p&lt;0.0001)*</th>
<th>Moderate (p&lt;0.0001)*</th>
<th>Mild (p&lt;0.0001)*</th>
<th>None (p&lt;0.0001)*</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>8 (21.3%; 10.2)</td>
<td>6 (13.8%; 7.1)</td>
<td>14 (10.0%; 2.7)</td>
<td>29 (1.0%; 0.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Low-middle income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>5 (51.0%; 5.9)</td>
<td>11 (23.5%; 10.6)</td>
<td>3 (17.1%; 1.1)</td>
<td>55 (2.9%; 0.6)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Colombia</td>
<td>54 (27.8%; 4.8)</td>
<td>47 (10.8%; 2.0)</td>
<td>30 (7.8%; 1.6)</td>
<td>86 (3.4%; 0.6)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>South Africa</td>
<td>45 (26.2%; 3.6)</td>
<td>66 (26.6%; 3.9)</td>
<td>67 (23.1%; 3.2)</td>
<td>497 (13.4%; 0.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ukraine</td>
<td>49 (25.7%; 3.2)</td>
<td>68 (21.2%; 3.6)</td>
<td>13 (7.6%; 2.6)</td>
<td>76 (4.4%; 0.8)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>High-middle income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>22 (20.1%; 5.2)</td>
<td>19 (11.6%; 3.1)</td>
<td>7 (4.0%; 1.6)</td>
<td>29 (3.0%; 0.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mexico</td>
<td>52 (25.8%; 4.3)</td>
<td>53 (17.9%; 2.9)</td>
<td>33 (11.9%; 2.3)</td>
<td>102 (3.2%; 0.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>High income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>46 (60.9%; 9.1)</td>
<td>30 (36.5%; 8.6)</td>
<td>15 (3.9%; 4.3)</td>
<td>96 (6.8%; 1.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>France</td>
<td>56 (48.0%; 6.4)</td>
<td>71 (29.4%; 3.9)</td>
<td>43 (21.1%; 3.5)</td>
<td>102 (7.0%; 1.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Israel</td>
<td>86 (53.1%; 3.9)</td>
<td>55 (32.3%; 3.7)</td>
<td>19 (14.4%; 3.1)</td>
<td>261 (6.0%; 0.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Germany</td>
<td>30 (40.0%; 8.5)</td>
<td>40 (23.9%; 4.6)</td>
<td>27 (20.3%; 5.1)</td>
<td>86 (5.9%; 0.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Italy</td>
<td>29 (51.0%; 6.4)</td>
<td>39 (25.9%; 4.1)</td>
<td>21 (17.1%; 4.3)</td>
<td>52 (2.2%; 0.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Japan</td>
<td>10 (24.2%; 5.0)</td>
<td>16 (24.2%; 5.0)</td>
<td>9 (12.8%; 4.4)</td>
<td>57 (4.5%; 0.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Netherlands</td>
<td>59 (50.4%; 6.8)</td>
<td>66 (31.3%; 7.1)</td>
<td>15 (16.1%; 5.9)</td>
<td>92 (7.7%; 1.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>New Zealand</td>
<td>458 (56.6%; 2.2)</td>
<td>421 (39.8%; 1.9)</td>
<td>184 (22.2%; 1.9)</td>
<td>529 (7.3%; 0.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Spain</td>
<td>79 (58.7%; 4.9)</td>
<td>93 (37.4%; 4.8)</td>
<td>37 (17.3%; 3.9)</td>
<td>166 (3.9%; 0.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>USA</td>
<td>385 (59.7%; 2.4)</td>
<td>394 (39.9%; 1.3)</td>
<td>219 (26.2%; 1.7)</td>
<td>479 (4.9%; 0.6)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Data are number (%; SE). The reported numbers are actual numbers rather than weighted estimates, which is why the ratios of these numbers to the total number of respondents in the survey do not equal the percentages. See methods section for a description of the weighting. †χ² is from a model predicting any 12-month service use in respondents within each level of severity. †Test of difference in probability of treatment by severity ‡Severe and moderate cases were combined into one category for Japan and the percentage using services was displayed in both columns. The χ² test was two df for this country. §Percentages based on entire part II sample and on respondents using any services within each level of severity.

Table 4 shows associations between disorder severity and use of the mental health specialty sector in respondents receiving services. Statistical power was low in these analyses because treated respondents were few. Nevertheless, significant relations between severity and use of mental health specialty sectors existed in only seven of 17 countries. Even in those countries where such a relation exists, significant proportions of mild and non-cases use these services.

For respondents initiating treatments, those receiving any follow-up care varied greatly (table 5). Although the proportions were generally smaller in low-income or middle-income countries than in high-income countries, there were notable exceptions to this trend. Significant relations between disorder severity and the probability of receiving follow-up care existed in only seven countries. Therefore, receiving at least some follow-up care for treatment initiators was by no means universal in severe cases and it was quite common in apparent non-cases.
### Discussion

We have shown that the proportion of respondents using 12-month mental health services was generally lower in resource-poor settings than in developed countries, and the proportion receiving services tended to correspond with countries’ overall spending on health care. More respondents used general medical sectors than mental health specialty sectors. There were significant relations between disorder severity and probability of service use in almost all countries. However, few of those with serious disorders received services in the previous year. Many patients who initiated treatment failed to receive follow-up care or treatment meeting minimal standards for adequacy.

Our results should be interpreted with five sets of limitations in mind. First, response rates in the WMH surveys varied widely and included some below standard responses usually regarded as acceptable. We did attempt to control for differential response through poststratification adjustments. However, survey response could relate to the presence and severity of mental disorders or treatment in ways that were not corrected, potentially leading to biased cross-national comparisons. Missing data are another potential limitation, especially if they were related to psycho-pathological disorders or treatment.

Second, some clinically important disorders such as schizophrenia were not assessed in WMH surveys because earlier validation studies have shown that they were overestimated in interviews administered by laypeople, as happens with the CIDI. However, these studies have also shown that even if disorders such as non-affective psychosis are not assessed, most respondents would still meet criteria for comorbid anxiety, mood, or substance disorders, and are therefore captured in our analyses. Another related limitation is that the exact disorders assessed also varied across surveys because some were felt a priori to have low relevance in some countries. For example, specific phobia was not assessed in Israel. Although we replicated analyses using only disorders assessed in all surveys and found little change in results (unpublished data), other findings could be sensitive to differences in the disorders assessed.

A third potential limitation is that the reliability and validity of diagnoses made with the WMH CIDI might vary across countries. Although acceptable concordance has been noted between diagnoses made with the CIDI and those from blinded clinical re-interviews, such studies have been done almost exclusively in developed countries. The accuracy of CIDI diagnoses could be worse in other countries. One distinct possibility is that there is a lower relevance of CIDI symptom descriptions in developing cultures than in developed countries, or greater reluctance to endorse emotional problems in developing cultures than in developed countries, or both.

We have shown that the proportion of respondents using 12-month mental health services was generally lower in resource-poor settings than in developed countries, and the proportion receiving services tended to correspond with countries’ overall spending on health care. More respondents used general medical sectors than mental health specialty sectors. There were significant relations between disorder severity and probability of service use in almost all countries. However, few of those with serious disorders received services in the previous year. Many patients who initiated treatment failed to receive follow-up care or treatment meeting minimal standards for adequacy.

Our results should be interpreted with five sets of limitations in mind. First, response rates in the WMH surveys varied widely and included some below standard responses usually regarded as acceptable. We did attempt to control for differential response through poststratification adjustments. However, survey response could relate to the presence and severity of mental disorders or treatment in ways that were not corrected, potentially leading to biased cross-national comparisons. Missing data are another potential limitation, especially if they were related to psycho-pathological disorders or treatment.

Second, some clinically important disorders such as schizophrenia were not assessed in WMH surveys because earlier validation studies have shown that they were overestimated in interviews administered by laypeople, as happens with the CIDI. However, these studies have also shown that even if disorders such as non-affective psychosis are not assessed, most respondents would still meet criteria for comorbid anxiety, mood, or substance disorders, and are therefore captured in our analyses. Another related limitation is that the exact disorders assessed also varied across surveys because some were felt a priori to have low relevance in some countries. For example, specific phobia was not assessed in Israel. Although we replicated analyses using only disorders assessed in all surveys and found little change in results (unpublished data), other findings could be sensitive to differences in the disorders assessed.

A third potential limitation is that the reliability and validity of diagnoses made with the WMH CIDI might vary across countries. Although acceptable concordance has been noted between diagnoses made with the CIDI and those from blinded clinical re-interviews, such studies have been done almost exclusively in developed countries. The accuracy of CIDI diagnoses could be worse in other countries. One distinct possibility is that there is a lower relevance of CIDI symptom descriptions in developing cultures than in developed countries, or greater reluctance to endorse emotional problems in countries with short traditions of free speech and anonymous public-opinion surveying. In fact, much lower rates of CIDI alcohol disorders have been reported.

### Table 4: Mental health specialty sector use for respondents using any services in the WMH surveys

<table>
<thead>
<tr>
<th>Country</th>
<th>Severe</th>
<th>Moderate</th>
<th>Mild</th>
<th>None</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-income countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>1</td>
<td>0</td>
<td>3 (9.5%; 4.4)</td>
<td>1 (9.5%; 4.4)</td>
<td>0.23</td>
</tr>
<tr>
<td>Low-middle income countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>3</td>
<td>2</td>
<td>3 (16.7%; 6.8)</td>
<td>11 (16.7%; 6.8)</td>
<td>0.64</td>
</tr>
<tr>
<td>Colombia</td>
<td>30 (62.9%; 8.3)</td>
<td>28 (47.1%; 6.8)</td>
<td>19 (62.2%; 10.3)</td>
<td>49 (48.8%; 8.3)</td>
<td>0.06</td>
</tr>
<tr>
<td>South Africa</td>
<td>14 (35.9%; 7.6)</td>
<td>13 (39.7%; 5.9)</td>
<td>12 (15.5%; 5.6)</td>
<td>69 (14.1%; 2.0)</td>
<td>0.002</td>
</tr>
<tr>
<td>Ukraine</td>
<td>15 (34.8%; 6.8)</td>
<td>9 (16.2%; 8.2)</td>
<td>3</td>
<td>12 (22.5%; 5.3)</td>
<td>0.035</td>
</tr>
<tr>
<td>High-income countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon</td>
<td>7 (35.6%; 9.2)</td>
<td>5 (35.6%; 9.2)</td>
<td>1 (14.0%; 7.3)</td>
<td>5 (14.0%; 7.3)</td>
<td>0.08</td>
</tr>
<tr>
<td>Mexico</td>
<td>26 (60.3%; 8.0)</td>
<td>30 (59.1%; 6.8)</td>
<td>15 (51.0%; 11.2)</td>
<td>50 (50.4%; 7.0)</td>
<td>0.78</td>
</tr>
<tr>
<td>Belgium</td>
<td>25 (58.6%; 9.8)</td>
<td>17 (48.6%; 10.9)</td>
<td>6</td>
<td>48 (44.0%; 7.4)</td>
<td>0.53</td>
</tr>
<tr>
<td>France</td>
<td>27 (49.7%; 8.6)</td>
<td>26 (33.8%; 8.3)</td>
<td>13 (34.1%; 7.0)</td>
<td>45 (40.1%; 6.9)</td>
<td>0.49</td>
</tr>
<tr>
<td>Germany</td>
<td>17 (46.4%; 12.1)</td>
<td>28 (68.9%; 8.8)</td>
<td>12</td>
<td>43 (47.4%; 6.2)</td>
<td>0.018</td>
</tr>
<tr>
<td>Israel</td>
<td>42 (47.4%; 5.6)</td>
<td>30 (53.2%; 7.0)</td>
<td>10</td>
<td>133 (50.7%; 3.2)</td>
<td>0.85</td>
</tr>
<tr>
<td>Italy</td>
<td>10</td>
<td>11 (31.7%; 10.1)</td>
<td>7</td>
<td>27 (65.8%; 7.4)</td>
<td>0.029</td>
</tr>
<tr>
<td>Japan</td>
<td>7</td>
<td>17</td>
<td>5 (34.2%; 6.0)</td>
<td>18 (34.2%; 6.0)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Netherlands</td>
<td>35 (64.9%; 7.1)</td>
<td>22 (45.2%; 15.5)</td>
<td>6</td>
<td>42 (47.5%; 9.2)</td>
<td>0.55</td>
</tr>
<tr>
<td>New Zealand</td>
<td>232 (57.4%; 2.9)</td>
<td>140 (34.7%; 3.4)</td>
<td>49 (26.3%; 4.3)</td>
<td>164 (32.0%; 2.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Spain</td>
<td>52 (65.4%; 7.3)</td>
<td>55 (61.3%; 5.5)</td>
<td>21 (45.2%; 10.4)</td>
<td>77 (47.5%; 6.5)</td>
<td>0.14</td>
</tr>
<tr>
<td>USA</td>
<td>250 (66.0%; 2.4)</td>
<td>182 (45.0%; 3.3)</td>
<td>91 (41.5%; 3.1)</td>
<td>215 (43.8%; 2.6)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Data are number (%; SE). The reported numbers are actual numbers rather than weighted estimates, which is why the ratios of these numbers to the total number of respondents in the survey do not equal the percentages. See methods section for a description of the weighting. ¶Test of difference in probability of treatment by severity. †One df χ² tests were done for Nigeria, Lebanon, Japan and China, where combined severe and moderate was compared against combined mild and none category. Three degree of freedom tests were done for all other countries. ‡Percentages were done for Nigeria, Lebanon, Japan and China, where combined severe and moderate was compared against combined mild and none category. Three degree of freedom tests were done for all other countries.
especially for respondents with distressing disorders.19,20 Service use might overestimate administrative records, groups. Earlier studies suggest that self-reports of sectors or clinical, sociodemographic, and cultural measuring a respondent’s commitment to the survey) would think carefully and answer honestly. Nevertheless, and excluding respondents who failed to say that they could have led to underestimation of unmet need WMH surveys did attempt to keep such inaccuracies to a minimum by using commitment probes (ie, questions on the magnitude and seriousness of concerns over differential diagnostic validity.

Fourth, without corroborating data for service use we cannot study the accuracy of self-reported treatment use or how this validity could differ across specific sectors or clinical, sociodemographic, and cultural groups. Earlier studies suggest that self-reports of service use might overestimate administrative records, especially for respondents with distressing disorders.21 In many countries, nearly a half of those initiating treatments failed to receive any follow-up care. Consistent with previous studies, WMH surveys did attempt to keep such inaccuracies to a minimum by using commitment probes (ie, questions measuring a respondent’s commitment to the survey) and excluding respondents who failed to say that they would think carefully and answer honestly. Nevertheless, potentially biased recall of service use remains possible and could have led to underestimation of unmet need for treatment, especially for those with serious disorders. Finally, despite the unprecedented scope and size of the WHO WMH survey initiative, some analyses consisted of small numbers of respondents, which might have rendered our conclusions less certain.

With these limitations in mind, our results show disturbingly high levels of unmet need for mental health treatment worldwide, even for people with the most serious disorders. The situation seems to be worst in less-developed nations, with only a few people with serious disorders receiving any form of care in the previous year; however, even in developed countries, roughly half of those with severe disorders receive no services. Additionally, the study limitations we describe that would lead to underestimation of unmet needs for treatment, especially in less-developed countries, compound these findings.

For the small number of people receiving some services, it seems likely that few are treated effectively. Some received non-health care from complementary and alternative medicine and human services sectors, despite growing questions about the effectiveness and safety of such treatments.22 In many countries, nearly a quarter of those initiating treatments failed to receive any follow-up care. Consistent with previous studies, in Ukraine than was expected from administrative data.18

<table>
<thead>
<tr>
<th>Countries</th>
<th>Any severity (p&lt;0.0001)*</th>
<th>Severe (p=0.051)*</th>
<th>Moderate (p&lt;0.0001)*</th>
<th>Mild (p&lt;0.001)*</th>
<th>None (p&lt;0.0001)*</th>
<th>p value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>47 (76.3%; 7.8)</td>
<td>6</td>
<td>6</td>
<td>13 (74.6%; 8.4)</td>
<td>22 (74.6%; 8.4)</td>
<td>0.50</td>
</tr>
<tr>
<td>Low-middle income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China†</td>
<td>56 (77.6%; 6.1)</td>
<td>4</td>
<td>6</td>
<td>3 (80.8%; 6.9)</td>
<td>43 (80.8%; 6.9)</td>
<td>0.33</td>
</tr>
<tr>
<td>Colombia</td>
<td>138 (72.0%; 4.3)</td>
<td>49 (92.6%; 3.5)</td>
<td>31 (73.1%; 7.9)</td>
<td>20 (61.7%; 11.3)</td>
<td>58 (66.6%; 7.9)</td>
<td>0.006</td>
</tr>
<tr>
<td>South Africa</td>
<td>601 (89.1%; 1.7)</td>
<td>42 (93.9%; 3.9)</td>
<td>63 (95.7%; 3.0)</td>
<td>58 (87.4%; 3.7)</td>
<td>438 (88.0%; 2.2)</td>
<td>0.39</td>
</tr>
<tr>
<td>Ukraine</td>
<td>167 (79.3%; 3.8)</td>
<td>44 (92.3%; 3.6)</td>
<td>51 (82.3%; 4.5)</td>
<td>14</td>
<td>58 (71.8%; 7.0)</td>
<td>0.006</td>
</tr>
<tr>
<td>High-middle income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lebanon†</td>
<td>62 (78.9%; 6.9)</td>
<td>17 (84.1%; 4.4)</td>
<td>15 (84.1%; 4.4)</td>
<td>7 (75.7%; 10.2)</td>
<td>23 (75.7%; 10.2)</td>
<td>0.367</td>
</tr>
<tr>
<td>Mexico</td>
<td>180 (74.5%; 4.4)</td>
<td>40 (85.5%; 4.2)</td>
<td>41 (76.6%; 6.7)</td>
<td>25 (84.3%; 6.9)</td>
<td>74 (67.8%; 7.7)</td>
<td>0.11</td>
</tr>
<tr>
<td>High income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>165 (84.3%; 3.9)</td>
<td>42 (84.4%; 9.5)</td>
<td>27 (84.3%; 10.4)</td>
<td>14</td>
<td>82 (83.1%; 5.3)</td>
<td>0.30</td>
</tr>
<tr>
<td>France</td>
<td>235 (86.0%; 3.9)</td>
<td>49 (87.5%; 4.7)</td>
<td>66 (97.3%; 1.6)</td>
<td>35 (89.7%; 4.4)</td>
<td>85 (80.0%; 6.9)</td>
<td>0.049</td>
</tr>
<tr>
<td>Germany</td>
<td>150 (70.2%; 5.1)</td>
<td>28 (89.2%; 8.5)</td>
<td>38 (97.1%; 0.7)</td>
<td>23</td>
<td>63 (61.1%; 7.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Israel</td>
<td>364 (86.1%; 1.8)</td>
<td>77 (88.2%; 4.1)</td>
<td>49 (89.2%; 4.1)</td>
<td>17</td>
<td>221 (83.6%; 2.3)</td>
<td>0.66</td>
</tr>
<tr>
<td>Italy</td>
<td>129 (94.5%; 3.5)</td>
<td>28</td>
<td>35 (93.5%; 3.4)</td>
<td>19</td>
<td>47 (94.4%; 2.5)</td>
<td>0.73</td>
</tr>
<tr>
<td>Japan†</td>
<td>83 (89.8%; 2.6)</td>
<td>9</td>
<td>13</td>
<td>9 (91.2%; 3.3)</td>
<td>52 (91.2%; 3.3)</td>
<td>0.33</td>
</tr>
<tr>
<td>Netherlands</td>
<td>183 (85.9%; 4.3)</td>
<td>55 (96.6%; 2.0)</td>
<td>35 (98.9%; 1.2)</td>
<td>15</td>
<td>78 (78.1%; 7.3)</td>
<td>0.006</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1394 (85.7%; 1.3)</td>
<td>421 (92.5%; 1.4)</td>
<td>368 (88.7%; 1.8)</td>
<td>151 (83.5%; 3.2)</td>
<td>454 (81.0%; 2.8)</td>
<td>0.002</td>
</tr>
<tr>
<td>Spain</td>
<td>341 (88.8%; 2.6)</td>
<td>73 (95.5%; 1.9)</td>
<td>86 (92.6%; 3.0)</td>
<td>35 (91.5%; 5.8)</td>
<td>147 (84.7%; 4.8)</td>
<td>0.12</td>
</tr>
<tr>
<td>USA</td>
<td>1313 (86.8%; 1.4)</td>
<td>362 (93.2%; 1.7)</td>
<td>354 (88.4%; 2.0)</td>
<td>187 (83.0%; 2.9)</td>
<td>410 (83.3%; 2.6)</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

Data are numbers (%; SE). The reported numbers are actual numbers rather than weighted estimates, which is why the ratios of these numbers to the total number of respondents in the survey do not equal the percentages. See methods section for a description of the weighting. *p value† is from a model predicting follow-up treatment among respondents in each level of severity that used any 12-month services †Test of difference in probability of treatment by severity. ‡One df χ² tests were done for Nigeria, Lebanon, Japan, and China, where combined severe and moderate was compared against combined mild and none categories. Three df tests were done for all other countries.

Table 5: Follow-up treatment for respondents using services in the WMH surveys§

Articles
few treatments were observed to meet minimum standards for adequacy.13–15,22

High levels of unmet need worldwide are not surprising, since WHO Project ATLAS’ findings of much lower mental health expenditures than was suggested by the magnitude of burdens from mental illnesses.16 Generally, unmet needs in low-income and middle-income countries might be attributable to these nations spending reduced amounts (usually <1%) of already diminished health budgets on mental health care, and they rely heavily on out-of-pocket spending by citizens who are ill equipped for it.23 Notable exceptions to the rule of greater unmet needs in developing countries might be attributable to these nations spending reduced amounts on health care. For example, South Africa’s high rates of treatment could indicate its greater levels of investment in health care. For example, South Africa, or New Zealand. ¶Minimally adequate treatment was defined as receiving eight or more visits to any service sector, or four or more visits and at least 1 month of medication, or being in continuing treatment at interview. Percentages are based on entire part II samples are those receiving minimally adequate treatment among those in treatment among respondents in each level of severity that used any 12-month services. †Test of difference in probability of treatment by severity. ‡The test was not done for Nigeria because there was only one (unweighted) case with adequate treatment. One degree of freedom χ² tests were done for Lebanon, Japan, and China, where combined severe and moderate was compared against combined mild and none category. Two degree of freedom test was done for the USA, where the mild and none categories were collapsed. Three degree of freedom tests were done for all other countries. §The questions on pharmacoepidemiology were not asked in Ukraine, South Africa, or New Zealand. ¶Minimally adequate treatment was defined as receiving eight or more visits to any service sector, or four or more visits and at least 1 month of medication, or being in continuing treatment at interview. Percentages are based on entire part II samples are those receiving minimally adequate treatment among those in treatment among respondents in each level of severity that used any 12-month services. †Test of difference in probability of treatment by severity. ‡The test was not done for Nigeria because there was only one (unweighted) case with adequate treatment. One degree of freedom χ² tests were done for Lebanon, Japan, and China, where combined severe and moderate was compared against combined mild and none category. Two degree of freedom test was done for the USA, where the mild and none categories were collapsed. Three degree of freedom tests were done for all other countries. §The questions on pharmacoepidemiology were not asked in Ukraine, South Africa, or New Zealand.

<table>
<thead>
<tr>
<th>Region</th>
<th>Low income</th>
<th>Moderate (p=0.0056)*</th>
<th>Any severity (p&lt;0.0001)*</th>
<th>None (p&lt;0.0001)*</th>
<th>Severe (p&lt;0.0001)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-middle income</td>
<td>China</td>
<td>19 (24.1%; 7.1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Colombia</td>
<td>33 (14.7%; 3.4)</td>
<td>11 (23.1%; 8.5)</td>
<td>7 (21.7%; 10.5)</td>
<td>3 (6.3%; 4.6)</td>
</tr>
<tr>
<td></td>
<td>South Africa§</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>High-middle income</td>
<td>Lebanon¶</td>
<td>18 (24.5%; 7.1)</td>
<td>5 (24.0%; 6.2)</td>
<td>3 (24.0%; 6.2)</td>
<td>3 (24.8%; 10.7)</td>
</tr>
<tr>
<td></td>
<td>Mexico</td>
<td>42 (15.2%; 2.7)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>High income</td>
<td>Belgium</td>
<td>78 (33.6%; 5.2)</td>
<td>23 (42.5%; 8.5)</td>
<td>12 (35.5%; 12.6)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>113 (42.3%; 5.4)</td>
<td>29 (57.9%; 8.5)</td>
<td>29 (36.5%; 6.6)</td>
<td>15 (41.5%; 9.7)</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
<td>91 (42.0%; 6.1)</td>
<td>21 (67.3%; 10.7)</td>
<td>22 (53.9%; 8.5)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Israel</td>
<td>118 (35.1%; 2.5)</td>
<td>31 (35.2%; 5.3)</td>
<td>23 (42.8%; 6.8)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>65 (33.0%; 5.1)</td>
<td>12</td>
<td>11 (33.4%; 9.1)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td>35 (31.8%; 6.8)</td>
<td>6</td>
<td>5 (27.9%; 7.0)</td>
<td>18 (27.9%; 7.0)</td>
</tr>
<tr>
<td></td>
<td>Netherlands</td>
<td>58 (34.4%; 5.0)</td>
<td>39 (67.2%; 9.0)</td>
<td>19 (31.4%; 10.2)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>New Zealand§</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>152 (37.3%; 3.3)</td>
<td>41 (47.5%; 7.5)</td>
<td>37 (43.5%; 6.6)</td>
<td>22 (48.5%; 9.8)</td>
</tr>
<tr>
<td></td>
<td>USA‡</td>
<td>302 (18.1%; 1.1)</td>
<td>160 (41.8%; 3.2)</td>
<td>101 (24.8%; 2.1)</td>
<td>41 (4.9%; 0.8)</td>
</tr>
</tbody>
</table>

Data are number (%; SE). The reported numbers are actual numbers rather than weighted estimates, which is why the ratios of these numbers to the total number of respondents in the survey do not equal the percentages. See methods section for a description of the weighting. ¶’Y’ is from a model predicting minimally adequate treatment among respondents in each level of severity that used any 12-month services. †Test of difference in probability of treatment by severity. ¶The test was not done for Nigeria because there was only one (unweighted) case with adequate treatment. One degree of freedom χ² tests were done for Lebanon, Japan, and China, where combined severe and moderate was compared against combined mild and none category. Two degree of freedom test was done for the USA, where the mild and none categories were collapsed. Three degree of freedom tests were done for all other countries. §The questions on pharmacoepidemiology were not asked in Ukraine, South Africa, or New Zealand. ¶Minimally adequate treatment was defined as receiving eight or more visits to any service sector, or four or more visits and at least 1 month of medication, or being in continuing treatment at interview. Percentages are based on entire part II samples are those receiving minimally adequate treatment among those in treatment among respondents in each level of severity that used any 12-month services. †Test of difference in probability of treatment by severity. ‡The test was not done for Nigeria because there was only one (unweighted) case with adequate treatment. One degree of freedom χ² tests were done for Lebanon, Japan, and China, where combined severe and moderate was compared against combined mild and none category. Two degree of freedom test was done for the USA, where the mild and none categories were collapsed. Three degree of freedom tests were done for all other countries. §The questions on pharmacoepidemiology were not asked in Ukraine, South Africa, or New Zealand.

