1. **What is going on at the FDA? • EDITORIAL**  
   *Page 1137*  
   The Lancet

2. **“I’m sorry” • EDITORIAL**  
   *Page 1138*  
   The Lancet

3. **Who takes responsibility for Zimbabwe? • EDITORIAL**  
   *Page 1138*  
   The Lancet

4. **Resistance to anti-influenza agents • DISCUSSION**  
   *Pages 1139-1140*  
   Yi Guan and Honglin Chen

5. **Sex, pregnancy, hormones, and HIV • DISCUSSION**  
   *Pages 1141-1142*  
   James A McIntyre

6. **The global challenges of birth defects and disabilities • DISCUSSION**  
   *Pages 1142-1144*  
   Richard H Carmona

7. **Lessons from Hurricane Katrina, tsunamis, and other disasters • DISCUSSION**  
   *Pages 1144-1146*  
   Joseph L Nates and Virginia A Moyer

8. **Depression and the risk of neurological disorders • DISCUSSION**  
   *Pages 1147-1148*  
   Andres M Kanner

9. **The lady aspirin for cardiovascular disease • DISCUSSION**  
   *Pages 1148-1150*  
   Freek WA Verheugt and Sidney C Smith, Jr
| 10. | **Operation “Restore Order” wreaks havoc in Zimbabwe** • NEWS |
|     | **Pages 1151-1152**  |
|     | Clare Kapp  |

| 11. | **Getting drugs to HIV-infected children in Cambodia** • NEWS |
|     | **Pages 1153-1154**  |
|     | Barbara Fraser  |

| 12. | **Claiming the right to health** • BOOK REVIEW |
|     | **Pages 1155-1156**  |
|     | Helen Epstein  |

| 13. | **Pfizer CEO's call to action** • BOOK REVIEW |
|     | **Page 1156**  |
|     | David Henry  |

| 14. | **Quack** • DISCUSSION |
|     | **Page 1157**  |
|     | Andrew Wear  |

| 15. | **Douglas Young** • DISCUSSION |
|     | **Page 1157**  |
|     | Priya Shetty  |

| 16. | **Margot Kruskall** • PERSONAL REPORT |
|     | **Page 1158**  |
|     | Ivan Oransky  |

| 17. | **Ciprofloxacin-resistant gonorrhoea in South Africa** • CORRESPONDENCE |
|     | **Page 1159**  |
|     | Prashini Moodley and A Willem Sturm  |

| 18. | **Prenatal diagnosis** • CORRESPONDENCE |
|     | **Pages 1159-1160**  |
|     | Caroline Mackie Ogilvie and Frances Flinter  |

| 19. | **Prenatal diagnosis – Authors' reply** • CORRESPONDENCE |
|     | **Pages 1160-1161**  |
|     | Allan Caine, Edna L Maltby, C Anthony Parkin, Jonathan J Waters, John A Crolla and on behalf of the UK Association of Clinical Cytogeneticists (ACC)  |

<p>| 20. | <strong>Prenatal diagnosis</strong> • CORRESPONDENCE |
|     | <strong>Pages 1161-1162</strong>  |
|     | Bernd Eiben and Ralf Glaubitz  |</p>
<table>
<thead>
<tr>
<th>21.</th>
<th><strong>Prenatal diagnosis</strong> • CORRESPONDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Page 1162</em></td>
</tr>
<tr>
<td></td>
<td>Margaret Somerville</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>22.</th>
<th><strong>Prenatal research: a very sensitive field</strong> • CORRESPONDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Page 1162</em></td>
</tr>
<tr>
<td></td>
<td>Sebastien Tassy, Guillaume Gorincour and Pierre Le Coz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>23.</th>
<th><strong>Eradication of poliomyelitis</strong> • CORRESPONDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Page 1163</em></td>
</tr>
<tr>
<td></td>
<td>T Jacob John</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>24.</th>
<th><strong>Eradication of poliomyelitis</strong> • CORRESPONDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Pages 1163-1164</em></td>
</tr>
<tr>
<td></td>
<td>Charles Shey Wiysonge, Alfred K Njamnshi, Emmanuel Nomo and Muki S Shey</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>25.</th>
<th><strong>Eradication of poliomyelitis</strong> • CORRESPONDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Page 1164</em></td>
</tr>
<tr>
<td></td>
<td>Noel W Preston</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>26.</th>
<th><strong>Department of Error</strong> • ERRATUM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Page 1164</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>27.</th>
<th><strong>Department of Error</strong> • ERRATUM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Page 1164</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>28.</th>
<th><strong>Efficacy and effectiveness of influenza vaccines in elderly people: a systematic review</strong> • ARTICLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Pages 1165-1174</em></td>
</tr>
<tr>
<td></td>
<td>T Jefferson, D Rivetti, A Rivetti, M Rudin, C Di Pietrantonj and V Demicheli</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>29.</th>
<th><strong>Incidence of adamantane resistance among influenza A (H3N2) viruses isolated worldwide from 1994 to 2005: a cause for concern</strong> • ARTICLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Pages 1175-1181</em></td>
</tr>
<tr>
<td></td>
<td>Rick A Bright, Marie-jo Medina, Xiyan Xu, Gilda Perez-Oronoz, Teresa R Wallis, Xiaohong M Davis, Laura Povinelli, Nancy J Cox and Alexander I Klimov</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>30.</th>
<th><strong>Increased risk of incident HIV during pregnancy in Rakai, Uganda: a prospective study</strong> • ARTICLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Pages 1182-1188</em></td>
</tr>
<tr>
<td></td>
<td>Ronald H Gray, Xianbin Li, Godfrey Kigozi, David Serwadda, Heena Brahmbhatt, Fred Wabwire-Mangen, Fred Nalugoda, Mohamed Kiddugavu, Nelson Sewankambo, Thomas C Quinn <em>et al.</em></td>
</tr>
</tbody>
</table>
| 31. | Post-surgery adjuvant therapy with intermediate doses of interferon alfa 2b versus observation in patients with stage IIb/III melanoma (EORTC 18952): randomised controlled trial • ARTICLE  
Pages 1189-1196  
|---|---|
| 32. | Obesity • REVIEW ARTICLE  
Pages 1197-1209  
David W Haslam and W Philip T James |
| 33. | Personal digital assistants in health care: experienced clinicians in the palm of your hand? • REVIEW ARTICLE  
Pages 1210-1222  
Daniel C Baumgart |
| 34. | Shaking the tree: mapping complex disease genes with linkage disequilibrium • REVIEW ARTICLE  
Pages 1223-1234  
Lyle J Palmer and Lon R Cardon |
| 35. | End of life: the humanist view • DISCUSSION  
Pages 1235-1237  
Julian Baggini and Madeleine Pym |
| 36. | Alcock's canal syndrome revealing endometriosis • SHORT COMMUNICATION  
Page 1238  
Hélène Nehme-Schuster, Cherif Youssef, Catherine Roy, Jean-Phillipe Brettes, Thierry Martin, Jean-Louis Pasquali and Anne-Sophie Korganow |
What is going on at the FDA?

The scene at the US Food and Drug Administration (FDA) headquarters last week, after the sudden resignation of Commissioner Lester Crawford on Sept 23, was depressingly familiar: a leaderless FDA was plunged into organisational uncertainty, while persistent claims of political meddling in scientific decisions rumbled on.

Announcing his resignation, Crawford—whose 2-month tenure earned him the dubious accolade of the FDA’s shortest-serving commissioner—said in a memo to all FDA staff that he was leaving because “it is time, at the age of 67, to step aside”. This reasoning has, understandably, been greeted with some scepticism. As Public Citizen, a US-based public-interest group, commented to *The Lancet*: “It doesn’t make sense that his age would be a problem now, when it didn’t seem to be 2 months ago.”

The Department of Health and Human Services (HHS), of which the FDA is a part, refused to provide additional details about Crawford’s decision to step down and declined to comment on whether he had been asked to resign. Crawford’s resignation came just 2 days after a Senate committee requested that the FDA’s conflict of interest policy be reviewed, citing concerns about use of waivers for conflict of interest requirements.

A further twist to the plot emerged with the almost instantaneous release of an announcement by the White House on Sept 23 saying that Andrew von Eschenbach, current director of the National Cancer Institute (NCI) of the National Institutes of Health, would take on the role of acting commissioner following Crawford’s resignation. von Eschenbach has, according to his colleagues, had a long-standing friendship with the Bush family and has attended gatherings at the home of former President George H W Bush and his wife Barbara, in Kennebunkport, Maine. Given the rapidity of the decision to appoint von Eschenbach—especially at a time when President George W Bush was focusing all his attention on the hurricane-hit southern states—it would be difficult to believe that the appointment had not been preplanned. At the time of writing, the White House had not responded to *The Lancet’s* requests for information about how far in advance of the Sept 23 announcement von Eschenbach was asked to take on the FDA role.

More intriguing facts have since emerged: according to an HHS spokesman there are no plans for von Eschenbach to relinquish his post at NCI when he takes on the role as acting FDA commissioner. Staff members at NCI were not aware of any plans to implement organisational changes that would help relieve von Eschenbach’s presumably substantial workload. Heading the FDA—and indeed the NCI—is a full-time job. How will von Eschenbach do both?