Table 6: Minimally adequate treatment use for respondents using services in the WMH surveys¶

We need to understand how the few mental health resources that nations do have can be best allocated. An overly simplistic view of our results could be that a meaningful number of services are going to those without apparent needs. Such potential diversion of limited treatment resources to individuals without apparent needs would be of concern in view of the magnitude of unmet needs for patients with clearly defined and serious disorders. The weak or absent relation between use of mental health services will also be important.26
on the few patients with access to specialty sectors.25 This finding could also suggest gatekeeping by primary care physicians employed in some countries to reserve specialty treatment for severe cases.26 Whatever the rationale, we need to ensure that mental health care received in general medical sectors is not of low intensity and adequacy, as has been recorded in other studies.22

Our results for predictors of service use are generally consistent with previous work. The young relative to middle-aged carers might be more dependent on others and therefore reluctant to access services;26 on the other hand, elderly people might avoid seeking mental health care because of the greater perceived stigma of mental disorders and treatments for people in this age range than for those who are younger.16 Higher rates of treatment for women than for men could be explained by women’s diminished perceptions of stigma and their greater abilities to translate non-specific feelings of distress into conscious recognition of having a mental health problem.31

Effects of greater income were variable, since service use increased in some countries but decreased in others. Substantial effects of financial barriers on seeking treatment could exist in countries where there is a positive association between income and service use.32 However, negative associations could be explained by the fact that only poor people qualify for entitlements in some countries.42 Respondents who are well educated might also have greater resources than those whose education was poor; alternatively, their higher treatment rates might show that some methods (eg, psychotherapies) place an emphasis on knowledge and cognitive processes. The generally increased use of mental health services in those not married could indicate the power of relationship loss, strife, or social impairments as motivators for seeking treatment.30

Our results have implications in several areas. First, alleviation of the difficulty of widespread undertreatment will almost certainly need expansion of treatment resources and governmental as well as private means of financing mental health services. Second, there is also a pressing need to devise rational, transparent, and ethical allocation rules. Should countries focus resources on those with the greatest needs rather than on increasing numbers with mild disorders (to prevent negative sequelae)? Should service be delivered through primary rather than specialty sectors, or inpatient instead of community settings? And should countries provide mental health services on parity with those for general medical disorders?33 Ideal these questions would be answered through formal analyses of the burdens from illnesses and the cost-effectiveness of treatments.16 Unfortunately, rigorous data to analyse disease burdens and weigh the costs and benefits of different regimens are largely scarce.27 Without such rational schemes, decisions about resource allocation are usually made on the basis of simple minimisation of costs and even attitudinal factors such as stigma and desire to punish people perceived as being personally responsible for their mental health problems.35

Finally, when rational, transparent, and ethical priorities have been set, policymakers need specific designs that they can implement to achieve their goals. Some techniques used in managed care systems (eg, gatekeeping, increased cost sharing, review of use, previous approval, etc) could presumably be brought to bear on unnecessary use but not underuse—in fact, some techniques could worsen unmet needs for treatment. Furthermore, these elements from largely developed nations such as the USA might not be translatable to other countries and circumstances. The effects of other policies, delivery system features, and means of financing that policymakers could implement, are essentially unknown. Therefore, gathering of detailed data for the mental health policies, delivery system features, and means of financing mental health care in different countries is a promising area for future work.21 When merged with WMH surveys on the use and adequacy of treatments, such combined data could shed light on the effects of policies, delivery system, and financing features, and help policymakers choose policies that achieve their desired goals.16

Contributors
All authors participated in the design of the manuscript. PSW, RCK, and MCL made substantial contributions to the analysis. All authors made contributions to the interpretation of the data. All the authors contributed to part of the content, took part in critical revision of the manuscript, and contributed to the acquisition of the data. All the authors had access to data from their own country, but only PSW, MCL, and RCK had access to the consolidated cross-national dataset. PSW and RCK had full access to all the data in the study and had final responsibility for the decision to submit for publication.

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

Acknowledgments
The surveys discussed in this article were done in conjunction with the WHO World Mental Health (WMH) Survey Initiative. We thank WMH staff for assistance with instrumentation, fieldwork, and data analysis. These activities were supported by the US National Institute of Mental Health (R01MH070884), the John D and Catherine T MacArthur Foundation, the Pfizer Foundation, the US Public Health Service (R15-MH066849, R01-MH069864, and R01 DA016558), the Fogarty International Center (FIRCA R01-TW09648), the Pan American Health Organization, Eli Lilly and Company, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline, and Bristol-Myers Squibb. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmh/. The Chinese World Mental Health Survey Initiative is supported by the Pfizer Foundation. The Colombian National Study of Mental Health (NSMH) is supported by the Ministry of Social Protection, with supplemental support from the Saldarriaga Concha Foundation. The ESEMeD project is funded by the European Commission (Contracts QLG5-1999-01042; SANCO 2004123), the Piedmont Region (Italy), Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Spain (FIS 00/0028), Ministerio de Ciencia y Tecnología, Spain (S AF 2000-158-CE), Departamento de Salud, Generalitat de Catalunya, Spain, Instituto de Salud Carlos III (CIBER CB06/02/0046, RETICS RD06/0011 REM-TAP), and other local agencies and by an unrestricted educational grant from GlaxoSmithKline. The Israel National Health Survey is funded by the Ministry of Health with support from the Israel National Institute for Health Policy and Health Services Research and the National Insurance Institute of Israel. The World Mental Health Japan (WMHJ) Survey is supported by the Grant for
Research on Psychiatric and Neurological Diseases and Mental Health (H13-SHOGAI-023, H14-TOKU/BETSU-026, H16-KOKORO-013) from the Japan Ministry of Health, Labour and Welfare. The Lebanese National Mental Health Survey (LEBANON) is supported by the Lebanese Ministry of Public Health, the WHO (LEBANON), anonymous private donations to IDRAAC, Lebanon, and unrestricted grants from Janssen Cilag, Eli Lilly, GlaxoSmithKline, Roche, and Novartis. The Mexican National Comorbidity Survey (MNCS) is supported by The National Institute of Psychiatry Ramon de la Fuente (INPREFMIDES 4280) and by the National Council on Science and Technology (CONACYT-630544-H), with supplemental support from the Pan American Health Organization (PAHO). Te Rau Hinengaro: The New Zealand Mental Health Survey (NZMHS) is supported by the New Zealand Ministry of Health, Alcoholic Advisory Council, and the Health Research Council. The Nigerian Survey of Mental Health and Wellbeing (NSMHW) is supported by the WHO (Geneva), the WHO (Nigeria), and the Federal Ministry of Health, Abuja, Nigeria. The South Africa Stress and Health Study (SASH) is supported by the US National Institute of Mental Health (R01-MH69575) and National Institute of Drug Abuse with supplemental funding from the South African Department of Health and the University of Michigan. The Ukraine Comorbid Mental Disorders during Periods of Social Disruption (CMDPSD) study is funded by the US National Institute of Mental Health (R01-MH69990). The US National Comorbidity Survey Replication (NCS-R) is supported by the National Institute of Mental Health (NIMH; U01-MH61220) with supplemental support from the National Institute of Drug Abuse (NIDA), the Substance Abuse and Mental Health Services Administration (SAMHSA), the Robert Wood Johnson Foundation (RWJF; Grant 094789), and the John W Alden Trust.

References

32. Wells KB, Manning WG, Duan N, Newhouse JP, Ware JE Jr. Sociodemographic factors and the use of outpatient mental health services. Med Care 1986; 24: 75–85.
Depression, chronic diseases, and decrements in health: results from the World Health Surveys

Saba Moussavi, Somnath Chatterji, Emese Verdes, Ajay Tandon, Vikram Patel, Bedirhan Ustun

Summary

Background Depression is an important public-health problem, and one of the leading causes of disease burden worldwide. Depression is often comorbid with other chronic diseases and can worsen their associated health outcomes. Few studies have explored the effect of depression, alone or as a comorbidity, on overall health status.

Methods The WHO World Health Survey (WHS) studied adults aged 18 years and older to obtain data for health, health-related outcomes, and their determinants. Prevalence of depression in respondents based on ICD-10 criteria was estimated. Prevalence values for four chronic physical diseases—angina, arthritis, asthma, and diabetes—were also estimated using algorithms derived via a Diagnostic Item Probability Study. Mean health scores were constructed using factor analysis and compared across different disease states and demographic variables. The relation of these disease states to mean health scores was determined through regression modelling.

Findings Observations were available for 245 404 participants from 60 countries in all regions of the world. Overall, 1-year prevalence for ICD-10 depressive episode alone was 3.2% (95% CI 3.0–3.5); for angina 4.5% (4.3–4.8); for arthritis 4.1% (3.8–4.3); for asthma 3.3% (2.9–3.6); and for diabetes 2.0% (1.8–2.2). An average of between 9.3% and 23.0% of participants with one or more chronic physical disease had comorbid depression. This result was significantly higher than the likelihood of having depression in the absence of a chronic physical disease (p<0.0001). After adjustment for socioeconomic factors and health conditions, depression had the largest effect on worsening mean health scores compared with the other chronic conditions. Consistently across countries and different demographic characteristics, respondents with depression comorbid with one or more chronic diseases had the worst health scores of all the disease states.

Interpretation Depression produces the greatest decrement in health compared with the chronic diseases angina, arthritis, asthma, and diabetes. The comorbid state of depression incrementally worsens health compared with depression alone, with any of the chronic diseases alone, and with any combination of chronic diseases without depression. These results indicate the urgency of addressing depression as a public-health priority to reduce disease burden and disability, and to improve the overall health of populations.

Introduction Depression is an important global public-health issue, both because of the relatively high lifetime prevalence ranging from 2% to 15% and because it is associated with substantial disability.\(^1\) Rated as the fourth leading cause of disease burden in 2000, depression accounted for 4.4% of total disability-adjusted life years (DALYs).\(^1\) It is also responsible for the greatest proportion of disease burden attributable to non-fatal health outcomes, accounting for almost 12% of total years lived with disability worldwide.\(^1\) Without treatment, depression has the tendency to assume a chronic course, be recurrent, and over time to be associated with increasing disability.\(^2\)

The comorbidity of depression with chronic physical diseases such as arthritis and diabetes is well recognised in developed countries.\(^3\)\(^,\)\(^4\) Several studies have shown that there is an increased risk of having major depression in people with one or more chronic diseases.\(^5\)\(^,\)\(^6\)\(^,\)\(^7\) The degree to which these comorbid states exist at the global level has not been shown. With a growing elderly population, and the associated increase in prevalence of chronic medical conditions, a concomitant rise in the prevalence of depression is to be expected. In fact, projections indicate that after heart disease, depression is expected to become the second leading cause of disease burden by the year 2020.\(^8\)

The increasing prevalence of chronic physical diseases and depression leads to the question of how these disorders compare in terms of their effect on overall individual health. The presence of self-reported chronic physical diseases such as angina, arthritis, asthma, and diabetes has been associated with reduced health-related quality of life scores.\(^9\)\(^,\)\(^10\) Lower health status has been reported in depressed patients than in those without depression, and this state is unequally distributed across population groups.\(^11\)\(^,\)\(^12\)\(^,\)\(^13\) Effects of depressive episodes have also been studied with regard to loss in productivity and poor health-related quality of life.\(^14\)\(^,\)\(^15\)\(^,\)\(^16\)\(^,\)\(^17\) Despite this evidence, depression, like other mental disorders, is often not deemed to be on a par with other chronic physical health conditions in terms of its effect on overall health.\(^17\)\(^,\)\(^18\) This view is perhaps one of the underlying reasons behind the lack of parity between mental and physical disorders in terms of access to health care.\(^19\)\(^,\)\(^20\)\(^,\)\(^21\) To our knowledge, there has been no worldwide...
comparison of depression with other chronic diseases and their effect, either individually or comorbid, on health.

We analysed data from the WHO World Health Survey (WHS) to address the following questions: how does the decrement in health state associated with depression compare with the decrement associated with other common chronic physical conditions; and what is the added effect on decrements in health of suffering from depression, over and above a chronic physical condition?34

Methods
Sample
Countries from the WHS were selected to represent all regions of the world, with 26 countries from the European region, 15 from the African region, six from the Americas, four from the eastern Mediterranean region, five from the southeast Asia region, and four from the western Pacific region, giving a total of 60 countries. The countries included in the survey programme represent those countries that were willing and able to participate in the survey. Countries with samples that were nationally representative, probabilistically selected, and which had sampling weights information available, were used in the analysis for this paper. To adjust for the population distribution as represented by the UN Statistical Division and for non-response, post-stratification corrections were made to the sampling weights.

Procedures
All respondents used in the analysis were interviewed with the standardised WHS survey, which included questions on sociodemographic and economic factors, a series of questions on health status, and questions related to whether the individual had ever been diagnosed with depression, asthma, arthritis, angina, and diabetes, whether the person had ever received or was currently on treatment for these conditions, and, with the exception of diabetes, a series of symptom questions related to each condition. All surveys were implemented as face-to-face interviews with the exception of Luxembourg and Israel, which were implemented as telephone interviews. All questionnaires were translated and back-translated using a standard WHO protocol. The quality of translations was independently verified by bilingual experts before field implementation. Informed consent was obtained from all respondents and the study was cleared by the ethics review committees at each site.

The health state measure presented in this analysis was developed by WHO based on its framework for measuring health. WHO assesses an individual’s state of health as a vector of capacities in multiple domains.34 For measurement in surveys, this information needs to be reduced to a parsimonious set of domains that are clearly defined and measured with reliable self-report questions. After an extensive review of existing survey questionnaires, none were deemed able to match the exact ideas or have the information needed to measure the distribution of health in the general population. Hence, a new measure was developed where the valuation could be estimated—ie, the relative disability weight assigned to different patterns of the health states, thus allowing for cross-country comparability.

The measure was based on 18 health-related questions, where the responses were recorded on a five point scale ranging from “no difficulty or problem” to “extreme difficulty/inability”. Two of the questions assessed general health: one asking overall self-reported health, and the other asking how much difficulty the respondent had in working or doing household activities during the past 30 days. These two items were analysed individually.

The remaining 16 questions were grouped into the following eight health domains: vision, mobility, self care, cognition, interpersonal activities, pain and discomfort, sleep and energy, and affect. These domains are included in many commonly used health outcome measures such as the Short Form 36 (SF36), the Health Utilities Index Mark 3 (HUI 3), and the Euroqol 5D.35–37

The health measure had been extensively tested as part of a similar survey done between 2000 and 2001, the Multi-Country Survey Study.38 The internal consistency of the health measure as assessed using Cronbach’s alpha was 0.91. The test–retest reliability of individual items, measured by the weighted Kappa, ranged from 0.48–0.62. Missing data for individual items ranged from 1.3% to 5.8%. Construct validity was also assessed, and respondents who were older or had a chronic condition reported worse health, whereas respondents with higher socioeconomic status and countries with higher life expectancies on average reported better health.39

A composite health status score was derived from the 16 self-reported health questions. Since the item responses were based on a five-point ordered categorical scale, a factor analysis using polychoric correlations was done to take into account the covariance structure of the responses to individual questions. The choice of one factor solution was justified by the high eigenvalue of the first factor (8.79, 74% as a cumulative percentage of the variance explained) and the high communalities of the original variables (between 0.43 and 0.69). We used the principal component method for factor extraction and the regression scoring method to obtain the factor scores. The factor score was transformed to a 0–100 scale, with 0 indicating worst health and 100 indicating best health.

To validate the use of symptom questions for diagnosing chronic diseases, WHO implemented in 2003 a diagnostic item probability study in seven countries. Patients were selected from clinics if they were positive for any of the specific disease conditions based on a gold standard diagnostic test, and these patients were considered to be true positives for that particular condition. The patients were then traced back to their homes and asked, for all
disease conditions, the same symptomatic questions as respondents from the WHS. Additionally, a sample of respondents matched by sex, age, and country of origin was drawn from the WHS if they had negative responses to all the self-reported diagnosis questions for depression, asthma, arthritis, angina, and diabetes. These respondents were considered to be the true negatives for the study. The individual response rates, calculated as the ratio of completed interviews in selected respondents in the sample, excluding ineligible respondents from the denominator, ranged from 63% in Israel to 99% in the Philippines (detailed response rates available on request).

The diagnosis of depression was based on the International Classification of Diseases tenth revision (ICD-10) diagnostic criteria for research for depressive episodes, and was derived from an algorithm that took into account respondents reporting symptoms of depression during the past 12 months. The individual questions used to assess these symptoms were based on the World Mental Health Survey version of the Composite Diagnostic Interview. The diagnosis for angina was based on the algorithm derived from the Rose questionnaire. For asthma and arthritis, the sensitivity and specificity of all potential combinations of answers to these symptomatic questions were checked based on responses from the diagnostic item probability study. The combination of answers that produced the best result based on the Receiver Operator Characteristic analysis was used to apply a diagnosis for each respondent in all 60 countries of the WHS sample. Respondents were regarded as positive for diabetes if they reported ever being diagnosed with diabetes. Questions about diabetes were asked in only 46 of the countries that implemented the long version of the questionnaire. All diagnoses of these chronic physical diseases applied to the past 12 months from the date of interview. For most of the analysis, respondents were grouped on the basis of their disease status into one of the following: respondents having none of the aforementioned health conditions, respondents having any of the single conditions alone, respondents having any of the chronic physical diseases alone in conjunction with depression, respondents having two or more comorbidities without depression, and respondents having two or more comorbidities with depression.

Analysis
The prevalence of each chronic physical disease was estimated, first alone—ie, without any of the other conditions present—then comorbid with depression but without any additional conditions present, and then two or more comorbid conditions with or without depression. The prevalence of depression in respondents who had any one of the conditions was also estimated. All these estimates were calculated using post-stratified probability weights. To make valid comparisons across countries, age and sex standardisations were done using WHO’s World Standard Population for age and the UN Statistical Division for sex ratio.

The mean of the health score was calculated using probability weights for the entire sample after stratification by sex, age, education, and income quintile, as well as respondents’ disease status. To test the statistical difference of health scores between each pair of disease groups, a one-way analysis of variance using a Scheffe test was done to adjust for multiple comparisons. Linear regression analysis was used on the pooled dataset of 46 countries to model the relation between respondents’ health state and whether they had depression, a chronic physical disease, or a combination thereof, after controlling for country of origin, sex, age, education, marital status, occupational status, income level, and any interaction between sex with the other demographic variables. To establish whether cultural differences in countries affected the relation between disease state and overall health state, interaction terms between marital status, education, and income quintile with the country variable were included in the model.

Responses to some health domains such as sleep and affect are likely to be influenced by whether the respondent is clinically depressed, which could lead to spurious conclusions about the decrements in health associated with depression. To test for this possibility, we did two further analyses. The first analysis explored the association of disease states with each of the two questions of health measure which assess overall health and do not include any symptoms of depression. Second, we used a recursive regression technique to model the effect of depression when the dependent variable, the health score, contained only two domains—mobility and vision. Then, progressively, other domains were added to estimate the health score. The effect of depression on each successive health score was assessed at each marginal addition of a domain by evaluating changes in the regression coefficient. This regression analysis was repeated until all eight health domains were included and their coefficients for depression compared. Minimum differences in the coefficient for each successive regression would corroborate the effects of depression on decrements in health as genuine and not the result of a systematic reporting bias. All analysis was coded and done using STATA version 9.2.

Role of the funding source
This study was funded by WHO. 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.
Results
The prevalence for depression alone, each chronic disease alone, depression with each chronic physical disease, and having multiple chronic conditions with or without depression for 60 countries was estimated with CIs (results not shown, available on request). At the worldwide level, the prevalence for having any one condition alone did not exceed 5·0%. The prevalence of having diabetes alone had the lowest overall prevalence of 2·0% (1.8–2.2). However, since diabetes prevalence was based on self-report, the role of reporting bias that possibly underestimates the true prevalence cannot be ruled out. Depression alone had the next lowest overall prevalence at 3·2% (3·0–3·5). Asthma alone had an overall prevalence of 3·3% (2.9–3.6), prevalence of arthritis alone was 4·1% (3.8–4.3), and angina alone 4·5% (4·3–4·8). There were variations across countries, but the range of differences in prevalence of any one condition did not exceed eight percentage points.

A significant percentage of respondents with any one of the chronic physical conditions also had depression. For respondents with diabetes, at a worldwide level, 9·3% (7.3–11.3) also had depression, 10·7% (9·1–12·3) with arthritis also had depression, 15·0% (12·9–17·2) with angina, and respondents with asthma had the highest prevalence of depression at 18·1% (15·9–20·3). For the 7·1% (6·6–7·6) of respondents who had comorbidity of two or more chronic physical conditions, nearly a quarter (23%) also had depression in addition to their existing comorbid conditions. Thus, the prevalence of depression in respondents with chronic diseases is significantly higher than in respondents without chronic diseases (3·2%, p<0·0001).

The figure shows the mean health score and the 95% CIs for each disease. Respondents without any of the chronic diseases or depression had the highest health score, 90·6—i.e., reported having the best health. Respondents with asthma, angina, arthritis, or diabetes alone, had mean health scores of 80·3, 79·6, 79·3, and 78·9, respectively, which were significantly different from having no disease but not from each other. Respondents with depression had the lowest health score among all the chronic disease conditions, 72·9 (p<0·0001). Respondents who had depression comorbid with another chronic condition had much lower mean health scores than respondents who had the chronic condition alone (p<0·01). For respondents who had two or more chronic conditions excluding depression, their mean health score was 71·8, lower than any of the disease conditions alone but higher than any disease state comorbid with depression. The lowest overall mean health score was for respondents with two or more chronic conditions comorbid with depression (56·1). These results show that comorbid depression is significantly associated with lower health states in respondents with chronic conditions in comparison to having chronic conditions, including multiple chronic conditions, without depression (p<0·0001).

We examined mean health scores by disease state in more detail by looking across sociodemographic variables (results not shown). The patterns are consistent, and for depression alone the mean health score is lower than for other chronic conditions alone, across all socioeconomic variables. For all comorbid depression, the mean health score is lower across all socioeconomic variables than for any of the chronic conditions alone or for depression alone. Thus, having depression comorbid with another chronic physical disease lowers health status substantially, irrespective of a respondent’s age, sex, and other demographic variables.

The coefficients of a regression model in the table summarises the relation between overall health and the different disease states and the sociodemographic determinants of sex, age, education, employment status, income quintile, and marital status. We also controlled for country of residence, interaction of country with education, martial status and income quintile, and interaction of sex with education, employment status, and income quintile (results not shown).

The results from the model indicate that lower coefficient values are associated with lower health scores. Results of the model show that being older is indicative of decreased health status, as is having less education, having lower income, and being unemployed. Women had a lower overall health score, and the decrements in health were greater for women who were unemployed, less educated, or widowed as indicated by the significant coefficients for the interaction terms of sex with these demographic variables (results not shown). There does not seem to be a significant difference in health status between being married and
never having been married or cohabitating. However, being separated, divorced, or widowed is associated with lower health scores.

Overall, disease status has a greater association with reported health scores than do sociodemographic characteristics. Controlling for all other factors, having depression is associated with the lowest health scores, either alone or comorbid with other chronic diseases (p<0·0001). The coefficient values from the model show that the comorbid state of depression with diabetes causes even greater decrements in health than the addition of the two conditions separately. This finding is suggestive of an interactive effect between depression and diabetes that causes an extra negative effect on health beyond the simple addition of each of the two conditions. Having more than two chronic diseases without depression (coefficient –11.97), although associated with a lower health score than having any one chronic disease, has much less of a negative association with health than does having depression alone (coefficient –13.89), or depression comorbid with one of the chronic diseases (coefficients range from –16.77 to –23.43). Respondents with two or more chronic diseases in addition to having depression had the lowest health scores of all the disease groups (coefficient –24.38).