Crawford’s resignation is particularly interesting given that he has successfully weathered several crises since becoming acting commissioner in March, 2004. His time in charge has been dogged by criticisms of presumed pandering to political pressure in delaying a decision about whether or not to approve over-the-counter sales of emergency contraception. The most recent delay, announced at the end of August, led Susan Wood, assistant FDA commissioner for women’s health, to resign in protest. Crawford has also drawn fire for presiding over the organisation at a time of unprecedented insecurity for FDA scientists and rising concern about the stringency of the organisation’s drug-safety review process.

So will a new commissioner solve the FDA’s problems? Unlikely. It remains uncertain whether von Eschenbach will be simply an interim leader or will eventually be nominated by the President to take up the full role. What is clear, however, is that the upheaval of becoming suddenly leaderless is not what the FDA needs. Staff morale—dented by judgments that go against the findings of scientific committees—is dismal. And public confidence in the organisation is at an all-time low, despite Crawford’s decision to commission a comprehensive Institute of Medicine review of drug-safety policy, which is due to be published in mid-2006.

However, it is the persistent allegations of political meddling in FDA decisions that are most unlikely to be resolved by a switch of leader. Crawford’s troubles with politics mirror the FDA’s predicament 4 years ago, when George W Bush first took office. At that time, rumours were rife that the appointment of a new FDA head was being held up by a debate over FDA’s approval of an abortion drug. Political persuasion in scientific decisions seems to be a recurring theme for Bush’s FDA. This issue must be tackled before the organisation can properly return to serving America’s health. ■ *The Lancet*
“I’m sorry”

Everyone makes mistakes. Tragically, when mistakes happen in a health-care environment the consequences may be injury or death. Although it is common sense to apologise after an error has occurred, and make every effort to prevent the same problem recurring, it seems that in medical practice around the world this is not happening. Patients remain dissatisfied with how medical errors are handled, and doctors are concerned that saying sorry will somehow make matters worse—for themselves.

Addressing this mire of cultural and professional confusion, the Being Open policy announced this month by the UK National Patient Safety Agency (NPSA) is a welcome foundation on which to build a consistent approach to medical errors. Not least, because the guidance issued is both evidence-based and practical. The NPSA’s aim is that all NHS Trusts should develop local medical-error management policies by June, 2006; the cornerstone of Being Open guidance is to “say sorry for what has happened” and “explain exactly what went wrong”.

For a patient, an apology and frank explanation when things go wrong must be a step in the right direction. But what of the doctors who have to make these admissions? Will they bear the brunt of patients’ and their lawyers’ wrath? And will their legal defence organisations continue to support them after such candid conversations? The Medical Defence Union (MDU), a key UK agency providing legal assistance to doctors, supports the Being Open policy and for the last 50 years has been advising doctors to provide full explanations when errors occur, and to apologise. In the USA, several health-care providers have already adopted this type of approach. The University of Michigan Health System in Ann Arbor has recorded a reduction of around 50% in the number of malpractice claims and lawsuits after the introduction of an openness policy.

However, although apologies from individual doctors are crucial, it is rare for one doctor to bear sole responsibility for an error. As the NPSA guidance indicates, systems failures must still be identified and addressed if errors are to be minimised in the future. ■ The Lancet

Who takes responsibility for Zimbabwe?

Zimbabwe is in crisis. Since May, Operation Murambatsvina (“drive away rubbish”) has led to the forced evictions and demolition of communities countrywide, leaving hundreds of thousands of people homeless. This mass destruction has exacerbated the problems of drought and malnutrition, increased the devastation of HIV/AIDS, and worsened national economic meltdown.

A UN report has estimated that over 79 500 people with HIV/AIDS were among those evicted, disrupting home-based care, and Zimbabwe’s antiretroviral programme. The crucial issue of adherence to drug regimens has been seriously threatened. In public-sector antiretroviral programmes, it is estimated that 30% of patients have experienced a break in drug supplies of at least 2 weeks. 2 weeks is enough to further the development of clinically significant resistance to nevirapine, the cornerstone of the government’s first-line antiretroviral protocol. Interruption in treatment—coupled with the disruption to social and safety mechanisms, overcrowding, lack of access to clean water, food, and shelter, especially with the onset of winter looming—make the sick even more vulnerable.

These conditions could cultivate an epidemic of nevirapine-resistant HIV. Moreover, they may lead to the spread of other communicable diseases such as tuberculosis, malaria, pneumonia, and outbreaks of diarrhoea, dysentery, and cholera. The potential public health crisis is all too apparent. Immediate needs are to provide access to antiretrovirals and good quality care for people with HIV/AIDS. In the future, resources should be invested to quantify the extent of nevirapine resistance in this region.

It is increasingly difficult for agencies to gain access to people in holding camps to assess the level of provision of basic shelter and sanitation, let alone medical care. But such an escalating health crisis warrants an immediate response to provide humanitarian assistance where it is desperately needed. WHO-AFRO should take the lead—it has so far failed to do so. ■ The Lancet
Resistance to anti-influenza agents

The current heightened state of awareness about the pandemic potential of H5N1 avian influenza viruses in southeast Asia should not detract from the fact that human influenza viruses continue to cause a substantial burden of disease. More than half a million people are estimated to die from influenza-associated complications every year.1 As with many infectious agents, control of influenza virus infection is sought in two main ways: prevention, either by vaccination or drug prophylaxis, and treatment with antiviral drugs.