Interacting sociodemographic variables with the country variable did not show statistically significant change the coefficients of depression and comorbid depression in the model, which suggests that cultural differences across countries and their interaction with sociodemographic characteristics does not affect the influence of depression on overall health.

To rule out the effects of depression on some of the health domains included in the health measure, we compared the mean scores by disease status for the two general questions on overall self-rated health and difficulties with work and activities (results not shown). Although the scores for the difficulties with work question were higher than the overall health question, the pattern across disease states for both questions was quite similar to the pattern seen for the overall mean health score. Respondents who reported no chronic conditions had the highest scores, respondents with depression alone or comorbid with another condition had the lowest scores overall. Even respondents with two or more chronic conditions but no depression scored higher than any respondent with depression alone or comorbid with another condition.

We also did a recursive regression as described in the methods section (results not shown). The coefficient for depression in the model with the least-related of the health domains, vision and mobility, had a value of –9.9. Adding a third domain, pain, the coefficient for depression rose slightly to –13.6. With each successive addition of a health domain, the coefficient does not exceed –13.9. The addition of more domains—even those that are likely to be most responsive to the presence of depression such as sleep or energy and affect—did not have an appreciable addition on the average effect of depression on health status. The coefficient remains fairly stable irrespective of the composition of domains that underlie the computation of the health score. These analyses show that our findings are unlikely to indicate a bias due to inclusion of items in the overall health status score that are related to depression.
Discussion

The worldwide prevalence of depression, asthma, angina, arthritis, and diabetes based on data collected in the World Health Surveys, and used in the analysis presented here, are similar to the data reported by WHO’s Global Burden of Disease study. The data show that comorbidity between chronic physical conditions and depression is common, and that people with chronic diseases are significantly more likely to suffer from depression than those without (p<0.0001). Our data indicate that depression is associated with a decrement in health that is significantly greater than those associated with the other chronic diseases in this study. Though depression has previously been shown to be associated with disability and declines in health-related quality of life, this is the largest scale study to our knowledge that shows this decline using direct comparisons across physical conditions in multiple countries with a common measurement strategy. Furthermore, we have also shown that depression comorbid with other chronic diseases produced significantly greater decrements in health than from one or more chronic diseases, and that this additive effect is substantially amplified in the case of depression comorbid with diabetes. These associations remained evident after adjustment for sociodemographic, country of origin, and economic factors.

Our findings are consistent with earlier studies that have shown a high degree of association between depression and disability. There are, however, few studies that have compared the effect of depression with other chronic diseases. One reason for our findings could be that depression is associated specifically with decrements in mental domains of health, which were included in the composite health score we computed for our analyses. However, in the recursive regression analysis, we showed that adding each health domain serially does not alter the substantive results, since the size of the depression regression coefficient is barely changed. This finding confirms that the measure is not biased towards depression. Another reason for our findings might be that depression is associated with a negative assessment of functioning in all domains and therefore what one is measuring is merely a negative frame of mind that leads to reporting biases. An illustration of such a response bias is shown in the study by Owsley and colleagues, who assessed the effect of depression in elderly individuals on their response to a vision questionnaire. After controlling for demographics, general health, and vision, depression was found to be associated with reduced scores on the questionnaire, suggesting negative reporting as a function of being depressed rather than actual vision ability. To address this bias, the WHS also included vignettes in the survey whereby each respondent was presented with a set of brief descriptions of individuals in a fixed level of health for a particular domain. Five vignettes per domain were presented ranging, for example, from quadriplegia at one extreme of mobility, to a marathon runner at the other extreme. Respondents were asked how they would rate their difficulty in that particular domain if they were the person described in each of the vignettes. The examination of these rating patterns show that respondents with and without depression showed the same pattern of rating, even though they differed in self-report of their own experiences in each domain. This finding suggests that depressed respondents were not reporting things more negatively for the same level of health, and further supports the absence of biased reporting due to depression (results not shown). Our vignette method was possibly not sufficiently sensitive in detecting systematic reporting biases: though we do not believe this to be the case, this possibility needs to be investigated in future studies. Additionally, the reporting of depressive symptoms or diagnosis could vary between countries because of cultural differences in reporting such symptoms. If respondents in some countries underreported their depressive symptoms, leading to an underestimate of the prevalence of depression in these countries, and if this was associated with denial of health problems, it would in fact narrow the difference in the reported decrements in health between those with and without depression. Our results show this situation was not the case, that even after controlling for country effects, the decrements in health due to depression, both in pure and comorbid states, continue to remain significant. So if underestimating of depression prevalence is occurring, it actually strengthens the findings that depression increases decrements in health. Thus, these reporting biases, if they exist, do not detract from our substantive findings.

The WHS was a cross-sectional study and did not include questions on onset and duration of illness, fluctuations in course and details of health-care use such as number and timing of contacts with health-care services, the reasons for contact, and the outcome following contact. Hence, we cannot establish what burden depression, and its comorbidity with other chronic diseases, places on the health system, how depression can modify the course of these disorders, and whether treatment of depression when present with these chronic physical diseases would alter their course. For the estimation of prevalence, the algorithms for the chronic physical conditions were based on a small validation study in a few countries that used negatives drawn from respondents in the WHS who self-reported no chronic diseases. Since these conditions are known to have an average prevalence in the general population of around 5%, the likelihood of these true negative respondents having any one of the above diagnoses is low. Both depression and angina were based on validated algorithms, but the algorithms of asthma and arthritis could benefit from a more comprehensive validation study, since the presence of some false negative respondents cannot be ruled out. The algorithms might
need to be modified if larger validation studies are done in more countries and if the true negatives were also identified on a gold standard test. Diabetes was based on self-report, and the role of reporting bias on the prevalence presented in our study is noted. However, we do not think that the possibility of a small misclassification bias would substantially alter the core message of this study that being sad is bad for one’s health.

In conclusion, we report the largest population-based worldwide study to our knowledge that explores the effect of depression in comparison with four other chronic diseases on health state. Our main findings show that depression impairs health state to a substantially greater degree than the other diseases. A significant percentage of respondents have depression in addition to their existing chronic physical conditions, a group that is often unrecognised and untreated.\textsuperscript{20, 45} This finding is of special importance, considering the presence of depression and its treatment is clearly related to the outcome of these chronic diseases.\textsuperscript{7, 46–48} Comorbidity with depression significantly worsens the health state of people with chronic diseases. The need for timely diagnosis and treatment of depressive disorders to reduce the burden on public health is imperative. In many primary care settings, patients presenting with multiple disorders that include depression often don’t get diagnosed, and if they do, often treatment is focused towards the other chronic diseases.\textsuperscript{4} Depression can be treated in primary care or community settings with locally available cost-effective interventions.\textsuperscript{13} On the basis of our results, addressing the further exacerbation of disability due to depression needs to be a priority of health systems worldwide. Primary care providers must be taught not to ignore the presence of depression when patients present with a chronic physical condition, in view of the marked effect that it has on an individual’s health. This goal can be accomplished in part by sending a message which, in addition to reducing the stigma surrounding mental illness, can alert providers and the public at large that depression is a disease at least on a par with physical chronic diseases in damaging health.

Contributors
SM, SC, and TBU contributed to the design of the study. SM, SC, EV, AT, and VP contributed to the analyses. All authors were involved in the development of the manuscript and approved the final version. The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the Asian Development Bank or the World Health Organization.

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

Acknowledgments
AFRO—Blaise Sondo (Burkina Faso), Diona Avoodsoomua (Chad), Sitti Djasoura Chihabeddine (Comores), Georges Moyen (Congo, Brazzaville), M Seka Monnay Firmin (Cote D’Ivoire), Makonnen Asea (Ethiopia), Richard Britwum (Ghana), Fredrick Otieno (Kenya), Sidon Konyani (Malawi), Mamadou Basséry Ballo (Mali), Maye Mint Haïdy (Mauritania), P Barhou (Morocco), M Zerrouk (Namibia), Babacar Drame (Senegal), Zaid Kimmie (South Africa), Isabel Thembi Zwane (Swaziland), Musonda Lembila (Zambia), Julita Maradzika (Zimbabwe), AMRO/PAHO—Célia Landmann Szwarcwald (Brazil), Marita Molina Achcar (Dominican Republic), Fernando Sacoto (Ecuador), Tim Farrell (Guatemala), G Olaz Fernández (Mexico), M W Torres Numbay (Paraguay), Jorge Agulla (Uruguay), EMRO—Mustapha Azeizmat (Morocco), Ashfaq Ahmed (Pakistan), Nourredinne Achour (Tunisia), Gohar Wajid (United Arab Emirates), Christine Kotarakos (INRA-ECO EU surveys), Aida Pilav (Bosnia and Herzegovina), Marijan Ereg (Croatia), Vlasta Mazzanova (Czech Republic), Andras Saar (Estonia), Roman R.Shakarishvili (Georgia), Jozsef Vitrai (Hungary), Bruce Rosen (Israel), Maksut Karimovich Kuhlzhanov (Kazakhstan), Ineta Pirkšina (Latvia), Jorun Ramm (Norway), Tamara Maximova (Russian Federation), MariaLetkovicova (Slovakia), Tat Albreht (Slovenia), Rosa Matau (Spain), Adnan Kisa (Turkey), Volodimir Olexandrivich Kolenenko (Ukraine) SEARO—P Aronisamy (India), Kyi Soe (Myanmar), Sanjoy Nandy (Nepal), Thushara Fernando (Sri Lanka) WPRO—Russel Blamey (Australia), Keqin Rao (China), Buongnong Bousha (Laos Peoples’ Democratic Republic), Maimunah A Hamid (Malaysia), Laurie Ramiro (Philippines), Duong Huy Lieu (Vietnam).


Global Mental Health 1

No health without mental health

Martin Prince, Vikram Patel, Shekhar Saxena, Mario Maj, Joanna Maselko, Michael R Phillips, Atif Rahman

About 14% of the global burden of disease has been attributed to neuropsychiatric disorders, mostly due to the chronically disabling nature of depression and other common mental disorders, alcohol-use and substance-use disorders, and psychoses. Such estimates have drawn attention to the importance of mental disorders for public health. However, because they stress the separate contributions of mental and physical disorders to disability and mortality, they might have entrenched the alienation of mental health from mainstream efforts to improve health and reduce poverty. The burden of mental disorders is likely to have been underestimated because of inadequate appreciation of the connectedness between mental illness and other health conditions. Because these interactions are protean, there can be no health without mental health. Mental disorders increase risk for communicable and non-communicable diseases, and contribute to unintentional and intentional injury. Conversely, many health conditions increase the risk for mental disorder, and comorbidity complicates help-seeking, diagnosis, and treatment, and influences prognosis. Health services are not provided equitably to people with mental disorders, and the quality of care for both mental and physical health conditions for these people could be improved. We need to develop and evaluate psychosocial interventions that can be integrated into management of communicable and non-communicable diseases. Health-care systems should be strengthened to improve delivery of mental health care, by focusing on existing programmes and activities, such as those which address the prevention and treatment of HIV, tuberculosis, and malaria; gender-based violence; antenatal care; integrated management of childhood illnesses and child nutrition; and innovative management of chronic disease. An explicit mental health budget might need to be allocated for such activities. Mental health affects progress towards the achievement of several Millennium Development Goals, such as promotion of gender equality and empowerment of women, reduction of child mortality, improvement of maternal health, and reversal of the spread of HIV/AIDS. Mental health awareness needs to be integrated into all aspects of health and social policy, health-system planning, and delivery of primary and secondary general health care.

Introduction

The WHO proposition that there can be “no health without mental health” has also been endorsed by the Pan American Health Organisation, the EU Council of Ministers, the World Federation of Mental Health, and the UK Royal College of Psychiatrists. What is the substance of this slogan?

Mental disorders make a substantial independent contribution to the burden of disease worldwide (panel 1). WHO’s 2005 estimates of the global burden of disease provide evidence on the relative effect of health problems worldwide. Non-communicable diseases are rapidly becoming the dominant causes of ill health in all developing regions except sub-Saharan Africa (table 1). The Global Burden of Disease report has revealed the scale of the contribution of mental disorders, by use of an integrated measure of disease burden—the disability-adjusted life-year, which is the sum of years lived with disability and years of life lost. The report showed that neuropsychiatric conditions account for up to a quarter of all disability-adjusted life-years, and up to a third of those attributed to non-communicable diseases, although the size of this contribution varies between countries according to income level (table 1). The neuropsychiatric conditions that contribute the most disability-adjusted life-years are mental disorders, especially unipolar and bipolar affective disorders, substance-use and alcohol-use disorders, schizophrenia, and dementia. Neurological disorders (such as migraine, epilepsy, Parkinson’s disease, and multiple sclerosis) make a smaller but still significant contribution. Of the non-communicable diseases, neuropsychiatric conditions contribute the most to overall burden (figure 1 and table 1), more than either cardiovascular disease or cancer.

Search strategy

We searched relevant databases (Medline, PubMed, Embase, and the Cochrane Library of systematic reviews and clinical trials) with the following Mesh terms: “mental disorders”, “substance-related disorders”, “cardiovascular diseases”, “cerebrovascular disorders”, “diabetes mellitus”, “diabetes complications”, “HIV infections”, “malaria”, “tuberculosis”, “genital diseases”, “female”, “infant nutrition disorders”, “and accidents”, together with the PubMed clinical queries algorithms for aetiology, prognosis, treatment, and systematic reviews. For non-communicable disorders (coronary heart disease, stroke, and diabetes), and communicable disorders (HIV/AIDS, tuberculosis, and malaria) we focused on index conditions that are especially salient to public health. We concentrated on papers published since 2000, and have prioritised evidence from low-income and middle-income countries and from systematic reviews and meta-analyses. We have cited subsequent publications if they provided new information.
Despite these new insights, ten years after the first WHO report on the global burden of disease, mental health remains a low priority in most low-income and middle-income countries. Developing countries tend to prioritise the control and eradication of infectious diseases and reproductive, maternal, and child health, whereas developed countries prioritise non-communicable diseases that cause early death (such as cancer and heart disease) above those that cause years lived-with-disability (such as mental disorders, dementia, and stroke). If mental disorders are not included in mortality data, they will be under-represented. The WHO’s 2005 estimate (which is the proportion of cases of disability that would not have occurred in the absence of mental disorders) could be as high as 0.6–9.3% which suggests that failing health and consequent disability could be the most important contributory cause for late-life depression. Two studies suggest that disability is an equally powerful, although less prevalent, prospective risk factor for depression in young people.17,18 Conversely, disability is an important prospective risk factor for depression in older adults.7,19 and mediates most of the effects of specific physical health conditions in this group.20–22 Social support is an effect modifier.10,11,16 The population-attributable fraction (which is the proportion of cases of disability that would not have occurred in the absence of mental disorders) could be as high as 0.6–9.3% which suggests that failing health and consequent disability could be the most important contributory cause for late-life depression. Two studies suggest that disability is an equally powerful, although less prevalent, prospective risk factor for depression in young people.17,18

Mental disorders also contribute to disability. According to WHO’s 2005 estimates, neuropsychiatric disorders account for 1.2 million deaths every year and 1.4% of all years of life lost; most of these are caused by depression, Parkinson’s disease, and epilepsy.4 Only 40 000 deaths were attributed to mental disorders (mainly unipolar and bipolar depression, schizophrenia, and post-traumatic stress disorder) and 182 000 to use of drugs and alcohol.4 These numbers are almost certainly underestimated, since the report attributes death by suicide to intentional injury.4 Every year, about 800 000 people commit suicide, 86% of whom are in low-income and middle-income countries, and more than half of whom are aged between 15 and 44 years. Even these figures might be underestimated, since official statistics in low-income and middle-income countries are not reliable. For example, studies in south India that used surveillance with validated verbal autopsy showed that rates of suicide were ten times greater than the official national estimates;23 that suicide was the leading cause of death in 10–19 year olds; and that suicides accounted for a quarter of all deaths in boys and up to three-quarters of all deaths in young women.24 A systematic review of psychological autopsy case-control studies identified mental disorders (depression, schizophrenia and other psychoses, and alcohol-use and substance-use disorders) as important proximal risk factors for suicide, with a median prevalence of mental disorders25 and, if unrecognised and untreated, these disorders may contribute to disability.26

Our first aim is to critically appraise the way that the burden of disability and premature mortality is apportioned, in WHO’s estimates of global burden of disease, between underlying conditions within groups of disorder, and, specifically, to assess whether these estimates account for the full contribution of mental disorder to mortality and disability. Our second aim is to review available evidence for interactions between mental disorders and other health conditions (such as medically unexplained somatic symptoms, communicable diseases, maternal and perinatal conditions, non-communicable diseases, and injuries). Our third aim is to discuss the implications of these links for the future orientation of health policies, health systems, and services.

### Contributions of mental disorders to disability and mortality

Mental disorders are an important cause of long-term disability and dependability. WHO’s 2005 report attributed 31.7% of all years lived-with-disability to neuropsychiatric conditions: the five major contributors to this total were unipolar depression (11.8%), alcohol-use disorder (3.3%), schizophrenia (2.8%), bipolar depression (2.4%), and dementia (1.6%).27 However, the interaction between mental disorder and disability is more complex and extensive than the WHO report suggests. Depression predicts the onset and progression of both physical and social disability.54 Conversely, disability is an important prospective risk factor for depression in older adults.27,28 and mediates most of the effects of specific physical health conditions in this group.29–31 Social support is an effect modifier.32,33 The population-attributable fraction (which is the proportion of cases of disability that would not have occurred in the absence of mental disorders) could be as high as 0.6–9.3% which suggests that failing health and consequent disability could be the most important contributory cause for late-life depression. Two studies suggest that disability is an equally powerful, although less prevalent, prospective risk factor for depression in young people.34,35

Mental disorders also contribute to mortality. According to WHO’s 2005 estimates, neuropsychiatric disorders account for 1.2 million deaths every year and 1.4% of all years of life lost; most of these are caused by depression, Parkinson’s disease, and epilepsy.4 Only 40 000 deaths were attributed to mental disorders (mainly unipolar and bipolar depression, schizophrenia, and post-traumatic stress disorder) and 182 000 to use of drugs and alcohol.4 These numbers are almost certainly underestimated, since the report attributes death by suicide to intentional injury.4 Every year, about 800 000 people commit suicide, 86% of whom are in low-income and middle-income countries, and more than half of whom are aged between 15 and 44 years. Even these figures might be underestimated, since official statistics in low-income and middle-income countries are not reliable. For example, studies in south India that used surveillance with validated verbal autopsy showed that rates of suicide were ten times greater than the official national estimates;23 that suicide was the leading cause of death in 10–19 year olds; and that suicides accounted for a quarter of all deaths in boys and up to three-quarters of all deaths in young women.24 A systematic review of psychological autopsy case-control studies identified mental disorders (depression, schizophrenia and other psychoses, and alcohol-use and substance-use disorders) as important proximal risk factors for suicide, with a median prevalence of mental

### Panel 1: WHO classification of mental and behavioural disorders

1.1 (F00–F09) Organic, including symptomatic, mental disorders (dementia, delirium, and brain injury)
1.2 (F10–F19) Mental and behavioural disorders due to use of psychoactive substances (alcohol-use and substance-use syndromes, including harmful use, dependence, and withdrawal)
1.3 (F20–F29) Schizophrenia, and schizotypal and delusional disorders
1.4 (F30–F39) Mood (affective) disorders (mania, hypomania, bipolar affective disorder, and depressive episodes)
1.5 (F40–F48) Neurotic, stress-related, and somatoform disorders (phobic anxiety disorder, panic disorder, generalised anxiety disorder, obsessive-compulsive disorder, post-traumatic stress disorder, adjustment disorder, dissociative disorder, and somatisation disorder)
1.6 (F50–F59) Behavioural syndromes associated with physiological disturbances and physical factors (eating disorders, sleep disorders, sexual dysfunction)
1.7 (F60–F69) Disorders of adult personality and behaviour
1.8 (F70–F79) Mental retardation
1.9 (F80–F89) Disorders of psychological development
1.10 (F90–F98) Behavioural and emotional disorders with onset usually in childhood and adolescence (hyperkinetic disorders, emotional disorders, and conduct disorders)
1.11 (F99) Unspecified mental disorders

Note: The term ‘common mental disorder’ (CMD) refers to the most prevalent conditions classified under depressive episode, neurotic, stress-related, and somatoform disorders. The term also recognises that mental disorders in the community are frequently characterized by comorbidity between these groups and shifting patterns of symptoms that resist precise classification.

---

**Declarations of interest:** None.
disorder of 91% in suicide completers, and a population-attributable fraction of 47–74%. Findings from psychological autopsy studies in India and China were similar. Therefore, prevention, identification, and appropriate management of mental health problems is an important element of suicide prevention.

Mental disorder is independently associated with a substantial excess in all-cause mortality risk. Most studies have focused on associations with depression: a meta-analysis of 15 population-based studies reported that depression diagnosis was linked with subsequent all-cause mortality, and yielded a pooled odds ratio (OR) of 1.7 (95% CI 1.5–2.0). Several studies report that this association is mediated partly through disability, but not through cardiovascular disease, cardiovascular risk factors, or antidepressant use. Increased all-cause mortality, excluding suicides, has also been reported for schizophrenia (relative risk [RR] 2.59, 95% CI 2.55–2.63), bipolar disorder (standardised mortality ratio [SMR] 1.9 for men and 2.1 for women), and dementia (RR 2.63, 95% CI 2.55–2.63). In a record linkage study of mental health service users from western Australia, mortality from ischaemic heart disease was linked with most mental disorders, especially dementia and schizophrenia and other psychoses, although rates of admission for ischaemic heart disease were similar. People with schizophrenia (RR for men 0.31 [95% CI 0.21–0.45] and for women 0.34 [0.18–0.64]) and people with dementia (RR for men 0.30 [0.07–0.26] and for women 0.53 [0.16–1.74]) were much less likely to undergo revascularisation procedures such as coronary artery bypass grafting.

In the general population, between 1980 and 1998, ischaemic heart disease mortality fell by 34% in men and 13% in women, but in users of mental health services the rate was stable in men and had fallen by 34% in women.

Table 1: Contribution by different health conditions to disability-adjusted life-years, by income level of countries

<table>
<thead>
<tr>
<th>Condition</th>
<th>World</th>
<th>High-income countries</th>
<th>Middle-income countries</th>
<th>Low-income countries</th>
<th>Projected for 2030</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Communicable, maternal, perinatal, and nutritional conditions</td>
<td>572 392 000 (38.6%)</td>
<td>66 470 000 (5.6%)</td>
<td>99 696 000 (20.2%)</td>
<td>465 948 000 (53.5%)</td>
<td>494 384 000 (30.0%)</td>
</tr>
<tr>
<td>II Non-communicable diseases</td>
<td>725 306 000 (48.9%)</td>
<td>102 311 000 (85.7%)</td>
<td>318 415 000 (64.7%)</td>
<td>304 773 000 (35%)</td>
<td>938 468 000 (56.9%)</td>
</tr>
<tr>
<td>Neurropsychiatric conditions</td>
<td>199 606 000 (13.5%)</td>
<td>32 717 000 (27.4%)</td>
<td>87 398 000 (17.7%)</td>
<td>79 490 000 (17.7%)</td>
<td>92 590 000 (17.5%)</td>
</tr>
<tr>
<td>III Injuries</td>
<td>185 262 000 (12.5%)</td>
<td>10 403 000 (8.7%)</td>
<td>74 439 000 (15.1%)</td>
<td>100 420 000 (21.5%)</td>
<td>217 777 000 (13.2%)</td>
</tr>
</tbody>
</table>

DALYs=disability-adjusted life-years. Data are DALYs (proportion of total DALYs), unless otherwise specified. *Proportion of non-communicable disease DALYs caused by neuropsychiatric conditions.

Figure 1: Contribution by different non-communicable diseases to disability-adjusted life-years worldwide in 2005

Data adapted from WHO, with permission.

www.thelancet.com Vol 370 September 8, 2007 861
increased by 40% in women. Although evidence from low income countries is scarce, a large population-based study in Ethiopia indicated very high mortality rates for major depression (SMR 3.55, 95% CI 1.97–6.39) and for schizophrenia (nearly 5% per year). The association between alcohol use and mortality is complex, with a U-shaped association, and different effects according to cause of death; nevertheless, in the UK 8.5% of years-of-life lost to age 65 in men and 4.0% in women have been attributed to drinking more than the recommended alcohol limits. In Russia, alcohol-related mortality contributed to substantial fluctuations in the overall mortality rate in the 1990s.

**Mental disorders interact with other health conditions**

**Medically unexplained somatic symptoms**

Typically, at least a third of all somatic symptoms remain medically unexplained, both in the general population and in general medical-care settings. Common medically unexplained symptoms include pain, fatigue, and dizziness. Syndromes that represent characteristic organ-specific groups of medically unexplained symptoms have also been defined: irritable bowel syndrome, fibromyalgia, chronic-fatigue syndrome, chronic pelvic pain, temporo-mandibular joint dysfunction, and sexual-discharge syndromes. Medically unexplained somatic symptoms and syndromes are strongly associated with common mental disorders; however, at least a third of those with somatisation have no comorbid mental disorder. About 15% of patients seen in primary care have somatisation, which is defined as medically unexplained somatic symptoms coupled with psychological distress and help-seeking behaviour. Somatisation is independently associated with poor health-related quality of life and greatly increased use of health care, after controlling for comorbid mental disorder. In the USA, somatisation is estimated to contribute US$256 billion to health-care costs every year. Evidence from randomised controlled trials supports the effectiveness of specific intervention strategies such as structured treatment recommendations, antidepressant medication, and cognitive-behaviour therapy for reduction of somatic symptoms and health-care use. Health-care costs can be reduced by as much as a third. A pilot trial of cognitive-behavioural therapy for medically unexplained symptoms in Sri Lankan primary care (the only published trial from a developing country) also showed that treatment was associated with significant reductions in medically unexplained symptoms, visits, and distress.