Although vaccines can reduce infections, influenza viruses are prone to a high rate of antigenic change in the haemagglutinin and neuraminidase surface glycoproteins, evolving rapidly to evade recognition by the adaptive immune system of the host.2 Vaccine designers of the WHO Global Influenza Programme are therefore forced to prepare new antigenic formulations every year, predicting likely candidate virus strains of avian H5N1 influenza have a special pandemic potential of H5N1 avian influenza viruses in southeast Asia should not detract from the fact that the importance of surveillance for the emergence and spread of resistant strains was given low priority. In this instance, resistance to adamantines in prevailing human H3N2 influenza viruses is the focus; however, the findings are also relevant to control measures for avian H5N1 influenza.

Bright and colleagues’ genotypic study shows a dramatic increase in the prevalence of amantadine-resistant H3N2 strains in some southeastern Asian countries, which has contributed to an increase of more than 30-fold in the frequency of worldwide viral resistance between 1994–95 and 2003–04. It seems that complacency had arisen from the low frequency of resistance observed in previous studies, and the importance of surveillance for the emergence and spread of resistant strains was given low priority. In the midst of increased vigilance towards emerging avian H5N1 viruses, Bright’s human influenza study serves as a timely reminder to watch this family of old foes closely. In view of the practical implications, the H3N2 amantadine-resistance study will be of great interest to those who have been involved in epidemic and pandemic preparedness and should encourage them to consider upgrading levels of surveillance.

Although the mechanism for generation of adamantane resistance has not been fully elucidated, it might be associated with inappropriate drug administration. The adamantanes rapidly induce resistance; resistant virus can be detected after just 3 days of treatment. Thus the wide use, and sometimes loose control, of particularly the older amantadine, which in some countries is available over the counter in “anti-flu” formulations. Drugs are an important component of plans for influenza containment, and stockpiling of both classes of drug has been advocated. A study by Rick Bright and colleagues,6 also in today’s Lancet, points to a potential major caveat for this plan: the emerging problem of resistance to antiviral drugs. In this instance, resistance to adamantines in prevailing human H3N2 influenza viruses is the focus; however, the findings are also relevant to control measures for avian H5N1 influenza.

Jefferson and colleagues’ findings suggest that improvements in vaccine coverage and formulations could further reduce initial influenza infections. While designers of vaccines targeting emerging pandemic strains of avian H5N1 influenza have a special challenge, in that the pandemic strain is not yet apparent and may appear with little warning, strategic vaccination coverage will be important to reduce infection and complications in at-risk groups.

Currently available anti-influenza drugs (table) include the adamantines, amantadine and rimantadine, which target the viral M2 ion-channel protein, and the neuraminidase inhibitors oseltamivir and zanamivir.4 These antiviral drugs are widely used, particularly the older amantadine, which in some countries is available over the counter in “anti-flu” formulations. Drugs are an important component of plans for influenza containment, and stockpiling of both classes of drug has been advocated. A study by Rick Bright and colleagues,6 also in today’s Lancet, points to a potential major caveat for this plan: the emerging problem of resistance to antiviral drugs. In this instance, resistance to adamantines in prevailing human H3N2 influenza viruses is the focus; however, the findings are also relevant to control measures for avian H5N1 influenza.

Table: Currently available antivirals for use against influenza

<table>
<thead>
<tr>
<th>Year of approval (US FDA)</th>
<th>Availability</th>
<th>Daily dose†</th>
<th>Resistance (in treated human beings)††</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amantadine 1966</td>
<td>Prescription*</td>
<td>2×100 mg</td>
<td>Regularly detected</td>
</tr>
<tr>
<td>Rimantadine 1993</td>
<td>Prescription*</td>
<td>2×100 mg</td>
<td>Regularly detected</td>
</tr>
<tr>
<td>Zanamivir 1999</td>
<td>Prescription</td>
<td>2×10 mg</td>
<td>Rarely detected</td>
</tr>
<tr>
<td>Oseltamivir 1999</td>
<td>Prescription</td>
<td>2×75 mg</td>
<td>Rarely detected</td>
</tr>
</tbody>
</table>

FDA=Food and Drug Administration. * Amantadine and rimantadine are available over the counter in some countries. †† Centers for Disease Control and Prevention suggested daily dose for treatment of influenza A in adults (age 13–64 years).

Published online
September 22, 2005
DOI:10.1016/S0140-6736(05)67340-0
See Articles
pages 1165 and 1175
amantadine could be to blame. Amantadine is a prescription drug in most countries, but is readily available over the counter in Russia, China, and perhaps other countries, where it may be consumed either knowingly or unknowingly at non-optimal doses. The recent outbreak of severe acute respiratory syndrome (SARS) might have contributed to the 2003 “spike” in adamantane resistance, which was most prominent in China and Hong Kong. During the outbreak, demand for antivirals, even those inappropriate for treatment of SARS, is likely to have increased. This kind of overreaction might also be observed should an H5N1 avian influenza outbreak occur.

The origin of amantadine-resistant H5N1 variants in Vietnam and neighbouring regions is still unclear. One possible explanation for the emergence of adamantane-resistant virus could be the alleged administration of the drug in agricultural practices in southeast Asia, in poultry farming in particular. However, this allegation has not been verified and is strongly denied by some governments. Naturally occurring resistance-associated mutations could also have arisen. Bright and co-workers observed that only two of 92 US patients shedding resistant H3N2 virus had a documented history of antiviral treatment before collection of virus. How the viruses affecting the remaining patients gained resistance remains open to conjecture. This relatively high frequency of resistance of unknown origin in patients raises the possibility that the resistant Asian viruses observed in Bright’s study may not have arisen solely as a result of exposure to drugs.

Is drug resistance now the norm? Assuming that distribution and administration of adamantanes is controlled, we predict that the “spike” of resistance to adamantane found in Bright’s study may wane over the next 2–3 years, perhaps due to continued viral evolution or, probably, competitive disadvantage of resistant virus. To ensure that resistance does subside, surveillance should be used as an active, rather than a passive, method to track the frequency of resistant viruses in real time, and not retrospectively. However, continued uncontrolled use of adamantanes might prevent the natural passing of this aberration, and continued misuse of drugs could cause resistant viruses to linger.

The studies published today reinforce the shortcomings of our efforts to control influenza. For too long, the development and manufacture of influenza vaccines and antiviral drugs has been of limited interest to drug companies. Since the existing defences are limited, it is critical that the most is made of them. Thus, although increasing resistance to adamantane is a cause for concern, it is still too early to call for market withdrawal or exclusion from stockpiling. What is needed is to improve control over the distribution and availability of the adamantane drugs, particularly in developing and southeast Asian countries, and to increase surveillance for resistance, so that these cheap and easily administered drugs can continue to play a part in our influenza control strategies. No one class of drug can be relied on: the newer neuraminidase inhibitors are also facing resistance issues, and despite their sometimes prohibitive expense, distribution should also be controlled. A wide-ranging approach must be taken to prevent and control endemic, epidemic, and pandemic influenza infection, with an emphasis on increased surveillance, better management of our existing tools of vaccination, M2 ion-channel blockers, and neuraminidase inhibitors, and the development of new agents to boost our arsenal.

*Yi Guan, Honglin Chen
Joint Influenza Research Centre (SUMC & HKU), Shantou University Medical College, Shantou, Guangdong 515031, China, (YG, HC); and Department of Microbiology and State Key Laboratory for Emerging Infectious Diseases, University of Hong Kong, Faculty of Medicine Building, Pokfulam, Hong Kong Special Administrative Region, China (YG, HC)
yguan@hkucc.hku.hk

We declare that we have no conflict of interest.

In today’s *Lancet*, Ron Gray and colleagues report an increased risk of incident HIV during pregnancy in Uganda. They draw on the unique research situation in Rakai to provide information on an important topic: does pregnancy increase the risk of HIV infection for women, and, if so, is this effect due to biology or behaviour? The results show a doubling in the risk of HIV acquisition during pregnancy—an HIV incidence of 2.3 per 100 person-years during pregnancy, compared with 1.3 per 100 person-years during lactation and 1.1 per 100 person-years in women who were not pregnant or lactating. Other investigators have reported high rates of HIV seroconversion in pregnancy, but the large number of women studied, the ability to follow them up through pregnancy and lactation, and the information on male and female sexual behaviour add to the strength of the Rakai study. The authors have been able to control for sexual behaviour to a greater extent than previous reports, and they show little effect of behavioural factors on the increased risk of HIV acquisition in pregnant women, suggesting that biological changes in pregnancy have an important role.

This possibility is worrying, and echoes some apprehensions about hormonal contraception and acquisition of HIV. But it is beyond our ability to change the biology of pregnancy, and any preventive strategy needs to tackle behavioural factors. Researchers have often ignored the reality of sex during pregnancy, and there are few data on sexual behaviour and HIV risk in pregnant women or their male partners at this time. Gray and colleagues report that 1% of pregnant women reported having two or more sex partners in the previous year, lower than the 2.7% of women who were not pregnant, whereas 36–39% of the male partners of pregnant women reported having other sexual partners. This dangerous intersection between behaviour and biology may be even worse in some other settings with high HIV prevalence. Two South African studies have reported transactional sex in more than 20% of pregnant women, with an attendant high risk of HIV infection. A report from one of these sites has also shown that 5% of pregnant women who tested HIV-negative at 27–28 weeks’ gestation were HIV-positive on retesting at 37 weeks.