**Non-communicable diseases**

Aside from neuropsychiatric disorders, the main contributors to disability and mortality from non-communicable disease are cardiovascular disease and cancer. Coronary heart disease and stroke account for 21% of disability-adjusted life-years in this group, and cancer for 12% (figure 1). Endocrine disorders (primarily diabetes) account for 3.7% of the disability-adjusted life-years attributed to non-communicable disease, and this proportion is predicted to rise sharply to 5.4% by 2030, with much of the increase in low-income and middle-income countries. Non-communicable diseases are a global challenge: they are the leading cause of death in all world regions other than sub-Saharan Africa, with 80% of deaths in low-income and middle-income countries.

**Cardiovascular disease**

A systematic review of evidence from population-based research reported moderate to strong prospective associations between depression (15/22 studies), anxiety (four of eight studies), and coronary heart disease. The outcomes studied included angina and non-fatal and fatal myocardial infarction. Population-based cohort studies also show that depression is an independent risk factor for non-fatal and fatal stroke. Follow-up periods in many of these studies were longer than ten years, which renders depression induced by preclinical cardiovascular disease an unlikely explanation. The effects were largely independent of risk factors for cardiovascular disease, since most of the cited studies comprehensively controlled for such factors.

The scarcity of evidence for risk mediation is surprising since mental health is strongly associated with cardiovascular risk exposures. Obesity, in a nationally representative survey in the USA, was associated with significant increases in lifetime diagnoses of major depression, bipolar disorder, and panic disorder or agoraphobia. Smoking, in population-based studies, is consistently shown to be associated with depressive and anxiety disorders and with schizophrenia. These associations might be bidirectional; prospective studies of young people indicate both that affective disorders and alcohol-use disorders could predict adoption of a daily smoking habit, and that tobacco use can be associated with the onset of common mental disorder. Findings from prospective population-based studies conflict as to whether mental disorders predict failure to quit smoking in those with the habit. In a study with a 7–16 year follow-up of participants, incident hypertension was independently predicted by both high depression scores (OR 1.8, 95% CI 1.2–2.8) and anxiety scores (1.8, 1.3–2.5) at baseline, after controlling for age, sex, education, smoking, body-mass index, alcohol use, history of diabetes or cardiovascular disease, and baseline systolic blood pressure.

The incidence of depression increases after myocardial infarction, to 15–30% for major depression, mostly in the first month after the event. Systematic reviews of prognostic studies report that comorbid depression is a consistent predictor of adverse outcomes (including recurrent coronary heart disease events, mortality from...
coronary heart disease, and all-cause mortality) after non-fatal myocardial infarction, after controlling for disease severity and treatment-related factors. Poor prognosis might be mediated partly by poor adherence by patients with depression to behaviour and lifestyle changes intended to reduce the risk of subsequent cardiac events. The evidence for anxiety as a prognostic factor is less strong. In a study based on the Maastricht stroke registry, the cumulative 1-year incidence of major depression was 23·3%. Two population-based incidence studies support a strong association between recent incident stroke and subsequent onset of depression, independent of disability. Depression after stroke is associated with poor functional outcomes and with a 3·4 times higher mortality over 10 years, after adjusting for baseline severity and type of stroke.

A Cochrane review of 36 trials of psychological interventions after myocardial infarction (18 of which focused on stress management) did not report an effect on total or cardiac mortality, but did show small reductions in anxiety and depression in patients with coronary heart disease. Few interventions have specifically targeted affective disorder. Antidepressants (selective serotonin-reuptake-inhibitors [SSRIs]) have been shown to be safe and moderately effective treatments for depression after myocardial infarction. A large trial of stepped-care cognitive behavioural therapy and SSRIs for depression and perception of low social support after myocardial infarction reported that the intervention was associated with significant improvement in mood and social support but not with improvement in event-free or overall survival. Therefore, more intensive and flexible patient-specific interventions have been advocated.

The evidence base for the effectiveness of antidepressants after stroke is weak. A Cochrane review of antidepressants as a preventive intervention reported no effect either on incident depression, or on reduction of disability or mortality. Another Cochrane review on pharmacological interventions for depression after stroke reported a reduction in symptoms, but not remission of diagnosable depression. Stroke recovery was not improved by pharmacological interventions. One trial subsequently published, with a 9-year follow-up, did show a sustained reduction in mortality after stroke, associated with antidepressant treatment.

Diabetes

Two US population cohort studies suggested that depression increases the risk for onset of type 2 diabetes, controlling for demographic, metabolic, and lifestyle factors; however, another large cohort study did not support this finding. The prospective associations might yet be explained by undetected diabetes leading to depression, or by help-seeking for depression leading to detection of diabetes. The evidence for comorbidity between mental disorder and diabetes is much stronger. The prevalence of diabetes in people with schizophrenia has consistently been shown to be about 15%, compared with a typical community prevalence of 2–3%. Much of this difference is probably explained by lifestyle factors, and some by the metabolic effects of typical and atypical antipsychotic medication. Abnormalities of glucose regulation were noted in people with schizophrenia before the use of antipsychotic medication, and independent of treatment in the modern era. The increased frequency of a family history of diabetes in people with schizophrenia also suggests an underlying mechanism specific to the disease. A meta-analysis of the association between depression and diabetes identified 20 controlled studies (of which 11 were population-based) with an OR for the association between the two conditions of 2·0 (95% CI 1·8–2·2); this ratio did not vary by type of diabetes, method for assessment of depression, or study design. Data on comorbid anxiety and diabetes were sparser, with only five controlled studies, one of which was population-based; the mean rate of generalised anxiety disorder in the clinical samples was 13·5%, which is much higher than the 3–4% typically seen in community studies. Comorbidity between diabetes and common mental disorder is important because of the implications for chronic disease management, and the effect on diabetic outcomes.

People with schizophrenia show poor adherence to oral hypoglycaemic therapy. Adherence to recommendations for diet and exercise, and to oral hypoglycaemic medication is low in diabetics with depression. In one study, however, attendance for screening by medical services to prevent complications was not affected by mood. Similar effects on adherence were noted for alcohol consumption in diabetics from ethnic minorities in Los Angeles. Poor mental health seems to have the greatest effects on patient-initiated behaviours that are difficult to maintain. The quality of diabetic care received by those with and without mental disorders, including serious mental illness, seems to be similar for most indicators, with the possible exception of those with substance-use disorders. Even so, meta-analyses suggest that both depression and anxiety are associated with poor glycaemic control. These cross-sectional associations are equally consistent with depression and anxiety being causes or consequences of poor glycaemic control. However, structural equation modelling in a prospective study suggested that the effect of depression on symptoms of glucose dysregulation is mediated through lower adherence to self-care. Depression in diabetes is consistently shown to be associated with diabetes complications, including retinopathy, nephropathy, macrovascular complications, and sexual dysfunction. Major depression (hazard ratio [HR] 2·3) and minor depression (HR 1·7) are significantly associated with mortality in type 2 diabetes. These associations were partly but not completely explained by extensive control for behavioural mediators and diabetes severity.

Evidence for the benefits of mental health interventions on these outcomes is mixed. Meta-analyses suggest that
psychological interventions in type 1 diabetes (in children only)\textsuperscript{103} and type 2 diabetes\textsuperscript{104} improve diabetic control. Participants in these trials were generally selected on the basis of risk factors for diabetes complications, such as poor glycaemic control, obesity, or inactivity, rather than depression. A large trial in nine US primary-care clinics reported that evidence-based collaborative depression treatment (consisting of pharmacotherapy, problem-solving treatment, or both in combination) for patients with diabetes and depression did not produce better effects than usual primary care on either glycaemic control\textsuperscript{105} or diabetic self-care,\textsuperscript{106} despite significant effects on depression outcomes.\textsuperscript{107} Similar findings were reported from two small randomised controlled trials of antidepressant treatment in diabetes.\textsuperscript{107,108}

### Communicable diseases

Communicable diseases continue to cause substantial death and disability in low-income and middle-income countries. HIV/AIDS (which causes 8.2% of all years-of-life lost) and malaria (which causes 4.5% of years-of-life lost) collectively account for nearly 13% of premature mortality and 39% of that attributable to communicable diseases. In 2004, about 34 million people were living with AIDS and over 3 million died of the disease. *Plasmodium falciparum* infects 500 million people each year and causes 2.7 million deaths, more than 90% of which are in young African children. The HIV epidemic and the emergence of strains with multiple drug-resistance has led to a resurgence of tuberculosis as a major public-health menace worldwide. In 2003, an estimated 8.8 million new cases of tuberculosis resulted in 1.7 million deaths; 27% of these cases and 31% of these deaths arose in Africa.\textsuperscript{109}

### HIV/AIDS

Some (mainly indirect) evidence shows that people with mental disorder are at heightened risk of contracting HIV/AIDS. Consistent evidence from the USA suggests that those with serious chronic mental illnesses have a high seroprevalence of HIV (5–7%), and that in those with schizophrenia, the mental illness generally precedes HIV infection.\textsuperscript{110} Behavioural risk factors identified (with a frequency of 30–60% in these high-risk groups) included high rates of sexual contact with multiple partners, low adherence to condom use, injected drug-use or sexual contact with injecting drug users, and unprotected sex between men.\textsuperscript{111} A large US cohort study of men who have sex with men provided more direct evidence: it identified use of alcohol and drugs before sex and depressive symptoms as independent predictors of seroconversion.\textsuperscript{112} Up to 10% of HIV cases worldwide are attributable to use of injection drugs.\textsuperscript{113} The evidence from low-income and middle-income countries is less clear; seroprevalence in psychiatric inpatients is often similar to that in the general population.\textsuperscript{114} Psychiatric inpatients in an Indian institution reported high rates of sexual and drug-related risk behaviours.\textsuperscript{115,116}

A fairly consistent association between infection with HIV and poor mental health has been reported. Several mechanisms might be implicated. Apart from psychological trauma (panel 2)\textsuperscript{117–123} the infection itself has direct effects on the central nervous system, and causes neuropsychiatric complications;\textsuperscript{118} depression,
mania, cognitive disorder, and frank dementia, often in combination. Although the incidence of HIV-associated dementia has halved since the advent of highly active antiretroviral therapy (HAART),125 and opportunistic infections of the central nervous system are rare,126 the incidence of HIV encephalopathy might have risen,127 suggesting continued infiltration of the central nervous system. Evidence for neurocognitive impairment in asymptomatic HIV-infected individuals has been found,128 although the severity and number of domains affected is greater in those with symptomatic disease.129 HAART, especially with efavirenz, can be associated with a range of side-effects on the central nervous system, including depression, nervousness, euphoria, hallucination, and psychosis.130 Patients with a previous history of psychiatric disorders could be at greater risk. Death by suicide has occasionally been reported. In a national probability sample of HIV-positive men and women in the USA, the 1-year prevalence of major depression was 36% and that of generalised anxiety disorder was 16%. These prevalences are five and eight times higher, respectively, than those identified by a national household survey that used the same assessment method.131 In a meta-analysis of studies that compared HIV-positive and HIV-negative control groups132 the difference in the prevalence of major depression (9.4% in HIV-positive vs 5.2% in HIV-negative) was significant (OR 2.0, 95% CI 1.3–3.0). A systematic review of the evidence from low-income and middle-income countries identified 13 studies of mental disorders in HIV-positive people; reported prevalence varied widely.133 The largest and best designed of these studies (which compared HIV-positive people who accessed HIV services with matched controls in Bangkok, Kinshasa, Nairobi, and Sao Paulo) reported that the rates of depressive disorder and depression symptoms were higher in symptomatic HIV-positive people, compared with either non-symptomatic cases or seronegative controls.134

Little evidence on associations between mental disorder and either help-seeking behaviour or uptake of diagnostic and treatment services for HIV/AIDS is available. In US women who were medically eligible, non-receipt of HAART was associated with substance use and with a history of childhood sexual abuse.135 Injection-drug use has consistently been shown to be associated with low uptake of HAART.136 Depression symptoms predicted drop out from a HIV-risk-reduction programme for socially deprived Latino women.137

Comorbidity affects prognosis. In US cohorts of HIV-positive women, chronic depressive symptoms were associated with increased AIDS-related mortality138,139 and with rapid disease progression,140 independent of receipt of treatment, and comorbidity use. Impairment in cell-mediated immunity (consisting of higher activated CD8 T lymphocyte counts and lower natural killer cell activity) might be implicated (panel 2).141 Cognitive impairment in HIV has been associated with greatly increased mortality independent of baseline clinical stage, CD4 cell count, serum haemoglobin, antiretroviral treatment, and social and demographic characteristics.142 Schizophrenia complicates treatment and has been associated with poor prognosis.143 The incidence of AIDS-defining illnesses in patients on HAART was reported to be especially high in injection-drug users.144

Adherence to HAART must be almost perfect to achieve lasting viral suppression. Adherence of less than 95% independently predicts viral resistance, hospital admissions, and opportunistic infections.145 Drug resistance can be transmitted to other people, which limits treatment options. Strong and consistent evidence from treatment programmes in developed countries now shows that adherence to HAART is adversely affected by depression,146–148 cognitive impairment,149–151 and alcohol-use and substance-use disorders.152 By contrast, adherence in the presence of serious mental illnesses can be good, presumably because of close medical supervision.153 We need to know more about adherence in low-income and middle-income countries.154 One study, from Uganda, which used a diagnostic assessment for depression, reported no association with adherence,155 whereas in Ethiopia depression was associated with less than 95% self-reported adherence in the week before interview.156 Data from a non-randomised US observational cohort study showed that antiretroviral adherence improved more in 6 months for those with depression who adhered to antidepressant treatment, compared with those not treated.157 We did not find any trials of the effect of antidepressants on adherence. Randomised controlled trials of motivational interviewing (for patients with alcohol problems) and adherence interventions (for those on methadone maintenance) suggested no sustained benefit for either approach.158 Modified directly observed treatment has been shown to improve adherence by substance users in one randomised controlled trial and one controlled trial.159

Findings on the effect of psychological interventions on psychopathology and HIV prognosis have been mixed. Group cognitive behavioural interventions have been tested extensively and shown to decrease depression-symptom scores,160 reduce herpes virus IgG titres,161 improve quality of life related to mental health,162 and reduce unsafe sexual behaviours.163 The evidence base for antidepressant treatment is surprisingly small, with only a few small randomised controlled trials and a much larger number of open-label interventions.164 Both tricyclic antidepressants and SSRI antidepressants seem to improve depression symptoms but have no effect on CD4 cell counts.165–167 Coverage and uptake are a challenge even in the USA, which has substantial resources; in a national survey of HIV-positive care recipients, about half those with depressive disorders did not receive antidepressants.168
Very few studies have investigated mental disorder as a predictor for HIV transmission, especially since the advent of HAART. One study of 168 HIV-infected men with resistance to antiretroviral drugs showed a high rate of high-risk sexual behaviour (such as unprotected anal or vaginal intercourse with an HIV-uninfected or status-unknown partner).158 These investigators reported strong evidence that depression, younger age, and sildenafil use predicted transmission, and moderate evidence that frequent alcohol use did so.

**Tuberculosis**

People with mental and substance use disorders might be at increased risk of contracting tuberculosis, although few studies have investigated this topic. A case registry study from Nagasaki suggested that the incidence of tuberculosis in patients with schizophrenia was high,155 similarly, high infection rates were recorded in people with serious mental illness in a psychiatric day programme in New York.156 Occasional reports of outbreaks in inpatients suggest that institutionalisation might contribute to risk of tuberculosis.157 A US population-based case-control study reported that heavy drinkers had twice the risk of tuberculosis infection of non-drinkers.158 Poor adherence to antituberculosis medication is an important barrier to global control of the disease, and increases the risks of morbidity, mortality, and drug resistance in both individuals and communities.159 Since treatment for multidrug-resistant tuberculosis is long (generally 2 years) and painful (consisting of daily injections for at least 6 months), with many unpleasant side-effects, adherence can be a challenge. A review of 13 treatment cohort analyses identified treatment-default rates of up to 39%, with an average of 12-6%.160 Alcohol-use disorder has also been reported to be associated with delayed treatment-seeking in Kiev; with poor adherence to directly observed therapy in New York;161,162 with unfavourable treatment outcomes for pulmonary tuberculosis in Kazakhstan163 and for multidrug-resistant tuberculosis in Tomsk, Russia;164 and with increased mortality in a US trial of directly observed therapy (HR 2-9).165 We identified only one report that did not find an association between psychiatric illness and substance use and poor adherence to tuberculosis treatment in a sample of homeless adults.166

Since depression has an important effect on adherence to treatment for many health conditions, the amount of research into comorbidity between tuberculosis and common mental disorders is surprisingly low. Multidrug-resistant tuberculosis, in particular, might be associated with poor mental health, attributed variously to loss of work and social roles and feelings of hopelessness and stigma.166 In Peru, the incidence of depressive disorder at recruitment into a treatment programme for multidrug-resistant tuberculosis was 52%, with further incidences of 13·3%, 12·0%, and 12·0% for depression, anxiety, and psychosis, respectively, during treatment.167 In an inpatient study in Turkey, the prevalence of depression, anxiety, or both was assessed to be 19% for recently diagnosed tuberculosis, 22% for defaulted tuberculosis, and 26% for multidrug-resistant tuberculosis.168 The prevalence of common mental disorders in 53 Nigerian tuberculosis patients recruited in a chest clinic was 30%, compared with 5% in healthy controls.169 A community-based study in Mali had suggested an even stronger association (OR 9-3), but with self-reported tuberculosis episodes.170

The failure of directly observed therapy to deliver improvement in treatment completion or cure171 has led to calls for rigorous investigation of extended interventions that address other factors known to influence adherence, such as quality of communication with treatment providers, patients’ health beliefs, patients’ education, and economic barriers.172 Since patients with multidrug-resistant tuberculosis face a range of difficulties, the development of strategies to support these patients will be essential to ensure treatment adherence. The information–motivation–behavioural skills model, which was originally developed to modify HIV-risk behaviour, has been recommended for use in tuberculosis treatment.172 Interventions based on cognitive behaviour therapy, which have proved helpful in management of various chronic diseases, have many similarities. In Peru, a non-randomised assessment of a group psychotherapy intervention, coupled with recreation, symbolic celebrations, and family workshops, was associated with a default rate of only 3·5% in a treatment cohort of 276 patients with multidrug-resistant tuberculosis.173 In India, a psychotherapeutic intervention based on behavioural-modification techniques was tested in a blind controlled trial with alternate allocation.173 Those who participated in the intervention were more likely than controls to complete treatment (72% vs 42%) and to be cured (72% vs 42%), and were less likely to default (17% vs 53%).173 The cost of the intervention was US$20 per patient, which was a quarter of the cost of the medication. In another non-randomised controlled trial, in Ethiopia, patients in tuberculosis clubs had significant improvements in treatment completion (69% vs 47%) and lower default rates (13% vs 41%), compared with controls.174

**Malaria**

No studies have investigated the possibility that mental disorders might increase susceptibility for malaria. Possible mechanisms could include effects on immunity, and on adherence to effective preventive measures. Severe *falciparum* malaria is associated with self-limiting psychiatric disorders,175 including depression,176,177 schizophrenic and manic syndromes, anxiety attacks,178 and confusional episodes.179 Treatment, especially with chloroquine, might be an associated factor.179 These syndromes might complicate and delay diagnosis.179 The extent of comorbidity between mental disorders and...
recurrent episodes of malaria, parasitaemia, or both has been very little studied. Dugbartey and colleagues compared 142 adult Ghanaians who had had a documented episode of malaria at least 12 months before the study with 30 community controls who had full medical records, no history of record of infection, and no existing parasitaemia. Patients with malaria had high scores for anxiety, depression, and total psychological symptoms, compared with controls. Carta and colleagues, in a small cross-sectional community survey in Mali, reported no association between acute malaria and common mental disorder. A systematic review provided strong evidence that malaria has both short-term effects on cognitive function and longer-term effects on cognitive development in children. Impairment is associated with the severity of the infection; cerebral malaria is especially important. Effects of non-severe malarial disease might be mediated through disrupted school attendance. In adults, a 1-year follow-up of cerebral malaria cases in Ghana reported no deficits, and a 20-year follow-up of Vietnam war veterans reported deficits in memory, language, and attention.

By comparison with work on tuberculosis, research on the effect of mental health on the prevention and effective treatment of malaria is scarce. Antimalarial programmes focus on intensification of preventive measures, (including use of insecticide-treated nets, which can reduce episodes of malaria in children by 50%), and encouragement of access to and uptake of affordable treatment within 24 h of onset. Recognised barriers to adoption of preventive health measures include poverty, inadequate education, knowledge and beliefs about malaria, and the complexity of preventive measures. For children, women are often the first to recognise the illness and have responsibility for illness management, although they might not have decisionmaking or financial control. Although the effects of illness-beliefs and attributions on help-seeking and self-treatment are increasingly well understood, three reviews suggest that mental health has not been regarded as relevant to help-seeking or self-treatment. Patient adherence is a major determinant of the therapeutic response to antimalarial drugs. A systematic review of 24 studies concluded that adherence was improved by interventions which focused on provider knowledge and behaviour, packaging, and provision of correct dosages. None of these studies discussed whether patients’ mental health (or maternal mental health status for children) would affect adherence to treatment.

Inappropriate overdiagnosis of malaria is also well documented; adverse consequences include drug side-effects, drug resistance, increased health-care costs, and failure to treat other causes of fever. In Africa, more than 70% of patients with suspected cases of malaria diagnose and manage their illness with traditional remedies or non-prescription drugs. A review suggests an average overestimation of 61% (range 32–96%) for clinical diagnosis, compared with a microscopy-based gold-standard diagnosis. In one series, 40% of those given a clinical diagnosis did not present with pyrexia. Somatisation might well account for a proportion of misdiagnosed cases.

Reproductive and sexual health
Women are at heightened risk for common mental disorders: a female to male sex ratio of 1-5 to 2-0 is typical. In Pakistan, the prevalence of common mental disorders in men is similar to that in other regions, but women are two to three times more likely than men to suffer from such disorders. Gender affects many of the determinants of mental health, including socioeconomic position, access to resources, social roles, rank, and status; and gender differences in mental disorders diminish after controlling for these mediators. The gendered disadvantage experienced by women in many parts of the world might be a relevant factor; for example, a large cross-sectional survey in Goa, India identified strong associations between common mental disorders and indicators of disadvantage, including early age at marriage, intimate partner violence and abuse, and absence of decisionmaking autonomy.

A systematic review identified 122 studies of the association between mental disorder and gynaecological morbidity. Sexual and other forms of abuse, anxiety, depression, and use of substances and alcohol were robustly and consistently reported to be associated with various reproductive health outcomes, including dysmenorrhoea, dyspareunia, and non-cyclical pelvic pain. Studies in south Asia, where abnormal vaginal discharge is a common complaint, report similar associations. Gynaecological complaints might be somatic idioms for common mental disorders; in the Goa study, the complaint of vaginal discharge was associated with symptoms of common mental disorder (OR 2-2, 95% CI 1-4–3-2) and somatoform disorders (6-2, 4-0–9-7), but not with reproductive-tract infection diagnosed with gold-standard laboratory tests (1-2, 0-9–1-6). In Asian cultures, explanatory models of reproductive health and mental health experiences might enhance the association between these health domains.

Maternal and child health
Maternal psychosis affects infant growth and survival. Maternal schizophrenia is consistently associated with preterm delivery and low birthweight. The effect of maternal psychosis on child survival has also been investigated—a meta-analysis linked maternal psychosis with a two-fold increased risk of stillbirth or infant mortality. Postpartum depression affects 10% to 15% of women in developed countries, with adverse consequences for the early mother–infant relationship and for children’s psychological development. In low-income and middle-income countries, the prevalence
of perinatal depression is, if anything, somewhat higher than in the developed world. Physical development of infants is a particular problem in Asia. An independent association between antenatal common mental disorder and low birthweight has been shown by two prospective studies: one from Pakistan (RR 1.9, 95% CI 1.3–2.9) and one from India (OR 1.4, 95% CI 1.0–2.1). Findings from high-income countries have been equivocal, with several negative reports. However, associations between maternal depression and preterm birth and between psychosocial stressors and low birthweight were reported from a disadvantaged African–American community.

In south Asia, two case-control and two cohort studies have consistently shown associations between perinatal common mental disorders and infant undernutrition at 6 months, after controlling for birthweight. However these studies did not assess the relative, independent contributions of antenatal and postnatal common mental disorders, and only one controlled for maternal nutrition. In the cohort study from Pakistan, 6-month old infants of antenatally depressed mothers were at much higher risk of being underweight (RR 4.0, 95% CI 2.1–7.7) and stunted (4.4, 1.7–11.4), after adjusting for birthweight, socioeconomic status, and frequent diarrhoea. In the same study, children of antenatally depressed mothers were also more likely to have had more than five diarrhoeal episodes in the first year of life (RR 2.3, 95% CI 1.6–3.1). In South Africa, neither postnatal nor current depression was associated with infant growth at 2 months, after adjusting for birthweight; however, there was a non-significant association at 18 months, and the study was small and had low power. A multicountry study that assessed common mental disorders in mothers contemporaneously with child growth at 6–18 months postpartum reported no cross-sectional association, in Ethiopia, between maternal mental health and child malnutrition, but did note that common mental disorders in mothers were associated with infant stunting in India, and with underweight infants in Vietnam. The longer-term effects of maternal mental health on infant growth or mortality have not yet been studied in low-income and middle-income countries.