The public-health implications of the Rakai findings are challenging. WHO has proposed a four-part approach to HIV in pregnancy and the prevention of mother-to-child transmission of HIV (PMTCT): preventing new infections in all people, preventing unintended pregnancy in HIV-infected women, preventing transmission to children of HIV-positive mothers, and providing care and support to infected women and their families. In addition to strengthening family-planning options for HIV-infected women, more attention needs to be given to preventing new infections during pregnancy, which carry a high risk of transmission to children. The prevailing view of PMTCT as a paediatric issue has hampered full implementation of these strategies: we have been too focused on an outcome measure of HIV-negative babies, and too caught up in the difficulties of implementing even simple nevirapine-based strategies to see the bigger picture. The provision of HIV counselling and testing in pregnancy provides a unique opportunity to keep HIV-negative women negative, but in many settings post-test counselling is provided only for HIV-positive women or is minimal, at best, for...
Comment

women who test negative. The need to provide risk-reduction counselling is underscored by the results from Rakai. Pregnant women need to be told of their increased risk of HIV infection and given information on protective strategies. The findings also highlight the futility of abandoning HIV testing in pregnancy and using universal PMTCT antiretroviral options.11 The missed opportunity to intervene to prevent new infections in women may counteract any short-term benefits of preventing transmission to babies.

Although it may not be easy to influence behaviour change, Gray and colleagues’ findings also suggest that HIV testing strategies for pregnant women need to be reconsidered. Most programmes currently test once in pregnancy, and will not detect new infections in late pregnancy, or provide antiretroviral interventions to these high-risk mothers. Some guidelines in the USA have recommended repeat HIV testing late in pregnancy,13,14 and a modelling exercise suggests that this would be cost effective in communities with an HIV incidence over 1 per 1000 person-years.12

The time has come to review management options for HIV-infected and uninfected pregnant women, and to improve use of the unique opportunity of antenatal care and counselling to protect both women and children.

James A McIntyre
Perinatal HIV Research Unit, University of the Witwatersrand, Diepkloof, Johannesburg, South Africa
mcintyre@pixie.co.za

I declare that I have no conflict of interest.


The global challenges of birth defects and disabilities

My office, the US Office of the Surgeon General, has declared 2005 “The Year of The Healthy Child”. This agenda includes all aspects of a child’s life—body, mind, and spirit—from preconception to prenatal care, and through all the developmental stages of childhood and adolescence. Advances in children’s health will require action to ensure the survival of even the youngest infants.

While additional research is needed, the public-health community, policymakers, and other partners must do all they can, individually and collectively, to apply what is already known to benefit the world’s children.

As noted in the Lancet’s series on child and neonatal survival, infections, preterm delivery, and asphyxia account for most neonatal deaths in developing countries.1 Numerous strategies are being generated to strengthen health systems and implement evidence-based low-cost programmes to address these causes and thus improve neonatal and maternal survival.2

In countries where infant mortality has already been reduced to less than 50 per 1000 livebirths, birth defects are emerging as the most common cause of neonatal mortality.3 We must begin to apply lessons learned about the prevention of birth defects in anticipation of the transition that developing nations will go through in addressing major causes of infant mortality.
An important step in this effort took place on Sept 11–14 in Beijing, China. Participants from all over the world and from several national and international organisations—the US Department of Health and Human Services, the Ministry of Health of the People’s Republic of China, WHO, the March of Dimes Birth Defects Foundation, and the Chinese Medical Association, among others—gathered for a major international conference on birth defects and disabilities in the developing world. A key goal of the conference was to identify steps to reduce gaps and disparities between developed and less developed nations to reduce the impact of birth defects and genetic diseases on infant mortality. As part of the conference, participants signed a declaration affirming the importance of birth-defect prevention and of health promotion for affected individuals.

The conference’s goals are timely. The reality is that children are still being born with birth defects that are preventable through effective low-cost strategies. Anencephaly, which causes infants to die before or shortly after birth, and spina bifida, which leads to lifelong disability, affect more than a quarter of a million newborn babies each year, despite the proven effectiveness of folic acid consumption in preventing up to 70% of such cases. Anencephaly, which causes infants to die before or shortly after birth, and spina bifida, which leads to lifelong disability, affect more than a quarter of a million newborn babies each year, despite the proven effectiveness of folic acid consumption in preventing up to 70% of such cases.4–6

Earlier this year, my office released new recommendations for folic acid consumption and encouraged all women of childbearing age to consume this important vitamin every day.7 Food fortification programmes also provide a low-cost means of increasing folic acid intake and preventing fatal and debilitating birth defects.8,9 In the USA, such programmes have helped reduce the prevalence of these birth defects by 27% in the span of a few years.10 Another powerful example of an underused prevention strategy is rubella immunisation, which provides a low-cost effective means of preventing unnecessary death and disability from congenital rubella syndrome.11,12

As more children survive, we must also focus on promoting the health of children and adults living with a disability. Not all disability is preventable, but people with disabilities often have preventable health problems. For example, in the USA, 30% of people with disabilities smoke, compared with 22% of those without disabilities.13 Similarly, 25% of persons with disabilities are physically inactive, compared with 13% of persons without disabilities.14

The Office of the Surgeon General recently released a call to action to improve the health and wellness of people with disabilities.14 The heart of this effort is to increase awareness—and from it action—that disability does not equal poor health. Unfortunately, such assumptions pervade the cultural environment through stigma and misunderstanding. Additionally, children and others with disabilities often face unnecessary physical barriers to taking full advantage of advances in public health. For example, inaccessible facilities and equipment create additional barriers for women with disabilities seeking to obtain mammography screenings. The results of such barriers are visible in the striking health disparities between those
living with and without disabilities. A cornerstone of the call to action is to promote understanding that people with disabilities can achieve healthier, more satisfying, and more productive lives. The four main goals of this effort (panel) aim to address several factors that affect the health of people with disabilities, including the social and physical environments.

In the USA alone, striving for and meeting these goals will improve the lives of the 54 million Americans currently living with a disability. If embraced worldwide, this initiative could improve the lives of countless more individuals. Continued efforts are needed to aggressively apply known public-health interventions to prevent infant mortality and disability where possible, and to promote whole health and wellness throughout the lifespan for people of all abilities. The promise of optimum health for people of all nations demands no less.

Richard H Carmona
United States Surgeon General, Parklawn Building 18–67, Rockville, MD 20857, USA
jessica.badger@hhs.gov

I declare that I have no conflict of interest.

1 Lawn JE, Cousens S, Zupan J, for the Lancet Neonatal Survival Steering

Lessons from Hurricane Katrina, tsunamis, and other disasters

A few weeks ago, Hurricane Katrina, a category 5 hurricane, stormed the Gulf coast of the USA, affecting an area as large as the UK. Katrina completely destroyed several communities, leading to the evacuation of millions of individuals. New Orleans, a major American city with a population of half a million inhabitants, was totally devastated and evacuated. This catastrophe follows on from the recent Asian tsunami. The worst, however, is that we knew of the catastrophic consequences of a powerful storm in that area of the USA and what we needed to do to reduce its impact. Last year, a simulation exercise in New Orleans, with a fictitious category 3 hurricane named Pam, predicted the scenario accurately,1 but funds to prevent the devastation were not allocated.

As we impotently watched television, the disaster unfolded. The hospital scenes of combined internal and external disasters already described in Houston2 repeated, despite the fact that we have learned how to prevent them. The medical centres that were supposed to take care of the sick became overheated traps with no electricity, water, communication, and other vital services; despite knowing the hurricane was approaching, they were unprepared to deal with the consequences of flooding and to evacuate their patients. Local and national leaders underestimated the storm and failed to act in time.

The initial costs of the disaster are estimated to reach or even surpass US$100 billion, and we still do not know the size of the tragedy in human lives and future ecological repercussions. Katrina left the affected region in chaos: confusion among the rescue and recovery teams, evidence of complete lack of preparedness with

References

insufficient immediately available physical and human resources, health-care systems incapacitated, urban anarchy, despicable crimes, while the world criticised American leaders and emergency organisations.

Early in the aftermath, the state of Texas absorbed about a quarter of a million evacuees in a matter of days. Houston, with the largest medical centre in the world (13 renowned hospitals, two medical schools, four schools of nursing, etc), an impressive city-wide health-care system (eg, Memorial-Hermann Healthcare System, 12 hospitals; Methodist Hospital System, four hospitals), and several colossal shelters (the George R Brown Convention Center, 1·8 million square feet; the Reliant Arena, 350 000 square feet; and the Astrodome, almost 400 000 square feet) fed, clothed, sheltered, and provided medical care and other services for many thousands of evacuees as they arrived. The initial medical response varied from complex services such as dialysis to more simple services (providing routine medications to people with chronic health problems such as arterial hypertension or diabetes mellitus), to treatment of infections acquired in contaminated water. Houston was only one of many cities that participated in the evacuation and relocation of affected people from Louisiana, Mississippi, and Alabama. Despite all the criticism about prevention and initial management, we cannot imagine another place in the world where this kind of response would be possible.