A review reported that the effect of maternal depression on child cognitive development has been studied less extensively in low-income and middle-income countries than in developed countries. In south India, maternal postnatal common mental disorder was negatively associated with mental-development quotient scores in infants at 6 months, but not with motor development. In Barbados, a long-term prospective study reported associations between maternal common mental disorder and impaired cognitive and motor development in infants at 6 months, and poor performance in high-school entrance examinations in children aged 11–13 years.

Strong but not consistent evidence from developed countries shows that maternal depression reduces adherence to child-health promotion and prevention

<table>
<thead>
<tr>
<th>Non-communicable diseases</th>
<th>MD is a risk factor for the HC</th>
<th>MD is a consequence of the HC</th>
<th>Comorbiditity (uncertain causal direction)</th>
<th>MD affects adherence to treatment for HC</th>
<th>MD affects prognosis or outcome of the HC</th>
<th>Treatment for MD affects mental health in those with HC</th>
<th>Treatment for MD affects physical HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression and CMD with coronary heart disease</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>−1</td>
</tr>
<tr>
<td>Depression with stroke</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>−1</td>
<td>−1</td>
</tr>
<tr>
<td>Common mental disorder with diabetes</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Schizophrenia with diabetes</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Communicable diseases</th>
<th>MD is a risk factor for the HC</th>
<th>MD is a consequence of the HC</th>
<th>Comorbiditity (uncertain causal direction)</th>
<th>MD affects adherence to treatment for HC</th>
<th>MD affects prognosis or outcome of the HC</th>
<th>Treatment for MD affects mental health in those with HC</th>
<th>Treatment for MD affects physical HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression and CMD with HIV/AIDS</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Serious mental illness with HIV/AIDS</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cognitive impairment and dementia with HIV/AIDS</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alcohol-use and substance use disorder with HIV/AIDS</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2*</td>
</tr>
<tr>
<td>CMD with malaria</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cognitive impairment with malaria</td>
<td>0</td>
<td>4</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alcohol-use disorder with tuberculosis</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Depression or common mental disorder with tuberculosis</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maternal and child health</th>
<th>MD is a risk factor for the HC</th>
<th>MD is a consequence of the HC</th>
<th>Comorbiditity (uncertain causal direction)</th>
<th>MD affects adherence to treatment for HC</th>
<th>MD affects prognosis or outcome of the HC</th>
<th>Treatment for MD affects mental health in those with HC</th>
<th>Treatment for MD affects physical HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal depression and CMD with impaired child growth and development</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Maternal psychosis with infant mortality</td>
<td>4</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
</tr>
</tbody>
</table>

MD=mental disorder. HC=health condition. CMD=common mental disorder. NA=data not available. 4=strong evidence from meta-analysis or systematic review. 3=consistent evidence from several studies. 2=evidence from one study only. 1=inconsistent evidence. 0=no evidence identified. −1=negative reports. *This disorder affects adherence to treatment.
Injuries

Injury and violence are important causes of death and disability worldwide. The 2005 WHO report estimated that 5·4 million deaths from injury accounted for 9% of deaths worldwide and 12% of the global burden of disease, and that such deaths would increase substantially by 2030. Mental health problems are both a cause and a consequence of injury. Injury and mental disorder also have many determinants in common, such as poverty, conflict, violence, and alcohol use. Any public-health approach to injury control must consider mental health.

Road-traffic accidents are responsible for about 1·2 million deaths and perhaps ten times as many permanent disabilities each year. Three-quarters or more of the deaths are in developing countries, where numbers of accidents and fatalities have been increasing at an alarming rate. In low-income and middle-income countries, poor people (pedestrians, passengers in buses and trucks, and cyclists) suffer a higher burden of morbidity and mortality from traffic injuries. In 1964, a US study showed that alcohol was a strong risk factor for involvement in road traffic accidents, and this finding has been substantiated by many epidemiological studies. Although data are scarce, alcohol is implicated in a large proportion of road traffic accident deaths in low-income and middle-income countries. Nevertheless, variations between countries are apparent; in China the proportion of alcohol-related traffic accidents might be as low as 1%. A proportion of unintentional injuries might be recognised suicide attempts; a US study noted that the rate of suicide was ten times higher in those with at least one previous hospital admission for injury, and almost three times higher for drivers who had been injured in a road traffic accident.

Earlier reports of cross-sectional associations between maternal depression and child injury risk have been supported by the findings of a 10 000-family cohort study in the UK; maternal postnatal depression was prospectively and independently associated with burns or scalds (1·29, 95% CI 1·01–1·64) and with two or more accidents during the follow-up period (1·39, 1·16–1·66). Up to 98% of child injury deaths happen in low-income and middle-income countries; one study reported strong and consistent cross-sectional associations between common mental disorder in caregivers and injuries in children in India, Peru, Vietnam, and Ethiopia. Evidence for an inverse association between maternal depression and self-reported accident-prevention practices is less consistent. US studies reported an inverse association with preventive practices (such as use of car seatbelts and electrical plug covers), but a UK study found no association in socioeconomically deprived families with practices such as use of fireguards, stair gates, smoke alarms, window locks, or safe storage of medicines.

Injury and violence are also important risk factors for mental disorder. Post-traumatic stress disorder is a recognised consequence of non-intentional injury; analysis of data from the 1958 British birth-cohort study showed that injury and burns were strongly associated with psychological distress. Child abuse is a potent risk factor for psychiatric disorders and suicidal behaviour; intimate partner violence is a risk factor for depression, anxiety, and suicide; sexual violence is a risk factor for mental health and behavioural problems; and collective violence is a risk factor for depression, substance abuse, and suicide. Conflict was responsible for an estimated 184 000 deaths in 2005. Post-traumatic stress disorder is a common psychological outcome of conflict, with a quarter or more of survivors affected.

Panel 3: Modelling the effect of extended coverage of treatment for depression on health outcomes

We did a modelling exercise to assess the possible benefits to public health of extension of the coverage of evidence-based treatments for depression in low-income and middle-income countries. We focused on observational research that had produced strong evidence for associations between depression and other health conditions: maternal depression as a risk factor for infant stunting in Pakistan and major depression as a risk factor for suicide in China.

We calculated the population-attributable risk with the method that Morgenstern and Bursic used to estimate deaths that could theoretically be prevented by better coverage of evidence-based management of diabetes. This method allowed us to factor in both partial coverage of the intervention (in a range from 25% to 75%) and partial effectiveness (with a conservative 40% net reduction attributable to the intervention in those covered by the intervention, and no reduction in those with no coverage). We based these estimates on findings of associations from observational research; the prevention benefits estimated in these models would only accrue in reality if the associations were causal and estimated free of confounding, and if the effective treatment for depression reduced the risk of a previously depressed person to the same level as that of someone who had never been exposed. Randomised controlled trials will be needed to establish the real extent of the benefit.

(Continues on next page)
To model prevention of infant stunting in Pakistan we based our calculations on a 25% prevalence of depression in mothers and a relative risk of 4.4 for the association of maternal depression with infant stunting at 6 months.219 We predicted that up to 8% of stunting would be averted at 25% coverage, rising to 20% at 75% coverage. Since 92 000 stunted infants are born each year in Pakistan (comprising 31% of all births), this estimate would translate into a maximum of 13 800 cases of stunting averted every year, if 50% coverage could be achieved (figure 2). The developmental and health consequences of stunting are expected to decrease an adult’s yearly income by 20% (US$144 at current income per head) which would imply a nationwide saving of US$1.99 million every year. Furthermore, for a 10% reduction in the number of stunted children, the number of children who completed primary school education would be expected to increase by 8%.214

To model prevention of suicide in China we based our calculations on a 4.3% prevalence of depression with infant stunting at 6 months.23 We predicted that up to 8% of suicides would be averted at 25% coverage, rising to 20% at 75% coverage. Since 325 581 suicides happen every year in China, we estimated that if 50% coverage with the intervention could be achieved, a maximum of 32 558 suicides would be averted every year. The potential economic effect could be substantial, with 5.8 million productive life-years lost nationally, which would translate to lost productivity of US$10.2 billion per year because of suicide (on the basis of GDP per head of US$1740 in 2006). If 50% treatment coverage was achieved, a 10% reduction in the suicide rate would save US$80 000 productive years of life, or US$1.0 billion per year. Alternatively, we used willingness-to-pay estimates268 and the estimate of US$24.458 as the value of a statistical life in China, to calculate a saving of US$1.1 billion per year.

**Figure 2: Proportion of health problems theoretically prevented by increased coverage of evidence-based treatment for depression**

To model prevention of suicide in China we based our calculations on a 4.3% prevalence of major depression260 and a relative risk of 4.4 for the association of depression with suicide.219 We predicted that a maximum of 6% of suicides would be averted at 25% coverage, rising to 20% at 75% coverage (figure 2). Since 325 581 suicides happen every year in China, we estimated that if 50% coverage with the intervention could be achieved, a maximum of 32 558 suicides would be averted every year. The potential economic effect could be substantial, with 5.8 million productive life-years lost nationally, which would translate to lost productivity of US$10.2 billion per year because of suicide (on the basis of GDP per head of US$1740 in 2006). If 50% treatment coverage was achieved, a 10% reduction in the suicide rate would save US$80 000 productive years of life, or US$1.0 billion per year. Alternatively, we used willingness-to-pay estimates268 and the estimate of US$24.458 as the value of a statistical life in China, to calculate a saving of US$1.1 billion per year.

**Implications for policy, practice, and research**

WHO estimates of the global burden of disease have helped to raise awareness of the enormous effect of mental disorders, both in their own right and relative to other health conditions. Much of this effect arises from the commonest disorders, especially depression and alcohol-use disorder. However, the Cartesian dualism that is implicit in the methods used to generate these estimates has meant that what began as a blessing is now, in some respects, a bane. In reality, the interactions between mental disorders and other health conditions are widespread and complex (table 2). Mental disorders are risk factors for the development of communicable and non-communicable diseases, and contribute to accidental and non-accidental injuries. For some infectious diseases, mental disorders in infected persons increase the risk for transmission. Many health conditions increase the risk for mental disorder, or lengthen episodes of mental illness. The resulting comorbidity complicates help-seeking, diagnosis, quality of care provided, treatment, and adherence, and affects the outcomes of treatment for physical conditions, including disease-related mortality. For many health conditions, mental illness makes an independent contribution to disability and quality of life.

Mental health is missing from the policy framework for health improvement—and poverty reduction; missing from health and social research; and missing from targets for interventions. Moreover, mental health has not been acknowledged as an obstacle to achievement of several Millennium Development Goals—notably, promotion of gender equality and empowerment of women, reduction of child mortality, improvement of maternal health, and reversal of the spread of HIV/AIDS, malaria, and other diseases.

Mental health awareness needs to be integrated into all elements of health and social policy, health-system planning, and health-care delivery. Sophisticated evidence-based arguments to increase resources for mental health care should be linked to evidence for its wider importance to public health.259 Integrated mental health policies, applied across disease categories, and to different levels of care and types of care setting, will maximise the effectiveness of the small number of mental health professionals available in most low-income and middle-income countries.258 Such policies will also mobilise the forces of public and community health to work for better mental health and reduce redundancies and budgetary and organisational inefficiencies in overstretched health systems. The strengthening of health-care systems to deliver mental health care should focus, where possible, on existing programmes and activities such as HIV prevention, antiretroviral treatment programmes, treatment of multidrug-resistant tuberculosis, campaigns against gender-based violence, antenatal care, integrated management of childhood illness, and innovative chronic-disease management.259

Mental health needs to be recognised as an integral component of practice in primary and secondary health care. Beyond this, primary health-care workers need to be trained in recognition and evidence-based treatment of mental disorders, and given suitable supervision and
support. Basic drug and psychotherapeutic treatments need to be made available at all levels of health care—the evidence for treatment of specific disorders is presented later in this Series.26 Primary and secondary care providers should overcome their reluctance to treat patients with severe mental illnesses, and learn effective ways to interact and communicate with these patients. Inequities in access and provision of good-quality physical-health care for people with mental disorders must be ended. We need to promote holistic models of care, which integrate psychosocial assessments and interventions seamlessly and routinely into the management protocols for major communicable and non-communicable diseases and reproductive and childhood disorders. For example, our modelling exercises indicated that up to 20% of infant stunting could be averted if maternal depression was treated more effectively, and that up to 15% of suicides could be averted by interventions to treat major depression (panel 3). By the same token, mental health professionals should routinely assess their patients to identify and monitor physical-health problems, should encourage them to attend regular checks in primary care, and should generally place a greater emphasis on lifestyle review and management. Current guidelines about the management of patients given antipsychotic drugs should be applied; for example, patients with schizophrenia should be weighed at every visit. Although more mental health specialists are needed, these might never be sufficient to meet the need, especially in low-income countries. The marshalling of this scarce resource will demand careful thought and planning, including clear protocols for referral from primary care.

Evidence for interactions between mental health and other health conditions comes overwhelmingly from the developed world, especially the USA. Whereas 80% of deaths from non-communicable diseases are in low-income and middle-income countries, all but four of the 59 papers cited in the non-communicable disease section of this review describe research from north America and Europe. Although 99% of deaths from HIV/AIDS are in low-income and middle-income countries, nearly all research on the interaction between mental disorders and chronic management of HIV infection comes from the USA. 99% of deaths from malaria are in low-income countries and 90% of these are in children aged younger than 5 years; we identified an absence of evidence, rather than evidence of absence, for what could be, by analogy with other evidence, important interactions between maternal mental health, adherence to malaria prevention measures, and prompt and appropriate help-seeking for childhood infections.

The first priority, therefore, is to increase the evidence-base for interactions between mental health and other health conditions in low-income and middle-income countries. Some existing evidence (eg, that which investigates mental disorders as risk factors and prognostic indicators for non-communicable diseases) might be generalisable to less well developed settings. However, the evidence on maternal depression and infant growth outcomes is reported mainly from low-income and middle-income countries. Only research that is conducted locally can be expected to affect awareness and lead to new policy development.

Second, we need to understand better the mechanisms that underlie interactions between mental health and other health conditions, if we are to develop effective public-health and clinical interventions (panel 2). We need to learn from the experience that, in many instances, interventions designed to treat common mental disorders are effective for reduction of the frequency of these conditions, but not for improvement of downstream physical-health outcomes with which associations had been reported.76–79,105–108,150–152 Explicit targeting of illness representations and associated behaviours through cognitive behavioural techniques might be effective.

Third, we are as yet at a very early stage in the development and trialling of adjunctive psychosocial, psychological, and mental health interventions. Despite strong evidence for relevance of mental health to HIV/AIDS, well designed trials to investigate effects of mental illness on the important downstream health outcomes are scarce; for example, presentation for voluntary testing and counselling, access to and acceptance into HAART programmes, adherence, adoption of low-risk behaviours, virological and immune status, and survival.

We have stressed the potential capacity for psychosocial interventions to improve physical-health outcomes (eg, as shown for glycaemic control in diabetes,104 and as modelled in panel 3 for infant stunting in Pakistan,219 and suicide prevention in China23). However, we need also to act immediately on the existing robust evidence that treatment of comorbid mental disorder is highly effective for improvement of mental health and quality of life outcomes across a range of disorders including cancer,250 diabetes,35 heart disease,7,79 and HIV/AIDS.146,148 The moral and ethical case for redressing the imbalance in provision for people with mental disorders can brook no delay.273 Practical steps such as those discussed in this Series must be accompanied, wherever possible, by high quality assessments of efficacy and cost-effectiveness.

Acknowledgments
We thank the Lancet Global Mental Health Group, Michael Dewey for advice on the modelling exercise, and Simon Wessely and Ian Roberts for comments and contributions to the draft manuscript. VP is supported by a Wellcome Trust Senior Clinical Research Fellowship in Tropical Medicine. AR is supported by a Wellcome Trust Career Development Fellowship in Tropical Medicine. The Lancet Global Mental Health Series was supported by a grant from the John and Catherine MacArthur Foundation. SS is an employee of WHO; the views expressed in this article do not necessarily represent the decisions, policy, or views of WHO.

References


Volume 370 | September 8, 2007 876
Resources for mental health: scarcity, inequity, and inefficiency

Shekhar Saxena, Graham Thornicroft, Martin Knapp, Harvey Whiteford

Resources for mental health include policy and infrastructure within countries, mental health services, community resources, human resources, and funding. We discuss here the general availability of these resources, especially in low-income and middle-income countries. Government spending on mental health in most of the relevant countries is far lower than is needed, based on the proportionate burden of mental disorders and the availability of cost-effective and affordable interventions. The poorest countries spend the lowest percentages of their overall health budgets on mental health. Most care is now institutionally based, and the transition to community care would require additional funds that have not been made available in most countries. Human resources available for mental health care in most low-income and middle-income countries are very limited, and shortages are likely to persist. Not only are resources for mental health scarce, they are also inequitably distributed—between countries, between regions, and within communities. Populations with high rates of socioeconomic deprivation have the highest need for mental health care, but the lowest access to it. Stigma about mental disorders also constrains use of available resources. People with mental illnesses are also vulnerable to abuse of their human rights. Inefficiencies in the use of available resources for mental health care include allocative and technical inefficiencies in financing mechanisms and interventions, and an overconcentration of resources in large institutions. Scarcity of available resources, inequities in their distribution, and inefficiencies in their use pose the three main obstacles to better mental health, especially in low-income and middle-income countries.

Introduction

Mental health is an integral and essential component of health.1 Human, social, and financial resources will be needed to achieve the World Health Report objective of adequate access to effective and humane treatment for those who suffer from a mental disorder.2 We review here the availability, distribution, and use of such resources for mental health care worldwide. We have summarised available evidence, including from the relevant WHO publications and databases. The limitations of our review include its selective, rather than systematic, nature and its focus on mental health services, rather than on prevention and promotion, which have been discussed elsewhere.3,4 The scope of our review is global but data show that the severest examples of scarcity of resources, inequity of distribution, and inefficiency of resource-use are in low-income and middle-income countries (as per the World Bank's classification).

These three themes—of scarcity, inequity and inefficiency—are inter-related and often seem to accentuate each other. For example, countries with fewer mental health resources commonly distribute them less equitably because they rely on private rather than collective financing mechanisms. In turn, the general neglect of mental disorders in under-resourced health systems can affect not only national productivity, but also individual quality of life.

Scarcity of resources

Information on resources for mental health care has been scant compared with information on prevalence, type, and burden of mental disorders. However, analysis of data from WHO’s Atlas project (panel 1) shows widespread, systematic, and long-term neglect of resources for mental health care in low-income and middle-income countries.6

Policy and infrastructure

Mental health policies and plans for their implementation are essential for coordination of services and activities to improve mental health and reduce the burden of mental disorders.11 The elements of such policies must be determined to some extent by local circumstances, but key components can be identified and recommended.12 About a third of all countries in the world have no such policy or plan,1 and in the African region, for example, this proportion is nearly half. Moreover, because nearly 40% of countries that do have policies have not revised them since 1990, these policies do not incorporate the substantial recent developments in mental health care.

A mental health policy framework must include legislation for protection of the basic human and civil rights of people with mental disorders, especially those in receipt of involuntary treatment. 135 (78%) countries, with 69% of the world’s population, have laws about mental health; the rest do not have specific legal protection for people with mental illness. About half the existing laws are more than 15 years old, and 16% were enacted before 1960, when the human rights of people with severe mental disorders began to receive greater attention.

Discrimination against people with mental disorders is widespread, often formalised, and sometimes even codified in law.15 For example, although most countries have some provision for disability benefits, 41 (22%) of countries worldwide, and 26 (45%) of low-income countries,
specifically exclude mentally ill people from such entitlements. Another example of systematic discrimination is exclusion of mental disorders from some social and private insurance schemes for health care, for example in the USA, some European countries, and China. Mental health services

A balance of community-based and hospital-based services has been shown to be the most effective form of comprehensive mental health care. Yet such a balance has only been achieved in a few high-income countries, where financial resources have been matched by the political will to increase community care. If community-based mental health care is defined broadly, as “any type of care, supervision and rehabilitation of patients with mental illness outside the hospital by health and social workers based in the community”, then only about half the countries in Africa, the eastern Mediterranean, and southeast Asia provide such care. Within countries, the balance of services varies widely; for example, community-based care is restricted to only a few areas in China, India, Paraguay, and Zambia. Overall about 52% of low-income countries and about 97% of high-income countries provide community-based care.

Evidence suggests that, in low-income and middle-income countries, support for primary care services to enable them to identify and treat people with mental disorders, with training, assistance, and supervision by available specialist mental health staff, is the best way to extend mental health care to the population. A systematic review of community-based models of care for adults with depression, schizophrenia, panic disorder, and bipolar disorders in low-income and middle-income countries has reported that such models did improve clinical outcomes, with some cost savings. Five key areas for expansion of primary care to achieve general adult mental health care have been identified, although they should be adapted to suit local conditions: outpatient or ambulatory clinics; mobile community mental health teams for outreach services; acute inpatient care; long-term community-based residential care; and rehabilitation, occupation, and work. Worldwide only 11% (59%) of all countries have facilities to train primary-care workers in mental health care. Appropriate training needs to be combined with continued supervision and support to achieve effective mental health care in primary-care settings.

A very large or very small number of psychiatric beds relative to the mental health budget can indicate that services for people with serious mental disorders are not adequate. If most of these beds are in large institutional settings such as mental hospitals, custodial care tends to be the standard mode for treatment of serious mental disorders, along with insufficient choice in treatments, services that are distant from the homes of residents, and inflexibility in the resource base. Most countries in Africa and Asia have too few beds, and a large proportion of these beds are in mental hospitals; a median of 0.34 beds are available per 10 000 population in Africa and 0.33 in South-East Asia Regions, with 73% and 83% in mental hospitals, respectively. South Africa has relatively few inpatient beds, although numbers vary between provinces, and yet services for people with mental disorders are perhaps better here than most countries in Africa. By contrast, the European region has a median of eight beds per 100 000 population, although in the low-income and middle-income countries of central and eastern Europe more than 80% of these are in mental hospitals.

Unavailability of essential medicines also constrains mental health treatment. About a quarter of low-income countries do not provide even basic antidepressant medicines in primary-care settings. In many others, the supply does not extend to all regions of a country or is irregular, despite the fact that effective pharmacological treatment for many disorders depends on continuous access to medication for extended periods. Since medicines are often not available in health-care facilities, patients and families can be forced to pay for them. Because of the disproportionate prevalence of mental health problems in lower income groups, mental health care can be unaffordable for some groups, and thus inequitable. Moreover, the cost of essential medicines is relatively high in low-income countries: for example, a 1-year supply of one of the least expensive antidepressant medicines costs only twice in high-income countries what it costs in low-income countries, whereas gross national product (GNP) per head in these countries differs by a factor of 12.5.

Panel 1: WHO’s 2005 Mental Health Atlas

In 2000, the World Health Organization (WHO) initiated the Mental Health Atlas project to fill gaps in global information on mental health resources and services. The objectives included collection, compilation, and dissemination of global information about mental health resources and services in each country. The first edition of Atlas was published in 2001 and the second revised edition in 2005. The primary sources for the Atlas information were the Ministries of Health of each country, or institutions nominated by them. Other sources (especially for the 2005 edition) included a literature search on mental health resources in low-income and middle-income countries, documents received from countries, travel reports submitted by WHO staff, country data collected by WHO Regional Offices, and feedback from experts and WHO Collaborating Centres within countries. Information was categorised into broad themes such as mental health policy, programmes, financing, and resource indicators such as beds, personnel, services for special populations, and availability of drugs. The Mental Health Atlas 2005 covers all 192 WHO member states and 11 associate members, areas, and territories, which represent about 99% of the world’s population. Limitations of the data include the fact that some countries could not provide specific details on all variables. Some of these data simply do not exist within the countries, and accurate national figures are unobtainable under the decentralised organisation of many countries. Varying definitions of terms across countries also made aggregation and comparisons difficult. Financial data are difficult to obtain from countries on a consistent basis because of differences in the way that health systems are structured. Beside the main study for Mental Health Atlas 2005, Project Atlas also covers child and adolescent mental health, psychiatric training, and the role of nurses in mental health care.
Community resources

Care of people with mental disorders draws on community resources that include formally structured bodies such as international and indigenous non-governmental organisations (NGOs); consumer and family associations; and informal resources of family, friends, and other social networks that often bear most of the burden of care.\(^2\)

Community resources also include traditional, indigenous, and alternative health-care systems and community-based social and rehabilitative services.

88% of countries have at least one NGO that is active in mental health.\(^3\) Common NGO activities include advocacy, mental health promotion, prevention of mental disorders, rehabilitation, and direct service provision.\(^4\) The NGO sector has sustained and increased its efforts towards the achievement of access to mental health care for all who need it. NGOs often innovate by development of new services, or supplement inadequate state infrastructure for mental health care. NGOs also provide care in countries affected by conflicts, wars, and disasters such as the 2004 Asian tsunami.\(^5\) However, in most low-income and middle-income countries the population coverage and the range of services provided by NGOs are not comprehensive. Sustainability of NGO activities will depend on efforts to build up locally controlled structures and effective collaboration between NGOs and governments. Consumer and family associations have also become more established and active in many low-income and middle-income countries,\(^6\) although they tend to be weak or fragmented in countries where need is greatest. People with mental health needs and their families tend to have few opportunities to participate in decision-making about treatment; this is true in all countries, but especially in those with low and middle incomes.\(^7\)

**Human resources**

Mental health care relies on professionals, rather than advanced technology or equipment. Shortages of psychiatrists, psychiatric nurses, psychologists, and social workers (see figures 1 and 2) hinder treatment and care in low-income and middle-income countries. Studies from...
several African countries show that inadequate numbers of health-care professionals are the main limiting factor in psychiatric care. The serious shortage of psychiatrists in low-income countries is illustrated by Chad, Eritrea, and Liberia (with populations of 9, 4·2, and 3·5 million, respectively), which have only one psychiatrist in each country, and by Afghanistan, Rwanda, and Togo (with populations of 25, 8·5, and 3 million, respectively), which have just two psychiatrists each.5

Low-income countries have a median of 0·05 psychiatrists and 0·16 psychiatric nurses per 100 000 population. High-income countries have a ratio of psychiatric health-workers to population that is about 200 times higher. These figures show the huge inequities in the distribution of skilled human resources for mental health across the world.

Concern about the scarcity of human resources for mental health in low-income and middle-income countries is accentuated by reports of large-scale migration of mental health professionals to countries with higher incomes.31,32 In low-income and middle-income countries, the general loss of health professionals is especially disruptive for mental health systems, because they tend to be underdeveloped. A review of education and training of mental health professionals has shown that, even in the absence of migratory depletion, training facilities in low-income and middle-income countries are grossly inadequate to make up for the scarcity of professionals.

Financial resources

Almost a third of countries (31%) do not have a specified public budget for mental health. Of the 101 countries that have a designated mental health budget, 21 (with more than 1 billion people), spend less than 1% of their total health budget on mental health (figure 3). In Africa and southeast Asia, most countries spend less than 1% of their small health budgets on mental health services. Figure 4 shows the global burden of neuropsychiatric disease, and figure 5 compares gross domestic product (GDP) per head with the proportion of the total health budget allocated to mental health. The logarithmic trendline shows that mental health in low-income countries faces a double disadvantage: the poorest countries spend the smallest proportion of their already scarce resources on mental health.

The table compares the relative burden of mental disorders with the relative budget assigned to mental health, and shows that the proportionate burden of mental disorders in low-income and middle-income countries is smaller than in high-income countries (mainly due to a larger burden of infectious diseases). However, comparatively, the budget for mental health in middle-income countries is even smaller. These gaps between burden and budget seem to merit action, by use of effective and affordable interventions; however, decisionmakers need to consider the relative cost-effectiveness of alternative uses of available resources.

The way in which available financial resources are used is crucial for provision of effective care to as many people as possible. Prepayment financing mechanisms—such as social insurance, voluntary health insurance, and tax-based arrangements—can pool risks; can redistribute benefits to people with the greatest need; and can be made progressive, so that poor individuals pay less for equivalent health care.

Figure 3: Proportion of specified budget allocated for mental health out of total health budget in each country
Redrawn from WHO Mental Health Atlas, with permission of WHO.
than rich people. But although out-of-pocket-payments can target need neither as effectively nor as equitably as these alternative systems, they are still widely used. Worldwide, the most common method of financing mental health care is taxation (60%), followed by social insurance (19%), out-of-pocket payments (16%), external grants (3%), and voluntary insurance (2%). But more than a third of low-income countries rely on out-of-pocket payments as a primary source of finance for mental health care, compared with only 3% of high-income countries. One reason for such reliance is that low-income and some middle-income countries do not have the infrastructure for introduction of prepayment mechanisms. Tax-based financing (eg, income tax) can be ineffective if employment is largely informal or if tax compliance and collection are poor. But indirect taxes (eg, sales tax) fall disproportionately on those low-income groups in which mental disorders are most prevalent. Problems with generation of sufficient tax revenues have prompted suggestions for alternative methods of financing. For example, social health insurance (SHI), has now been introduced in many East European and former Soviet states. In SHI systems, salary-based contributions by workers to so-called sickness funds are administered and managed by public or quasipublic bodies to provide cover for unemployed, retired, and other disadvantaged or vulnerable people. Governments usually make transfers to such funds from general taxation, and employers can also contribute. Payments are usually progressive, so that higher earners pay more, but not adjusted for health risk. However, the benefits of SHI are generally restricted to those who contribute. The high rates of unemployment and disrupted working patterns for people with serious mental illness mean that many people with common mental disorders cannot access SHI. In many countries (eg, in parts of South America) SHI cover is only available to urban populations; in Mozambique, only civil servants are covered. In some East European

![Figure 4: Percentage of neuropsychiatric DALYs out of total DALYs (2002)](Reproduced from WHO, with permission of WHO.)

![Figure 5: Association between specified budget for mental health as a proportion of total health budget and gross domestic product per capita](Logarithmic trendline, y=1.1041ln(x)−4.9884, R²=0.2507. Reproduced from WHO Mental Health Atlas, with permission of WHO.)
countries, the revenue generated by those in employment is insufficient to provide health-care cover for the eligible population, and heavy subsidies are needed from tax revenues. More generally, linkage of health-care entitlements with employment could constrain job mobility and hence reduce the competitiveness of the national economy.

Voluntary, or private, health insurance is not available in most low-income and middle-income countries, although use by wealthier individuals in some countries is growing. More generally, linkage of health-care entitlements with employment could constrain job mobility and hence reduce the competitiveness of the national economy.

Out-of-pocket user payments provide an immediate, flexible, and low-maintenance source of revenue, but do not protect individuals against disproportionately high costs, or distribute benefits towards those with greater needs. User charges disadvantage the poor, and are open to corruption. Furthermore, people who are already reluctant to seek help for a mental health problem (eg, because of stigma) might be forced by the high cost of out-of-pocket payments to delay treatment until their needs are acute and the necessary care is even more expensive. In India, the risk of out-of-pocket payments exceeding 10% of household income was much higher for women than for men, yet these groups have low access to appropriate services. Within communities, disadvantaged populations such as homeless people and refugees tend to have high rates of mental disorder; as do the indigenous populations of countries with colonial histories, even when these countries have a high average income per head.

**Socioeconomic status**

Poverty is linked to poor health status. Poverty is more than low-income or low consumption; it encompasses non-monetary aspects such as social exclusion, social vulnerability, and denial of opportunities and choice. For example the UN Development Programme has designed the Human Development Index as a comparative aggregate measure of life expectancy, literacy, education, and standard of living, and the World Bank has identified increased opportunity, empowerment, and security as means to overcome poverty. Children born into poverty face various risk factors for mental and physical illness. Risk factors in poor children’s families and communities combine with scarcity of protective factors to increases the likelihood of mental health problems and developmental disabilities. Relative poverty and inequality within communities are associated with increased risk of mental health problems.

Poverty and its associated psychosocial stressors, such as violence, unemployment, and insecurity, are correlated with the onset of adult mental disorder. Epidemiological data from five studies in low-income and middle-income countries showed that people with low education and low income were most vulnerable to common mental disorders, irrespective of the society in which they lived, and that relative poverty was a risk factor for common mental disorders. However, not all studies have reported a link between poverty and mental illness.

Another review, of 11 community studies in six low-income and middle-income countries in Africa, Asia, and Latin America, reported a consistent association between poor education and high rates of mental disorders. This finding was replicated in Chile and Columbia. The investigators suggested that poor education could be a marker for childhood adversity, which increases the risk of mental illness.

Sex is also an important determinant of mental disorders, help-seeking, and the need for services. In many countries, more women than men meet criteria for common mental disorders such as anxiety and depression. In Chile, Araya and colleagues showed that women, and especially those with little education and in low social classes, had high rates of common mental disorders. Harpham and colleagues recorded the same link in Columbia. Patel and co-workers showed that nearly half of people who attended primary care in India had common mental disorders, and that such disorders were associated with poverty and female sex, after controlling for other social and demographic variables.

**Inequities in access to mental health care**

Not only are resources for mental health scarce, but they are also distributed inequitably: between countries, between regions, and within local communities. Need and access tend to vary inversely—those with highest need have least access to care. The rate of mental disorders and the need for care are highest in poor people, those who are least educated, women, young people, and rural communities; yet these groups have low access to appropriate services. Within communities,
Panel 2: Resources for child and adolescent mental health care

About half of all lifetime mental disorders begin before the age of 14 years. Worldwide prevalence rates for child and adolescent mental disorders are around 20%, and similar types of disorders are reported in different cultures. The WHO Atlas on Child and Adolescent Mental Health Resources gathered information about child and adolescent mental health resources from 66 countries. The results showed that the regions in the world with the highest percentage of the population under the age of 19 years were those with the lowest level of resources. Specific child and adolescent mental health policies were generally absent, despite the finding that mental health programmes will not be adequately developed without relevant policies. A designated institution or a governmental entity with overall responsibility for child and adolescent mental health services could only be identified in less than a third of all countries. Acknowledgment of the UN Convention on the Rights of the Child (often seen as a corollary of child mental health policy) far exceeded its use in policy or programme development. Of 66 countries identified a national policy incorporating child rights, most often with a specific focus on abuse, rather than more general child mental health needs. This gap is important, since advocacy for child mental health services is complex and often an adversarial process.

No countries had adequate numbers of providers, trained to implement effective treatments. Most low-income and middle-income countries had one child psychiatrist for every 1–4 million people. Other relevant deficiencies were absence of standards for training; failure to use available potential resources; and inability to implement supplemental training for those in contact with children who might need care. Standards for training were non-existent in many regions and lacked enforcement in many others. Only 10 of 66 countries reported that more than 25% of their paediatricians had mental health training, although paediatricians were identified as providers of mental health care in 37 of 66 countries. Despite obvious need, countries failed to identify the training of primary health-care professionals as a resource for child mental health services. Less than 10% of child and adolescent mental health services were provided by primary-care clinicians. Retraining or supplemental training of adult psychiatrists has also lagged in many countries.

Rural populations also have inadequate access to care, since mental health professionals in most low-income and middle-income countries tend to live in and around the largest cities. Of 20 countries that assessed their mental health systems with the WHO Assessment Instrument for Mental Health Systems (AIMS) method, 12 reported that rural populations were under-represented among users of outpatient services. Similarly, six of 13 countries reported that ethnic and religious minorities were under-represented in the use of outpatient services. The main reason for this barrier to access was that services did not use strategies to deliver care equitably to all groups.

Stigma and discrimination

Though under-provision of resources remains the most important barrier to effective mental health care, even in the highest-income countries, most people with mental disorders receive no effective care; for example, in the USA, two-thirds of people with mental disorders received no treatment (and paradoxically half who did receive treatment did not meet diagnostic criteria for a mental disorder). Use of mental health care is therefore constrained by demand as well as supply. Stigma and discrimination are important factors in the reluctance of many people worldwide to seek help, or even to accept that their difficulties relate to mental illness. In Ethiopia, for example, 75% of relatives of people with diagnoses of schizophrenia or mood disorders said that they had experienced stigma because of the presence of mental illness in the family, and 37% wanted to conceal the fact that a relative was ill.

A survey in South Africa reported a general public perception that mental illnesses were related to either stress or insufficient willpower, rather than medical causes. People therefore believed that such problems could be dealt with by discussion, rather than consultation with health professionals. Similar views have also been reported in countries such as Turkey, Siberia, and Mongolia. Such attributions are associated with blame and rejection, as opposed to sympathetic or helpful responses to people with mental illness. In China, a large-scale survey reported that more than half the family members of people with schizophrenia said that the effect of stigma on them and their family was such that they had decided to conceal the mental illness in their family. Stigma was greatest in urban areas and for people with the most education. In India, relatives of people with schizophrenia were sufficiently concerned about the effects of stigma on marital prospects and the possibility of rejection by the community that they hid the condition from others. Indian women with mental disorders reported the highest levels of stigma, in addition to that associated with separation or divorce, and were especially disadvantaged since they often received no financial support from their former husbands. In India, psychiatrists are the least preferred option for people seeking help for mental illness.

Research on help-seeking by young people has not fully explained the very low rates of consultation by those who are mentally ill. Potential explanations for avoidance of health care by young people include low levels of so-called mental health literacy (ie, the ability to correctly identify mental illness in oneself or one’s peers) and negative emotional responses or attitudes to people with mental illness (ie, stigma). In many countries, young people are ignorant about mental illness. Young people who seek and receive mental health care also face barriers to care. Compared with adults, young people have less favourable attitudes towards people with mental illness, and young people with mental illness might be exposed to more stigma than adults. Because young people are often embarrassed about mental illness and believe that it should be handled privately, they tend to seek help less often. Stigma is therefore a barrier to help-seeking by young people for mental illnesses.

In many countries a sense of shame contributes to inhibition about seeking help for mental disorders. Cultural differences in stigmatisation and help-seeking...
include a reliance on religious authority figures in Muslim countries,\textsuperscript{76,81,114–120} So-called structural discrimination, in which people with mental illness are not considered to have the same value as people who do not have mental illness,\textsuperscript{121,122} is exacerbated by popular misunderstandings of mental illness, which affect people’s ability to seek help and disclose their problems.\textsuperscript{123} Experiences of shame about self and blame from others have been widely reported.\textsuperscript{124} Mental illnesses are more stigmatised than physical disorders,\textsuperscript{125,126} and indeed have been referred to as the “ultimate stigma”.\textsuperscript{127} Rejection and avoidance of people with mental illness seem to be universal phenomena.\textsuperscript{13}

Stigma leads to avoidance and under-use of mental health care, and exacerbates inequity, since individuals in greatest need of help, such as mentally ill people who are homeless, experience the most stigma, and are hence effectively excluded from care.\textsuperscript{128} National public education campaigns and local interventions based on direct social contact with people with mental illness might be effective for reduction of stigma.\textsuperscript{129,130}

Human rights

Many people with mental disorders experience outright abuses of their human rights, and sometimes even within treatment facilities. This type of inequity is much less common in the treatment of other medical conditions. Most countries routinely report violations of human rights of psychiatric patients,\textsuperscript{131} and few countries have legislation that adequately protects the rights of people with mental disorders. WHO data showed that in a substantial number of low-income and middle-income countries, patients in mental hospitals were physically restrained or secluded for long durations.\textsuperscript{69} The abuses associated with involuntary detention vary from country to country in both the private and public sectors. A study in four Central American countries\textsuperscript{132} showed that mental health workers employed in the private sector had a greater awareness of patient rights than those employed in the public sector, and that this difference was greater than the difference in awareness between countries.

In accordance with the objectives of the UN Charter and other relevant international agreements, human rights are the fundamental basis for mental health legislation, together with equality and non-discrimination, the rights to privacy and individual autonomy, freedom from inhuman and degrading treatment, the principle of least restrictive environment, and the right to information and participation.\textsuperscript{133} Jones\textsuperscript{134} argues that existing international human rights laws can be applied to the human rights abuses experienced by people with mental disorders. Such laws could be used as crude but useful instruments to apply pressure to protect people with mental disorders from abuses of their human rights.

Inefficiencies in use of resources

Inefficiencies characterise all health systems, but are perhaps most visible when resources are already scarce and inequitably distributed. If mental health systems are analysed from the perspective of allocative efficiency (ie, whether the distribution of resources best meets a society’s needs) the extent of untreated psychiatric morbidity and the high associated burden of disability suggest that resources for mental health are not distributed efficiently. Most low-income and middle-income countries give low priority to mental health policies despite evidence that mental disorders cause a high and growing disability burden and long-term effects on quality of life, and that treatments for mental disorders are relatively cost-effective, compared with those for other conditions.\textsuperscript{135} Governments still need to be persuaded to allocate much larger proportions of public resources to mental health, and not to rely on the inevitably patchy, uncertain and time-limited initiatives of donor countries and NGOs. Although, to prioritise mental health, governments would need to make difficult trade-offs between investments in different public policies, mental health policy has probably been unfairly disadvantaged by the endemic stigma attached to mental illness.

Even those governments that are committed to improvement of mental health are hindered by the scarcity of research about the link between resources expended and outcomes achieved. Most studies on the effectiveness or cost-effectiveness of mental health interventions have been in North America, western Europe or Australia, with very few in low-income or middle-income countries.\textsuperscript{136–139} Even if evidence for the effectiveness of interventions could be generalised from one country to another—which has been contested\textsuperscript{140}—cost-effectiveness evidence travels especially badly between countries. Cost-effectiveness can be affected by differences in health systems and other relevant systems such as education, housing, criminal justice, and income support, differences in financing arrangements and incentive structures, and differences in relative price levels.\textsuperscript{141} What works and what seems cost effective in the USA, for example, might not be cost effective in a low-income country where, in relative terms, salaries are low and medications very expensive.

Lack of evidence is not the only reason for a poor return in terms of health improvements and quality-of-life gains from the resources actually committed to treatment and support for people with mental illnesses (ie, poor technical efficiency). For example, many middle-income countries that have made substantial investments in large asylums are reluctant to replace them with community-based interventions and inpatient facilities in general hospitals. However, well planned community-care arrangements that have sufficient resources to ensure appropriate accommodation and adequate support staff are more cost effective than asylums for people with long-term mental health problems and associated needs.\textsuperscript{142,143} Data from the WHO Atlas showed that two-thirds of all mental health beds were still in specialist mental hospitals,\textsuperscript{134} and WHO-AIMS data\textsuperscript{14} showed that 18 of 19 low-income and middle-income
countries spent more than half of their mental health budgets on mental hospitals. One barrier to a transition to community-based care is the need for “double funding” as hospitals run down and community-based systems develop. Such investment is likely to pay for itself in the longer term but can appear unaffordable in the short-term. Another barrier might be resistance from high-level staff at mental hospitals, who might protect their vested interests by attempts to maintain the institutions that provide them with prestige, power, and financial control.139

Some evidence-based pharmacological treatments are unaffordable in most low-income and middle-income countries. For example, newer generations of medications (such as the atypical antipsychotics and the newer forms of antidepressant) that are only marginally more effective than other drugs are unlikely to be cost-effective (according to the criteria suggested by the Commission on Macroeconomics and Health)140 because of their substantially higher prices.17,141 Irrespective of whether they are cost-effective, effective psychosocial interventions132 are feasible only if there are sufficient staff to deliver them. Shortages of appropriately trained staff in many countries make it impossible to implement many of the evidence-based interventions used in high-income countries. WHO has argued that the most efficient interventions for common mental disorders are as cost-effective as interventions for other chronic, non-communicable conditions, and are very affordable (on the grounds that each healthy year of life gained costs less than a year of average income per head).7 Community-based interventions that use older drugs for severe mental disorders are also relatively affordable (on the grounds that each healthy year of life gained costs less than three times the average income per head).7

Arrangements for financing of mental health services often create inefficiencies. As described earlier, risk-pooling arrangements can be infeasible or very difficult to implement: tax-based financing can be ruled out by informal employment and low tax compliance; social health insurance can exclude people with chronic mental health problems who have disrupted work patterns; and voluntary health insurance is not affordable for most people in low-income and middle-income countries. Therefore, many countries rely on out-of-pocket payments, which are not only grossly inequitable but also create perverse incentives. For many people, mental health treatments are unaffordable; studies have shown that introduction of out-of-pocket payments or increases in costs reduce use of mental health services.8 Moreover, because the individuals who pay for these services generally have little understanding about the effectiveness of available treatments, they are vulnerable to misinformation from unscrupulous providers.

Many mental health problems affect not only an individual’s health but also their family relations, their employment, their needs for income and accommodation, and their behaviour (eg, illnesses that cause withdrawn or disruptive behaviour). Even if financial risks can be pooled, the breadth of these effects can create other barriers to efficiency, since—in principle—they should be addressed by many systems, such as social services, income maintenance, housing, and criminal justice. Coordination of action across so many areas of public policy is difficult even in those high-income countries that have recognised the challenge. Professional rivalry, budget protection, and narrowly defined performance criteria can inhibit system-wide cost-effectiveness. Paradoxically perhaps, this “silo budgeting” problem becomes more acute as services are shifted from hospital to community-based arrangements and as countries recognise the need to address non-health needs as well as health needs of people who use mental health services.

**Implications for policy and practice**

Scarcity of resources for mental health, inequity in access to them, and inefficiencies in their use have serious consequences, the most direct of which is that people who need care get none. The treatment gap—the proportion of those who need but do not receive care—is too high for some mental disorders. As many as one in three individuals with schizophrenia and other non-affective psychoses do not receive any treatment.142 The treatment gaps for depression and dysthymia, bipolar illness, panic, generalised anxiety, and obsessive compulsive disorders are all greater than 50%. The challenge is greatest in developing regions of the world: WHO has reported that the treatment gap for serious disorders is 35–50% for developed countries and 76–85% for low-income and middle-income countries.143 Even for those who receive some treatment, the proportion who receive effective and humane treatment is small. We cannot escape the conclusion that most of those who need care for mental disorders do not receive effective care. The consequences include an enormous amount of disability, human suffering, and economic loss.

In the past few years policymakers have become more aware of the importance of mental health, especially in countries with low and middle-incomes. This increased awareness has been demonstrated by the publication of World Health Report 2001,1 a call for action from Ministers of Health144 and the passing of a strongly worded resolution by the World Health Assembly145 and by European Ministers of Health.146 However, increased awareness has not yet been translated into greater investment of resources. We suggest that the issue of resources for mental health is complex: severe scarcity of resources is further compounded by inequity in their distribution and inefficiencies in their use. These obstacles need to be surmounted before real gains can be made for mental health care—worldwide and in low-income and middle-income countries. Innovative, concerted, and sustained efforts are needed to remove these obstacles and achieve better mental health.
Contributors
The views expressed are those of the authors, and do not represent the views of the organisations they work for. All authors have seen and approved the final version.

Conflict of interest statement
GT received funding for research on stigma from Lundbeck UK, for whom she has sat on advisory boards. She has also sat on advisory boards for the United States National Institute of Mental Health. She declares no other conflicts of interest.

Acknowledgments
We thank Marco Garrido Cumbirera, Tarun Dua, Robert Pratt, Pratap Sharan, and Rosemary Westermeyer for assistance in collection of material and organization of the manuscript. We thank other authors of the Lancet Series on Global Mental Health who provided comments on this paper in the draft stage. SS is an employee of WHO; the views expressed in this article do not necessarily represent the decisions, policy, or views of WHO.

References


Human papillomavirus and cervical cancer

Mark Schiffman, Philip E Castle, Jose Jeronimo, Ana C Rodriguez, Sholom Wacholder

Summary

Cervical cancer is the second most common cancer in women worldwide, and knowledge regarding its cause and pathogenesis is expanding rapidly. Persistent infection with one of about 15 genotypes of carcinogenic human papillomavirus (HPV) causes almost all cases. There are four major steps in cervical cancer development: infection of metaplastic epithelium at the cervical transformation zone, viral persistence, progression of persistently infected epithelium to cervical precancer, and invasion through the basement membrane of the epithelium. Infection is extremely common in young women in their first decade of sexual activity. Persistent infections and precancer are established, typically within 5–10 years, from less than 10% of new infections. Invasive cancer arises over many years, even decades, in a minority of women with precancer, with a peak or plateau in risk at about 35–55 years of age. Each genotype of HPV acts as an independent infection, with differing carcinogenic risks linked to evolutionary species. Our understanding has led to improved prevention and clinical management strategies, including improved screening tests and vaccines. The new HPV-oriented model of cervical carcinogenesis should gradually replace older morphological models based only on cytology and histology. If applied wisely, HPV-related technology can minimise the incidence of cervical cancer, and the morbidity and mortality it causes, even in low-resource settings.

Burden of cervical cancer

There were about 500,000 incident cases of and 275,000 deaths due to cervical cancer worldwide in 2002, equivalent to about a tenth of all deaths in women due to cancer. The burden of cervical cancer is disproportionately high (>80%) in the developing world. Not only is cervical cancer the most prevalent and important cancer in women in several developing countries, but also the societal importance of the disease is accentuated even further by the young average age at death, often when women are still raising families. Cases are often detected at late stages due to non-existent or inadequate screening, and the standard treatment options are often absent or unaffordable. Promising approaches to cervical cancer prevention have resulted from our new understanding that almost all cases are caused by persistent infection with about 15 genotypes of human papillomavirus (HPV). We review recent advances and current issues regarding HPV and cervical cancer.

The cervical transformation zone

Cervical cancer usually arises from a ring of mucosa called the cervical transformation zone (figure 1). For reasons that we do not understand, persistent HPV infections cause cancers mainly at the transformation zones between different kinds of epithelium (eg, cervix, anus, and oropharynx). Illustrating the importance of the transformation zone, cancer-associated (carcinogenic) HPV infections are equally common in cervical and vaginal specimens; however, cervical cancer is the second most common tumour in women worldwide, whereas vaginal cancer is exceedingly rare. The position of the cancer-susceptible transformation zone is dynamic, gradually shifting over years towards, and into, the endocervical canal as stratified squamous epithelium replaces the mucus-producing glandular epithelium.

Prevention of cervical cancer after abnormal screening results depends on the destruction or excision of the...
entire transformation zone epithelium, not specific precancerous lesions; this method is effective in about 80–95% of cases.8,9 The site of a biopsy showing cervical precancer is not necessarily the exact site of subsequent cancer development but rather is evidence of a field of increased risk. Therefore, exfoliative cytological and virological measurements of the transformation zone can sometimes predict cancer risk even when histopathology from a colposcopically derived biopsy does not confirm the presence of a precancer.20

**Histopathology**

In poorly screened populations, squamous cell carcinomas constitute most cases of cervical cancer. In regions with good cervical cancer screening programmes, the proportion of adenocarcinomas is increased (15–20%) compared with unscreened populations, presumably because they arise from the poorly sampled glands of the canal or from poorly recognised precursor lesions.11 Beyond the relative increase, absolute rates of cervical adenocarcinomas are thought to have increased in various countries over the past two to three decades,12,13 for uncertain reasons. Infection with a carcinogenic HPV is a necessary cause of both squamous cell carcinoma and adenocarcinoma. However, the distribution of carcinogenic HPV types and variants detected in these two histopathological types (eg, adenocarcinoma is more strongly linked with HPV18) and the roles of non-viral cofactors (eg, smoking and parity) differ.14,15

**Basics of HPV virology**

Papilloma (wart) viruses have co-evolved with animal hosts over millions of years and the life cycle of each genotype of HPV is tied closely to the differentiation of its specific epithelial target (eg, sole of foot, non-genital skin, anogenital skin, anogenital/oropharyngeal mucosa).16 The relations between HPV genotypes can be expressed in the form of phylogenetic trees based on DNA sequence and protein homologies, which serve as unifying tools in understanding HPV classification and behaviour.17 HPV16 and HPV18 are the two most carcinogenic HPV types, and are responsible for 70% of cervical cancer and about 50% of cervical intraepithelial neoplasia (CIN) grade 3 (CIN3);18 by contrast, HPV6 and HPV11 are responsible for about 90% of genital warts.19 When we refer to HPV infection in this Seminar, we are referring to the genetically related group of genotypes that are linked to cancer risk—ie, the carcinogenic types—unless specified. For cytopathology, we refer to the 2001 Bethesda System19 and for histopathology, we use the WHO classification.20

The human papillomavirus genome codes for only eight genes (figure 2).21 E6 and E7 are the primary HPV oncoproteins. Each has numerous cellular targets,22,23 with E5 and retinoblastoma tumour suppression protein (pRB) being the most important. E6 inhibition of p53 blocks apoptosis, whereas E7 inhibition of pRB abrogates cell-cycle arrest. E7 is the primary transforming protein. Both proteins are expressed at low levels during the infectious process. At some still undefined point in progression to

[Figure 2: The HPV genome and its expression within the epithelium](#)
precancer, E6 and E7 expression is deregulated by either genetic or epigenetic changes, leading to their overexpression in the full-thickness epithelial lesion.

Development of cervical cancer

Cervical cancer arises via a series of four steps—HPV transmission, viral persistence, progression of a clone of persistently infected cells to precancer, and invasion—that can be reproducibly distinguished and which provide a rational starting point for any discussion of optimum prevention efforts (figure 3). Backward steps occur also, namely clearance of HPV infection and the less frequent regression of precancer to normality. The molecular virology underlying HPV persistence, progression, and invasion is not well understood, but this causal model is supported by epidemiological and laboratory data and does not require unreliable morphological distinctions like histological CIN grade 1 (CINI) or cytological atypical squamous cells of undetermined significance (ASC-US) analogous to borderline dyskaryosis.25,26

HPV transmission

Anogenital HPV infections are transmitted mainly by skin-to-skin or mucosa-to-mucosa contact.27,28 The probability of infection per sexual act is not known but is clearly high,27 with no known difference between HPV types. Because of their common transmission route, HPV types tend to be transmitted together,29,30 resulting in a high proportion (20–30%) of concurrent infections with several different types when women in the general population are sampled.31 Men are also often infected with several HPV types concurrently, implying that a sexual act could transmit several types at once.

Independent of type, infecting viral particles reach the germinial cells in the basal layer presumably via tiny tears to the mucosa.4 Male circumcision might decrease male HPV infection and carriage, possibly due to the toughness of keratinised epithelium, thereby reducing transmission.32 Penetrative sexual intercourse is not strictly necessary for transmission and HPV types can apparently be transferred to the cervix from original infection at the introitus.33

Most women in the world are probably infected with at least one if not several types of HPV during their sexual life.4 Total exposure is difficult to measure because DNA detection is usually transient and serology is not accurate.35 Thus, a substantial proportion of HPV DNA negative, seronegative women have nonetheless been exposed.

While looking for uncommon, significant cervical lesions, pathologists and clinicians encounter a huge assortment of abnormalities that are minor or, even more commonly, equivocal (figure 4). Many millions of women are diagnosed every year with such abnormalities.38 However, these abnormalities cannot be ignored because most precancers and cancers are diagnosed in women with equivocal or mildly abnormal cytological findings.39

Only about a third of women with HPV infections detectable by DNA testing have recognised cytopathology.40
Cytological abnormalities are less sensitive for detection of HPV than molecular testing. HPV16 and related types are most likely to produce high-grade squamous intraepithelial lesions; by contrast HPV18 (the second most common type in cancers) causes a disproportionately low fraction of such lesions. A lack of HPV18-induced high-grade squamous intraepithelial lesions could explain, at least in part, the poor performance of screening for endocervical or glandular lesions and the increased proportion of adenocarcinoma, which are associated with HPV18, in well-screened populations.

In longitudinal studies of cytologically normal adult women who are HPV DNA positive at enrolment, the cumulative risk of incident equivocal and minor cytological abnormalities rises to a high level (about 25–50% of smears) 1–2 years after enrolment and declines thereafter, returning to baseline (<5% of smears) at about 4 years. The smaller cumulative risk of precancer and cancer continues to rise for as long as we have been able to observe prospectively (≥15 years), suggesting that a few women remain persistently infected. How often precancer arises from an evident mild lesion versus an equivocal lesion or cytologically normal, HPV-infected tissue is not known.

**HPV clearance versus persistence**

Most cervical HPV infections (with cytological abnormality or not) are cleared or suppressed by cell-mediated immunity within 1–2 years of exposure (figure 5). The most persistent HPV types tend to be the most common. This correspondence is to be expected because prevalence equals incidence multiplied by duration (ie, persistence). The prevalence of different types of HPV is modified by differential censoring due to detection and treatment, which are more common for lesions caused by HPV16 than other types.

With longer HPV persistence of a given type, the probability of subsequent clearance over a fixed interval decreases and the risk of precancer diagnosis increases. However, the average persistence of some non-carcinogenic types (eg, HPV61) can also be long. Prevalent infections detected in cross-sectional screening persist longer in older women than in younger women, probably because they are more likely to represent infections already of long duration. The median time to clearance of HPV infections detected during screening studies is 6–18 months. There is no accepted definition of clinically important persistence, but follow-up strategies targeting abnormalities lasting more than about 1 year (and especially 2 years) seem to distinguish infections and associated lesions posing greater risk to the patient from transient infections. The small proportion (about 10%) of carcinogenic infections persisting for several years is strongly linked to a high absolute risk of diagnosis of precancer.

Ongoing cohort studies with up to 10 years of follow-up data have shown that, after clearance, the same HPV type can occasionally re-appear. Whether infections resolve by complete viral clearance or by maintenance of a latent state in the basal-cell epithelium, in which the virus replicates at extremely low levels without full viral expression, is unclear. The appearance of many HPV infections among immunosuppressed HIV-positive individuals suggests that latency is a possibility. In populations with secondary peaks of HPV infection—eg, post-menopausal women—re-emergence from latency due to senescence of cell-mediated immune control could have a role, as well as new sexual partners (of the women or their partners) or cohort effects. However, older women with a long period without cytological signs of HPV infection show very small risk of subsequent cervical cancer, suggesting that re-activation from latency typically does not cause harm.

**Progression to cervical precancer**

In terms of histopathology, precancer includes the fairly reliable morphological diagnoses of CIN3, severe dysplasia or dyskaryosis, or carcinoma in situ (figure 6). In precancer, undifferentiated cells with fixed genetic abnormalities have replaced almost the full thickness of the cervical epithelium. To discuss what lesions do not represent precancer is also important for diagnostic specificity. CIN grade 2 (CIN2) is heterogeneous: it is sometimes produced by non-carcinogenic types of HPV and, therefore, is equivocal in cancer potential. CIN1 is an insensitive histopathological sign of HPV infection, and is not precancer. Careful study of cases of...
Figure 6: What defines precancer?
Heterogeneity in biology (and definition) still exists in precancer, even as it remains the prime target of screening programmes and preventive treatment as well as the scientific surrogate for cancer risk. As the functional definition of precancer has expanded to include smaller and less serious lesions due to difficulties distinguishing the true cancer precursors, the risk (predictive value) of the diagnosis of precancer as a surrogate for predicting invasive cancer has declined. These changes can alter the effect and assessment of prevention programmes. The most certain surrogate for invasive cancer is full-thickness carcinomas in situ; however, many CIN3 lesions detected by screening are very small and less certain to pose an eventual risk of invasion. Nonetheless, CIN3 shares the true cancer precursors, the risk (predictive value) of the diagnosis of precancer as a surrogate for predicting progression to precancer given viral persistence.17,45

HPV16 is remarkably carcinogenic, with an absolute risk of a precancer diagnosis approaching 40% after 3–5 years of persistent infection.17,46,47 The total risk of precancer for a woman carrying several HPV types is increased compared with women infected with any one of the HPV types she carries, but it is not clear whether her risk is greater than the sum of the risks posed by individual HPV types.11

Viral load measurements are not clinically useful. Levels detectable only by PCR (eg, below the threshold of detection of the commercially available Hybrid Capture 2 [Digene Corporation, Gaithersburg, MD, USA]) are associated with microscopic normality and with low risk of subsequent precancer or cancer, but increasingly high viral loads do not imply increasing prospective risk,14 except for HPV16.31,42 The amount of HPV DNA measured in scrapes of the cervical epithelium is a complex sum of the number, size, and grade of the HPV-associated lesions,53 and therefore the meaning of viral load is ambiguous. Some of the highest viral loads can be attributed to recently acquired minor lesions producing large amounts of virus, analogous to benign warts.

Risk of cervical cancer is mainly a function of HPV infection and lack of effective screening. External factors (apart from screening) are minor compared with the extremely high primary risks of the most carcinogenic HPV types. Smoking,6 smoking,64 multiparity,65 and long-term use of oral contraceptives66 can double or triple the risk of precancer and cancer among women infected with carcinogenic types of HPV. The role of chronic inflammation, especially due to coinfection with Chlamydia trachomatis, is less certain.67 Further, there has been no confirmation of a role for any one micronutrient in observational studies and supplementation trials,68 although there is some evidence of a possible protective association between higher folate and the risk of precancer.69 Among HPV-infected women, low socioeconomic status might remain a risk factor for precancer even when recent medical care is taken into account.69 Interestingly, a preliminary association between condom use and decreased persistence or progression has been seen in a few studies.70,71 The mechanisms of action for HPV cofactors (whether immune, genotoxic, or hormonal) are not well understood.

Poorly understood cellular immune responses strongly affect whether an infection is ultimately cleared or persists to pose a risk of precancer. Efforts to identify specific subsets of T cells responsible for clearance remain inconclusive.5 Epidemiological studies in diverse populations have shown the human leucocyte antigen DRB1*1301 to be protective.72 In addition to acquired immune responses, innate immune responses—the first line of mucosal defence—could also have an important role.73

HPV infection in people living with HIV/AIDS has been addressed in detail elsewhere.12,74,75 In brief, HIV
status, defined by CD4+ T-cell counts and HIV viral load, strongly affects the early natural history of HPV. HIV-positive individuals experience increased HPV prevalence and persistence, and decreased viral clearance compared with HIV-negative individuals for types other than the uniquely persistent and carcinogenic type HPV16. HIV-infected individuals are often infected with more HPV types than are non-infected individuals, suggesting type differences in the success of immune suppression. The probability of invasion is not strongly affected by HIV. Highly active antiretroviral therapy (HAART) does not seem to affect HPV natural history or to reduce the risk of cervical precancer and cancer; however, the relation between such therapy and risk is perhaps confounded because HIV-infected women on HAART are living longer and allowing the cervix to have prolonged exposure to carcinogenic HPV in the context of relative immunosuppression.

Invasive cervical cancer
In unscreened populations, the peak risk of invasive cervical cancer occurs earlier than for most adult cancers, peaking or reaching a plateau from about 35 to 55 years of age. This distribution is due to the fact that cervical cancers originate mainly from HPV infections transmitted sexually in late adolescence and early adulthood. The average time between HPV infection and establishment of a (small) precancer seems to be much shorter than the average duration of precancer growth leading to invasion. There are many more precancers than cancers, suggesting that only a minority invade. The precise magnitude and timing of risk of invasion, if precancers were left untreated, will remain unknown because contemporary cohort studies, in which treatment of precancer is mandated, cannot study invasion ethically. Crude estimates from early studies of large precancers suggested a 20–30% risk of invasion over a 5–10-year time frame.

Apart from age, risk factors for invasion are unknown except for viral type; in particular, HPV16, HPV18, and HPV45 are found in a higher fraction of cancers than in precancers than are other HPV types. The integration of the HPV genome into the host genome is associated with invasive cancer and might be an important biomarker distinguishing HPV infection from precancer. However, integration might not be necessary to cause invasion, since not all women with invasive cancers have measurable integration. Continued transcriptional activity of the HPV oncoproteins is needed to maintain the cancer.

Prevention of cervical cancer
Risk as a guiding principle of prevention strategies
The steps in cervical cancer pathogenesis can guide prevention and management. Short-term risk of CIN3 is a scientifically valid, ethically justified surrogate for long-term cancer risk, and can be estimated in prospective studies and clinical trials. To base clinical decisions on knowledge of risk of such lesions makes sense; the clinical response should be uniform irrespective of what clinical test is used to define risk (panel 1). For example, finding HPV16 on an HPV DNA test conveys slightly higher risk of subsequent CIN3 than does cytological identification of low-grade squamous intraepithelial lesions. The best way to predict individual risk is to use the risk estimate from a large stable group of women with similar characteristics.

Panel 1: Replacing clinical protocols with risk stratification
Total cost and total benefit are the key statistics needed to assess and compare old and new technologies for screening and diagnosis. For public-health programmes like screening and management of very common abnormalities, the costs and benefits should be described on a population basis; for example, the number of cases of disease or death averted and the total cost—ie, financial and iatrogenic medical consequences—summed over all women receiving the test. These integrated costs and benefits, and therefore prioritisation of who will receive the tests, will vary with the indication for the screening and with geography, age, and sexual behaviour. Sensitivity and specificity without weighting by frequency of disease are not enough to capture integrated costs and benefits adequately. By contrast, positive predictive value, which is the probability of disease among women testing positive, and negative predictive value, which is a measure of the reassurance of no disease among women testing negative, are expressed in terms of population risks and are helpful.

There are several corollaries of thinking in terms of risks. First, there should be no distinction between a clinical and a molecular test. For example, the utility of a colposcopic examination, a Pap test, and an HPV test should be compared on an even footing on the basis of performance and cost. Second, for a test with fixed sensitivity and specificity, the clinical or population subset with the higher risk of disease should have higher priority for intervention, given equal cost. Third, no one-dimensional comparison of tests or programmes, like an odds ratio or even a risk difference, can capture the information needed to choose between them. Finally, the cost effectiveness of a programme is measured by comparison with alternative programmes, including no programme. Thus, visual inspection with acetic acid could be cost effective in a resource-poor area where there is no alternative programme in place, despite its low accuracy.

Adoption of a-priori thresholds based on risk for deciding who needs closer surveillance, colposcopy, and treatment will aid clinicians in making decisions that maximise the health benefit to women. For example, a society might decide that women at less than 2% absolute risk of precancer within the subsequent 2–3 years are normal and can stay in regular interval screening, women with a 2–9% risk should be re-screened in a year, women with a 10–39% risk need intensive colposcopic assessment immediately, and women with 40% or greater risk need immediate treatment. The choices of such cutoff points, once accepted, could guide management as new risk biomarkers are validated.

There are a few implications of strategies based on absolute risk. First, no screening strategy is efficient among young women, who have very high prevalence of HPV infection and of its cytological signs, both of which are very likely to clear without intervention. Strategies to prevent the rare but sometimes fatal rapidly invasive cancers among young women require screening and aggressive management of huge numbers of ultimately normal young women; prevention strategies in regions seeking a nearly perfect level of prevention can be so expensive that the entire programme is no longer cost effective. Second, using HPV testing to detect very low viral loads and marginally carcinogenic types (eg, HPV53) should be avoided to preserve the predictive value of a positive test.
Panel 2: Important questions regarding the new HPV16/18 vaccines

1. What is the duration of protection and the total effect on cancer incidence?
2. Are boosters safe and effective if needed?
3. Do the vaccines provide cross-protection against a few related types, as previously suggested?
4. What is the efficacy of fewer than three doses of vaccine, as will sometimes occur in vaccination programmes?
5. Does the vaccine prevent infection in men, and reduce the transmissibility of HPV from men to their partners?
6. When immunity wanes and incident infections occur, are the natural history of HPV16 and HPV18 infections and the related risks for precancer and cancer the same as in unvaccinated women who typically acquire infections by these types as young women?
7. What will be the effect of HPV vaccination on compliance with screening programmes, which are needed for prevention of the 30% of cancers against which the vaccines do not provide protection?
8. How great will the negative effect be of the reduced prevalence of HPV16 and HPV18 in post-vaccinated populations on the clinical performance and cost-effectiveness of screening assays and diagnostic procedures?
9. Will prevention of infection with HPV16 or HPV18 alter the natural history of other carcinogenic types and the number of cervical cancers they cause?
10. Do these vaccines protect against other HPV-related cancers such as oropharyngeal and anal cancers?
11. In developing countries, where 80% or more of cervical cancer occurs, who can afford to get vaccinated, even with tiered pricing, in view of competing health priorities?

Primary prevention of HPV infection

There is some evidence that health education programmes that promote abstinence, conscientious condom use, or both, could reduce the risk of cervical cancer at the population level.28 However, mutual abstinence until marriage is far from universal, and even strict condom use is not completely protective against HPV transmission because the male anogenital skin is not completely covered.29 Thus, the development of HPV L1 virus-like-particle (VLP) vaccines is a potentially major advance in prevention of cervical cancer. These vaccines are based on the self-assembly of recombinant L1 protein into non-infectious capsids that contain no genetic material.30 Intramuscular injection of the vaccine induces high titres of neutralising antibody, more than 50 times the titres induced by natural infection.31 Protection at the cervix against the same types in the VLP vaccine is probably mediated by antibodies transudated into the secretions that bathe the epithelium, serum antibodies directly exuded at the site of microscopic trauma thought to be involved in transmission, or both32 (Schiller J, National Cancer Institute, Bethesda, MD, USA; personal communication).

Two VLP vaccines have been developed for primary HPV vaccination. Gardasil (Merck and Co, Bluebell, PA, USA) has gained regulatory approval in several countries. Cervarix (GlaxoSmithKline, Rixensart, Belgium) has been approved in Australia, is pending approval in the European Union, and applications for approval have been submitted to regulatory agencies in the USA and other countries.32 Both vaccines target HPV16 and HPV18; Gardasil, which includes a standard alum adjuvant, also targets HPV6 and HPV11.33 Cervarix uses a new proprietary adjuvant intended to boost immunogenicity. In populations of young adult women without known exposure to the target types, both vaccines have shown near perfect efficacy against HPV infection and related cytological and histological endpoints for up to 5 years.34-37

Important questions remain (panel 2). For example, the regulatory approval for Gardasil was predicated on rather short-term efficacy data for 15–26-year-old women and data showing good antibody titres after vaccination for 9–15 year olds. At present, girls aged 11–12 years are being targeted in the USA, before entry to middle school. Several jurisdictions are considering mandatory vaccination for 11–12-year-old girls, but this is controversial, especially regarding cost-effectiveness analyses (table 1).38 Ideally, to ensure that a vaccination programme will protect young women through the age of greatest risk of HPV exposure, we would know that durability will be 10–15 years or greater or that boosting will be safe and effective; waiting for certainty, however, would reduce the benefit for the cohorts of girls born between about 1995 and early 2005.

The value of universal vaccination in the upper age range, 19–26 years, is even more controversial.39-41 Women who have had several sexual partners and have already been exposed to the target types are at least partly immune and cannot be distinguished from unexposed women by DNA testing, because both groups will be negative and serology is not reliable or accurate as a biomarker of past exposure and protective immunity.41,42 As women age, they are more likely than younger women to have established monogamous relationships that reduce future risk. The vaccines do not treat existing infections or lesions,43,44 and cross-protection against other HPV types is partial45 or non-existent.46 Therefore, the current HPV vaccines are most certain to yield the greatest public-health benefit (population effectiveness) in girls at an age before most have begun sexual activity. At a certain, still-undetermined age that might vary by region, screening might be more cost effective than vaccination if a trade-off is considered.

We believe that broad recommendations for widespread vaccination in adult women should await independent, population-based effectiveness trials and cost-utility assessments.
Multivalent vaccines (which target an expanded range of types) and different approaches to produce immunising HPV proteins are being tested. The effect of expanding the number of types on vaccine price and safety are unclear, as is the acceptability of the current vaccine once available. Implementation will create the needed vaccinated cohort to permit assessment. Implementation will create the needed vaccinated cohort to permit assessment.

Population benefit of vaccination overlaid on screening
Making vaccination mandatory would increase coverage of prevention among the poorly screened, highest-risk population that would benefit most from vaccination. Further reductions in incidence and mortality (already low in the USA) would be difficult to achieve. Adding vaccination to screening might even reduce compliance with screening because women might falsely assume that they are protected.

Urgency of mandate
If we delay, cohorts of girls will miss the benefits of vaccination. It is better to proceed slowly, accumulate more data and public acceptance based on voluntary vaccinations, and to move to mandates when public-health benefit is established.

Safety
Good safety profile to date. Rare effects can not be ruled out until many more girls are vaccinated.

Known durability
Established durability of about 5 years with sustained serotitre for HPV16. Peak risk of sexual exposure lasts for more than 10 years after suggested mandatory vaccination age.

Trends in protection in years after vaccination
No evidence for decreasing efficacy over 5 years. HPV18 serotitre fall within 2–3 years of vaccination with Gardasil, which might herald a subsequent decrease in protection.

Feasibility of boosting
One unpublished, small study suggests good antibody response to booster among young adult women. General lack of large-scale evidence, and lack of well-formulated strategies regarding how and when boosting could be done. Lack of data on the safety of boosting. No serology assay widely available for monitoring if titres proved predictive.

Choice of vaccine
Gardasil is already approved in many countries. Cervarix is approved in Australia and might be approved elsewhere within a year. Weighing its relative benefits might make sense. Are the two vaccines interchangeable and compatible?

Impending development of second-generation vaccines with more types, longer durability, or lower cost
It is best to start now and replace or boost with newer vaccines when available. Especially for lower-resource regions, second-generation vaccines could reduce number of doses, need for boosting, and cost.

Ethical and family issues
There is no evidence that vaccination would promote sexual activity. Parents could opt out. Vaccination might encourage onset of sexual activity and deny the parents their right to choose.

Cost-effectiveness
Analysis assuming lifelong durability already indicates cost-effectiveness at current prices. Unknowns include durability, need for boosting, integration with screening schedules, and possibly reduced performance of screening tests. Strain on public-health resources.

Table 1: Arguments for and against mandatory HPV vaccination of girls before the average age of sexual debut

<table>
<thead>
<tr>
<th>For mandatory vaccination</th>
<th>Against mandatory vaccination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do we have enough information to promote a major public-health mandate?</td>
<td>Many vaccination programmes are implemented before long-term durability and safety data are available. Implementation will create the needed vaccinated cohort to permit assessment.</td>
</tr>
<tr>
<td>Population benefit of vaccination overlaid on screening</td>
<td>Making vaccination mandatory would increase coverage of prevention among the poorly screened, highest-risk population that would benefit most from vaccination.</td>
</tr>
<tr>
<td>Urgency of mandate</td>
<td>If we delay, cohorts of girls will miss the benefits of vaccination.</td>
</tr>
<tr>
<td>Safety</td>
<td>Good safety profile to date. Rare effects can not be ruled out until many more girls are vaccinated.</td>
</tr>
<tr>
<td>Known durability</td>
<td>Established durability of about 5 years with sustained serotitre for HPV16. Peak risk of sexual exposure lasts for more than 10 years after suggested mandatory vaccination age.</td>
</tr>
<tr>
<td>Trends in protection in years after vaccination</td>
<td>No evidence for decreasing efficacy over 5 years. HPV18 serotitre fall within 2–3 years of vaccination with Gardasil, which might herald a subsequent decrease in protection.</td>
</tr>
<tr>
<td>Feasibility of boosting</td>
<td>One unpublished, small study suggests good antibody response to booster among young adult women.</td>
</tr>
<tr>
<td>Choice of vaccine</td>
<td>Gardasil is already approved in many countries.</td>
</tr>
<tr>
<td>Impending development of second-generation vaccines with more types, longer durability, or lower cost</td>
<td>It is best to start now and replace or boost with newer vaccines when available. Especially for lower-resource regions, second-generation vaccines could reduce number of doses, need for boosting, and cost.</td>
</tr>
<tr>
<td>Ethical and family issues</td>
<td>There is no evidence that vaccination would promote sexual activity. Parents could opt out.</td>
</tr>
<tr>
<td>Cost-effectiveness</td>
<td>Analysis assuming lifelong durability already indicates cost-effectiveness at current prices. Unknowns include durability, need for boosting, integration with screening schedules, and possibly reduced performance of screening tests. Strain on public-health resources.</td>
</tr>
</tbody>
</table>

Screening
Cervical cancer prevention, as practised in high-resource regions, includes: screening; triage of equivocal results; colposcopically guided biopsy of abnormal screening results; decision whether to treat; treatment; and post-treatment follow-up (including eventual return to routine screening intervals if appropriate). Cervical cancer prevention programmes vary widely by country and could be radically improved by new technologies.

Whichever validated screening method is chosen, broad coverage and full follow-up of abnormalities are the key requirements for reducing the incidence of cervical cancer by screening. Screening in the absence of a treatment programme is unethical. The appropriate programme for a given setting (when to begin, the proper intervals between screens, when to stop) depends on affordability, differing societal demands for protection against cancer risk, and desire to prevent iatrogenic complications among women who are at low risk of cervical cancer. However, a few principles follow from the established fact that the cause of cervical cancer is persistence of sexually transmitted HPV infection of the transformation zone. Screening women within 5–10 years of first sexual intercourse, when the risk of finding benign HPV infections is very high but risk of cancer is vanishingly low, cannot be cost effective. Similarly, it is not cost effective to screen women after total hysterectomy for reasons unrelated to cancer or older women with...
Panel 3: Why HPV16 deserves individual consideration in prevention programmes

1. HPV16 causes half the cases of cervical cancer worldwide and is the major carcinogenic type in almost every country surveyed.
2. HPV16 is the most common carcinogenic type in the general population, accounting for about 20% of infections among cytologically normal women, and 20% among women with equivocal lesions, and 26% among those with mild abnormalities.
3. Although HPV16 is no more likely to cause cytological abnormalities than other carcinogenic types, it disproportionately causes changes suggesting precancer and accounts for about 45% of those severe interpretations.
4. Prospectively, HPV16 persists longer on average than any other type and persistence is highly associated with precancer (about 40% of women with persistent HPV16 are diagnosed within 5 years with precancer).
5. HPV16 is the main HPV type that causes other anogenital and oropharyngeal cancers that are not common enough to merit screening but might be prevented by an effective vaccine.

repeated negative cytology, HPV tests, or both, assuming endocervical sampling is adequate.

The development and implementation of organised and effective cytology-based cervical cancer screening—eg, Papanicolaou (Pap) tests—for detection and treatment of precancerous lesions and earlier stage, treatable cancers has led to significant decreases in the incidence and mortality of cervical cancer. The effect of cytology programmes has been best documented from ecological correlations of cervical cancer with screening activities in populations, mainly in Nordic countries. Many early screening programmes targeted the peak ages of cervical cancer risk and produced decreases in incidence confined mainly to women aged 30–70 years. When screening coverage is extended to younger and older women, rates at all ages decrease, although cost per cancer averted rises, as seen in US Surveillance and End Result (SEER) data. A fraction of the earliest cases of cervical cancer are probably rapid onset; their shorter time of development permits fewer rounds of screening that could detect precancerous lesions.

The success of well-established cytology programmes in detecting cervical precancer and treatable cancer is attributable to repeated, fairly insensitive screening of women during the slow progression from incident HPV infection to easily diagnosed precancer (5–10 years) and from precancer to cancer (typically ≥10 years). The need for repeated screening cycles makes cytology-based cervical cancer screening programmes expensive.

Further substantial, cost-effective improvements in cytology programmes could be difficult to achieve, although automated screening of liquid-based cytology might someday prove itself. In choosing between cytology techniques, there is no convincing evidence that liquid-based cytology is more accurate than conventional pap smears, especially when adjunctive HPV testing is done. Nonetheless, liquid-based cytology might reduce the proportion of inadequate smears, especially in settings where conventional smears are prone to air-drying (eg, the tropics) or where widespread cervical inflammation is a problem. Although not abnormal, an inadequate smear can similarly increase anxiety of women.

Assays for HPV have been introduced to improve the efficiency and maximise the sensitivity of cervical cancer screening. There is convincing evidence that testing for carcinogenic HPV DNA is cost effective and sensitive for detection of precancerous lesions in women with equivocal cytology, more sensitive but less specific than cytology-based methods for primary cervical cancer screening, can be added usefully to the follow-up of women post-colposcopy when precancer is not found, and can guide assessment of cure post-treatment. Most importantly, testing negative for carcinogenic HPV provides greater reassurance against cervical precancer and cancer than does cytology-based methods. The greater reproducibility of current tests for carcinogenic HPV types is an added advantage over cytology. In the USA, HPV testing to triage equivocal cytology is commonly used. HPV is also approved in primary screening in women 30 years and older, who are past the peak of self-limited infections, and in whom the positive predictive value is higher than in younger women. The International Agency for Research on Cancer has
endorsed the use of carcinogenic HPV testing alone as an option in primary cervical cancer screening.\(^\text{139}\)\(^\text{140}\)

At least four assays indicating current infection give roughly similar results when used to assay the major carcinogenic HPV types as a pool: Hybrid Capture 2; the MY09/MY11 primer set and its improvements like PGMY; the GP5+/GP6+ primer set; and SPF10/LiPA PCR-based methods.\(^\text{123}\)\(^\text{129}\)\(^\text{130}\) Only Hybrid Capture 2 is approved in the USA by the Food and Drug Administration (FDA); it does not provide individual typing information. Two PGMY-based systems have been submitted for FDA approval, one of which is a pooled test like Hybrid Capture 2, while the other provides genotyping. We believe that only completely standardised assays should be used for routine practice because standardisation is, with sensitivity, the major advantage of HPV testing.\(^\text{13}\) Without rigorous standardisation, subtleties of HPV tests can greatly affect analytical performance.\(^\text{132}\)

HPV16 is by far the most important HPV type worldwide, and its effect is reflected in all aspects of cervical cancer prevention (panel 3). If we could eliminate HPV16 infection or reliably identify and destroy all its cytopathological or colposcopic manifestations, we could prevent up to half of cervical cancer cases.\(^\text{1}\) Table 2 shows the relative importance of different HPV types. An important goal is to improve specificity of HPV testing while maintaining its clinical sensitivity. Possible advances include type-specific detection of HPV16 and HPV18, which has been shown to identify women at the greatest risk of developing precancer and cancer.\(^\text{136}\)

Detecting persistence of these most carcinogenic HPV types would be an even more specific marker of clinically important infections, theoretically, but clinical use of genotyping will require robust assays and workable clinical protocols. Other promising screening assays under development detect carcinogenic HPV E6/E7 mRNA\(^\text{137}\) and p16\(^\text{INK4a}.\(^\text{138}\)

Cervical-vaginal self-collection permits the use of molecular testing outside clinical settings with clinical sensitivity for precancer and cancer that seems as good as cervical cytology but lower than clinician-directed sensitivity for precancer and cancer that seems as good as molecular testing outside clinical settings with clinical reference standard for diagnosing precancer or even to make finer distinctions such as CIN grade 1, 2, or 3.\(^\text{3}\) However, the choice of biopsy site and the histopathological diagnosis of resultant biopsies tend to be variable and subjective.\(^\text{14}\) Clinicians rely on colposcopy to determine the presence or absence of epithelial lesions, find the area of the cervix with the highest degree of disease, and direct the biopsy for histological diagnosis. Colposcopic assessment also provides information about location and extent of disease, which is important for planning treatment.

Although the sensitivity of screening has improved considerably in the past decade, colposcopy has not advanced, given the weak correlations between visual changes and disease severity and lack of reproducibility among assessors.\(^\text{14}\) Even highly experienced assessors have false negative colposcopy rates as high as 20–40% in patients with a histological diagnosis of precancer.\(^\text{14}\)\(^\text{14}\) Two factors affect this false negative rate: first, CIN3 lesions missed by colposcopy are smaller and involve fewer quadrants of the cervix than do lesions that are detected visually;\(^\text{4}\) and second, patients with precancer related to non-HPV16 carcinogenic types are more likely to have equivocal visual lesions.\(^\text{14}\)\(^\text{14}\) Therefore, the use of colposcopy might be even more limited when HPV testing and vaccination become more widely used.

Colposcopic sensitivity increases significantly if more than one non-random biopsy is taken, irrespective of training or expertise of the assessor.\(^\text{14}\)\(^\text{14}\) More studies are needed to determine whether the additional biopsies should be taken only from apparent lesions, from areas of minor epithelial changes, or even from seemingly normal quadrants of the cervical epithelium.

**Treatment of cervical precancer and invasive cancer**

The effect of behavioural factors on the clearance of HPV or precancer is poorly understood. However, consideration of smoking is always important for reasons of public health. There is some evidence that smoking cessation promotes resolution of HPV-induced cytopathology.\(^\text{15}\) Genotoxic smoke constituents are secreted at high levels into the cervical mucus.\(^\text{16}\) Enhancement of cellular immunity is also probably involved. In any case, it makes sense to encourage women with precancerous screening abnormalities to stop smoking in the context of a broader programme for prevention of smoking-related cancer and health problems.

At present, the usual practice for treatment of precancers is to treat the entire transformation zone of women diagnosed with equivocal (CIN2) or more definite (CIN3)
Figure 7: A comparison of community pathology biopsy diagnoses to quality control pathology review diagnoses
Comparison of biopsy diagnoses made by clinical centre pathologists (community diagnosis)156 with biopsy diagnoses rendered by the expert quality control pathology. Patterned bars highlight the proportion of agreement for a given community diagnosis. Note that many biopsies diagnosed as CIN2 by the clinical centre pathologist, the threshold for excisional treatment in many countries, were not called CIN2 on expert review.

Cryotherapy with nitrous oxide, a low-cost ambulatory procedure, is almost as effective as ambulatory excisional procedures to treat small precancerous lesions. Cryotherapy is widely used in low-resource areas mainly because it can be provided without local anaesthesia or electricity. One noteworthy drawback is the typical weeks of recovery marked by discharge with some possibility of infection. Carbon dioxide is often used as an alternative gas to nitrous oxide because it is less expensive and easier to find in remote areas; however, technical improvements are needed to overcome serious deficiencies such as blockage of equipment and poor depth of tissue necrosis.

While CIN2 is a poorly reproducible diagnosis (figure 7) and the accuracy of colposcopic biopsy itself is in some doubt, we might soon be able to treat the transformation zone based on more exact virological assessment of risk. Specifically, if a carcinogenic HPV type persists for a number of years in an older woman for whom fertility issues are not important, the risk posed by LEEP might be warranted to address the possibility that a precancer is being missed by colposcopy. The use of a molecular test to guide treatment will require extremely careful study (panel 1) and we suggest it for discussion, not immediate adoption.

Cytology and HPV testing are useful to assess cure after treatment by LEEP. Women successfully treated usually test HPV negative. Those testing HPV-negative 4–6 months after LEEP have no appreciable risk of recurrent CIN2 or worse within the subsequent couple of years, although the relevant studies have lasted only around 2 years, so for how long negative tests after LEEP can be interpreted in this way is not clear. Those testing positive must be monitored more closely, although the proper immediate management must be individualised.

Any proposed outpatient non-surgical method must work very well, because excisional procedures are 90–95% effective with minimal side-effects. Current HPV vaccines do not treat existing HPV infections or precancerous lesions. A better understanding of the molecular underpinnings of HPV and cervical carcinogenesis could lead to the rational design and development of an array of targeted, lower-morbidity non-surgical treatments such as therapeutic vaccines, topical immunotherapies (eg, imiquimod and resiquimod for treatment of condyloma), and topical chemotherapies (eg, siRNA and apoptosis inducers).

There is a pressing need to educate health professionals and the public regarding the natural history of HPV as we move towards HPV-based prevention strategies. As evidenced by the recent alarmist reaction to a report that HPV is very common in the general US population, many still conflate the one-time detection of HPV DNA with high risk of cervical cancer (unpublished data). Unwarranted psychosocial damage can follow detection of HPV. Many women would probably still prefer to be notified of a mildly abnormal Pap than of an HPV infection because the connection between abnormal Paps and sexual behaviours was previously not well understood, although the two test results address the same biological processes.

We have yet to agree as a health community on the full set of messages that should accompany HPV screening. It is not clear which health professional will have the time, training, and interest to lead the education effort in different regions, especially as messages change following the advent of vaccination. A critical example of an
educational issue is what to tell women with normal cytology and a positive HPV test. We propose that patients should be informed that although HPV exposure is extremely common, almost all infections go away within a year or two; many are gone within 6 months. Patients should make sure that they get retested, and if the infection does not clear, then they will need a full examination (colposcopy with multiple biopsies) and possibly treatment to prevent precancer and later risk of cancer.

Two important advances in the treatment of invasive cancer deserve brief mention. Radical hysterectomy has been the preferred treatment for stage I cases, but during the past few years, minimally invasive surgery has become an option for young women with small tumours who desire fertility. Currently there are enough data to conclude that radical vaginal trachelectomy with laparoscopic pelvic lymphadenectomy is a safe procedure with an acceptable recurrence rate (4%). In one study, pregnancies occurred in 31 (43%) of 72 treated patients and 36 (72%) of 50 pregnancies reached the third trimester.140

Radiotherapy is still the best choice for stage II–IV patients, but several randomised studies have shown improvement of survival with concurrent chemotherapy.81 Cisplatin seems to be the best drug for advanced squamous carcinoma as a single agent or in combination with other cytotoxic drugs. The optimum treatment of adenocarcinoma is less clear.

Fitting prevention strategies into available resources and existing programmes

New cervical cancer prevention methods must be introduced with consideration of added value and added cost. Otherwise, the rich could easily be over-treated, while the poor at higher risk are neglected. For example, the addition of HPV testing to cytology for screening, if repeated every year, cannot be cost effective and will lead to excessive interventions.82 Similarly, new preventive vaccines, if adopted with high acceptance,rationally must lead to less frequent screening to be cost effective.83 With less HPV16 and HPV18, the predictive value of positive screens will fall as the number and relative proportion of important lesions decreases, and the marginal gain in reassurance from negative screens will decrease as well.134 In general, regional planners must decide which prevention strategies do well, and which additions (or replacements) are most worthwhile.

To reach their potential, new cervical cancer prevention methods will need to be accessible and affordable to women who are currently underserved and at the greatest risk. The average cost per year of life saved with cytology-based screening programmes in countries where it has been successful is higher than most resource-limited countries’ annual gross per head income.

Cervical cancer screening programmes requiring one to three interventions in a woman’s lifetime are the most cost effective; more visits (common in high-resource regions) have a notably reduced cost-effectiveness.85 High programme coverage and immediate, effective treatment of positive cases (to minimise loss to follow-up) are crucial to achieve a reduction in cervical cancer mortality. Other factors that increase cost include the woman’s time requirement, need for transportation and the availability and cost of treatment of cancer cases. The relative importance of each of the different factors varies by country; thus programme selection has to be adjusted accordingly.86

One of the least expensive, easiest, and most widely assessed screening approaches is visual inspection with acetic acid or with Lugol’s solution. Visual inspection with acetic acid has been better studied, and its sensitivity estimates vary considerably (40–90%), partly due to lack of technique standardisation but also to the use of different gold standard methods.107–109 Additionally, the technique provides instant results that, if combined with treatment options such as cryotherapy, allows for 1-day see-and-treat schemes that decrease the overall cost of screening programmes substantially. Despite its low sensitivity, specificity, and predictive values when used as a stand-alone test, in scarce-resource areas visual inspection with acetic acid is a realistic screening method when the only alternative is no screening.171

We face an important challenge to apply HPV-based technology widely at low cost. The effect on cancer of vaccines against HPV infection will not be felt until about 20–30 years after a countrywide programme is introduced.165 The introduction of such programmes will probably require the involvement of donors like WHO, the Pan American Health Organisation, the GAVI Alliance, or the Bill & Melinda Gates Foundation to make vaccines available and affordable.

Like HPV vaccines, existing HPV tests are unaffordable and need to be done in specialised laboratories. A new HPV DNA test has been developed for low-resource regions by the Program for Appropriate Technology in Health (PATH) through a grant from the Gates Foundation. This test will provide results within a few hours with sensitivity and specificity similar to current commercially available tests, but at a cost of under US$5.172 Additionally, there are few infrastructure and reagent requirements, making HPV testing a practical possibility as a stand-alone screening method. Validation studies are currently underway.

We believe that a logical prevention strategy in regions with scarce resources would combine vaccination before sexual debut (if reduced cost or donated vaccine is available) and screening at an optimum age around 35 years with cryotherapy of all HPV-positive women except for those needing expert care—eg, for obvious cancers.88 Region-specific age rates of HPV prevalence should be considered to guide the ages of vaccination and screening. Combining vaccination and screen-and-treat strategies would reduce overall HPV endemicity and provide lasting population benefit.
Future directions

There are a number of important, active research topics that will soon affect clinical management of cervical HPV and precancer: the average risk and timing of clearance versus persistence of each type of HPV; the risk and timing of diagnosis of precancer given persistence of each of the types; the effect, if any, of age at infection on these rates of clearance, persistence, and progression; the risk, if any, of occasional re-appearance of an HPV type via reinfection or latency, if such a state exists for HPV, following initial clearance; the origin and significance of age-specific HPV prevalence curves that differ by region; the unique carcinogenicity of HPV16, including its molecular mechanism and natural history; the occult nature of HPV18 infection and related lesions and the increased importance of HPV18 in the development of adenocarcinoma; the validation of new molecular markers with better predictive values than CIN2 diagnoses for distinguishing between HPV infections with and without concurrent precancer and, among those without concurrent precancer, distinguishing those that are most likely to become precancer in the next 5 years; the immune response that prevents HPV reinfection and the response that underlies HPV clearance, including genetic influences on the success of these responses; and HPV natural history in immunosuppressed individuals.

The advent of highly efficacious prophylactic vaccines against HPV16 and HPV18 has irrevocably changed the landscape of research into cervical cancer prevention. No-one can predict how quickly we can move towards the goal of a vaccine that protects against all carcinogenic HPV types, with a safe and inexpensive, universally applicable route of delivery. However, each improvement in vaccines will force a reconsideration of the whole prevention effort, relative to resources. Based on the current vaccines, some new clinical directions are already evident.

Although approval of the Merck vaccine has proceeded rapidly in many countries, actual adoption has been irregular by country and, at least in the USA, by state. Mandatory vaccination, catch-up vaccination of older girls and young women, and vaccination of boys remain controversial issues. The protection offered by two rather than three doses of vaccine is being determined, which will affect cost. The trade-offs from the quadrivalent coverage of the Merck vaccine and the novel adjuvant of the GSK vaccine will be considered in cost-effectiveness analyses once more longer-term, type-specific efficacy data are released.

The use of HPV testing for primary screening will certainly increase, but the relative roles of cytology and HPV testing (alone or combined) will vary by country for years to come. We predict that HPV genotyping, first for HPV16, will eventually enter clinical practice as an important prognostic biomarker, and hope that its introduction is preceded by clear validation of reliable typing assays and useful follow-up algorithms. Once genotyping is reliable, type-specific viral persistence will immediately become an appealing and powerful prognostic biomarker. The discussion will shift to proper time intervals for defining persistence. Even if data demonstrate high risk in women with persistent infections in the absence of clearly diagnosed precancer, excisional treatment based on molecular tests alone will probably be very controversial and adopted as policy only in societies favouring aggressive clinical management.

Successful, widespread vaccine programmes will motivate reconsideration of optimum screening techniques and strategies. It is evident that the most clear-cut cytopathological and colposcopic abnormalities are caused by HPV16. The predictive values of screening protocols including HPV testing depend to a major extent on the risks associated with HPV16. Screening protocols will need to change with time as the population prevalence of HPV16 is gradually reduced.

Improvements in diagnosis will need to be made to match the improvements in screening. Colposcopically directed biopsy, in which the clinician targets the most abnormal lesion, is not sufficiently reliable or accurate to diagnose precancer in women referred by the combination of HPV testing and cytology. To foster better diagnosis for clinical practice and as a reference standard of disease, there is an urgent need for clinical trials to assess how biopsies should be taken to improve sensitivity and the reassurance of a negative colposcopic examination. Otherwise, the benefits of improved screening will not be fully realised.

Developing a simple and safe treatment for persistent HPV infection, including small precancerous lesions, would be an important breakthrough of immediate importance worldwide. We can now reliably and sensitively detect infection with carcinogenic types of HPV to identify women at risk of cervical cancer, but viral clearance is too common to justify immediate treatment, especially at younger ages. In some regions, for women past the peak ages of HPV prevalence, the most effective screening-based prevention strategies mandate immediate rather than delayed treatment (which results in losses to follow-up). Immediate treatment requires safe, inexpensive, simple destruction of the transformation zone and surrounding epithelium (or an equivalently safe and simple non-surgical approach). Finding a treatment better than the current forms of cryotherapy and loop excision should be a high priority.

We project that improved vaccines, screening tests, and management strategies will continue to emerge without evident end. New vaccine candidates and molecular biomarkers will supplant the prevention tools we have discussed. How quickly the institutions supporting cervical cancer prevention can react to
evolving opportunities is unclear. For example, recent studies of visual inspection with acetic acid indicate some benefit in detection of cervical cancers and large precancers. Adoption of this technique could reduce the incidence of, and mortality due to, cervical cancer in very low-resource regions. However, any large prevention effort based on this technique should take into account further improvements such as inexpensive HPV testing, which will be available within a few years. To frame and disseminate public-health messages at this pace is difficult, even when the change represents advances.

Conclusions

Much of the cervical cancer problem can be solved with existing or soon-to-be available technology, sufficient will, and modest resources. There is an enlarging repertoire of options for cervical cancer prevention for regions with varying needs and values, based on innovative technology and clear understanding of cervical carcinogenesis. Because of the importance of the problem and the feasibility of ameliorating it, we hope to see a major decrease in the numbers of women affected with this cancer within our lifetimes.

Conflict of interest statement

ACR is supported by an appointment to the Senior Fellowship Program at the National Institutes of Health (NIH). The programme is administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the NIH. The other authors are supported by the Intramural Research Program of the National Cancer Institute, NIH, and the US Department of Health and Human Services. None of the authors has a personal financial conflict of interest to report. The National Cancer Institute has a clinical trials agreement with GlaxoSmithKline (Rensselaer, Belgium) in which we are autonomously assessing their bivalent prophylactic HPV vaccine. The company will be provided the data as part of their efficacy demonstration, while we will independently analyse and publish the findings. The bulk of funding for this large trial, apart from the provision of vaccine and components required for the vaccine’s regulatory approval, is provided by the National Cancer Institute.

Acknowledgments

We are grateful for important contributions from Allan Hildesheim and Diane Solomon.

References

Seminar


An unexpected cause of fever and seizures

Dhanpat Kumar Kochar, Deepak Pakalapati, Sanjay Kumar Kochar, Parmendra Sirohi, Mahesh Pal Khatri, Abhishek Kochar, Ashis Das

In June, 2007, a 27-year-old man was brought to our emergency department by ambulance, having regained consciousness after a generalised tonic-clonic seizure. He had been having fever, chills, and rigors, on alternate days, for the previous 8 days. He had no past history of convulsions, head injury, febrile convulsions during infancy, birth trauma, meningitis, encephalitis, or psychiatric illness. There was no other past medical history of note. Until he fell ill, he had been working in a jewellery shop in Surat, Gujarat—a city where both falciparum and vivax malaria are endemic.

On examination, nothing abnormal was found. The patient’s full blood count was normal; biochemistry tests, including a blood glucose measurement, also gave unremarkable results. Electrocardiography, ophthalmoscopy, examination of the cerebrospinal fluid, and CT of the head all showed nothing of note. However, examination of the blood film showed trophozoites of *Plasmodium vivax*, at a density of 16 200 per μL (figure). A rapid diagnostic test (FalciVax, Zephyr Biomedical Systems, Goa, India) indicated the presence of parasite lactate dehydrogenase, specific to *P vivax*, and the absence of histidine-rich protein 2, specific to *P falciparum*.

6 h after he arrived, the patient had another generalised seizure. He was immediately given intravenous quinine, as per the WHO guidelines for severe vivax malaria; in addition, anticonvulsant drugs were given. Over the next 12 h, the patient had a total of eight generalised seizures, with intervals of 30–120 min, without regaining full consciousness. 48 h after treatment began, the fever subsided, and the patient became fully conscious.

Further blood tests—for dengue fever, leptospirosis, and HIV—gave negative results; repeat CT of the head, and electroencephalography, showed nothing remarkable. PCR, which was done as described by Kochar and colleagues, confirmed that the patient had been infected by *P vivax*, but not *P falciparum*. The patient was discharged 8 days after his arrival. When last seen, in August, 2007, he was entirely well.

*P falciparum* is known to cause cerebral malaria, which can manifest with seizures. The parasite multiplies in red blood cells, which adhere to the walls of small blood vessels, causing reduced cerebral blood flow. *P vivax* is less likely than *P falciparum* to cause severe illness—indeed, the typical 48 h interval between fevers, and benign course, have led to vivax malaria being termed “benign tertian malaria”. Classically, *P vivax* has not been thought to cause cerebral malaria. However, it is now known that severe *P vivax* infection can cause cerebral malaria—although, to our knowledge, this is the first case in which the cause of seizures has been confirmed as *P vivax* alone. How *P vivax* causes cerebral malaria is unclear, but recent studies indicate that the mechanism may be similar to that triggered by *P falciparum*. Other causes of seizures in malaria include hypoglycaemia, hyponatraemia, lactic acidosis—and other illnesses, such as epilepsy.

References