In the past 25 years, drought and famine have killed more than 300 000 people and adversely affected almost 300 million throughout the world; floods have killed almost 70 000 people and adversely affected more than 300 million; and earthquakes, volcanic eruptions, windstorms, and landslides have killed almost 200 000 people and adversely affected more than 60 million. The earthquake off the coast of northwest Sumatra on Dec 26 last year left the region with a death toll of more than a quarter of a million people, with millions displaced in poor conditions and at risk of disease outbreaks, tens of billions (in US dollars) in economic losses, and the need for many more resources for reconstruction.

As we have seen with the Asian tsunami and Hurricane Katrina, as populations increase in vulnerable areas, the problem is getting worse. The Centre for Research on Epidemiology of Disasters (CRED) at the Catholic University of Louvain in Belgium and the US Office of Foreign Disaster Assistance (OFDA) have collaborated to create a joint Emergency Disasters Database (EM-DAT) (figure 1). The trends shown are alarming, and although the exponential rise in the number of

Figure 1: Natural disasters reported, 1900–2004

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of disasters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>0</td>
</tr>
<tr>
<td>1920</td>
<td>0</td>
</tr>
<tr>
<td>1940</td>
<td>0</td>
</tr>
<tr>
<td>1960</td>
<td>0</td>
</tr>
<tr>
<td>1980</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>0</td>
</tr>
</tbody>
</table>

EM-DAT created (1988)
CRED created and OFDA began compiling (1973)
OFDA created (1964)
disasters could be biased by over-reporting and other factors, the vast and increasing number of people affected demands our attention (figure 2). In January, the UN took a big step forward in dealing with the prevention, management, recovery, and other critical aspects of disasters by adopting the Hyogo Framework for Action 2005–2015 Resolution. This resolution, which was partly based on the Yokohama Strategy and the Johannesburg Plan, addresses the specific gaps in present responses and challenges that disasters pose to communities around the globe. At the same time as organisations such as the Cochrane Collaboration and the US Centers for Disease Control and Prevention have groups dedicated to gather evidence-based data to better respond to these contingencies, we wonder what role all these resolutions and knowledge played in the decisions not to repair the New Orleans levees.

As we look at the whole picture, it appears that the poor outcome in many of these disasters is not the result of lack of knowledge but rather the result of inaction and poor implementation of the necessary measures to prevent, contain, or mitigate the impact of natural disasters on the populations exposed; this, of course, after discounting the enormity of the catastrophes involved. Memory of previous events in history is short; George Santayana once said, “He who forgets history is destined to repeat it”. It seems to us that if we do not react soon with rapid and effective changes to our current emergency responses and leadership, we will knowingly and sadly be repeating history in many more opportunities to come.

*Joseph L Nates, Virginia A Moyer
MD Anderson Cancer Center (JLN), and Center for Clinical Research and Evidence Based Medicine (VAM), University of Texas, Houston, TX 77030, USA
jlnates@mdanderson.org

We declare that we have no conflict of interest.

The relation between depression and neurological disorders is generally construed as unidirectional and expressed as an increased risk of depression in the presence of a neurological disorder. In the past 15 years, however, data from several studies have suggested the existence of a bidirectional relation between depression and various neurological disorders, including dementia, epilepsy, stroke, and Parkinson’s disease. This relation is expressed by an increased risk of occurrence or a worse prognosis of these disorders in individuals with a history of depression. For example, many studies have shown an increased prevalence of depression in patients with dementia. However, using data from a case-register study of almost 23000 patients with an affective disorder, L Kessing and P Andersen suggested that the relation expressed as the number of major depressive episodes leading to an inpatient admission increased the risk of developing dementia. Thus patients with three admissions had close to a three-times increased rate of dementia (95% CI 0.64–13.2), compared with patients with only one admission. In a separate study of 1003 elderly patients (all with a Mini Mental State score ≥26), the presence of significant depressive symptoms at baseline predicted a higher risk of cognitive decline 4 years later. Structural and functional neuroanatomical changes in primary major depressive disorders might help to explain these data.

Sheline et al compared the hippocampal volumes on high-resolution MRI in patients with major depressive disorders with those of healthy controls matched for age and sex, and found bilateral hippocampal atrophy in the patients. They also found a significant inverse correlation between the duration of depression and left hippocampal volume, suggesting that patients with more chronic and active disease were more likely to have hippocampal atrophy. The effect of severity of depression on the development of hippocampal atrophy has been supported by other investigators who found that patients with treatment-resistant major depressive disorders were more likely to have hippocampal atrophy than those who responded to therapy and healthy controls.

Hippocampal atrophy is not necessary, however, to yield memory deficits in major depressive disorders. For example, MacQueen et al found lower verbal memory scores on neuropsychological testing of patients with major depressive disorders. These verbal memory deficits were identified in patients with a single as well as multiple major depressive episode, but only patients with multiple episodes had hippocampal atrophy. As in Sheline’s studies, Macqueen’s group found a significant correlation between the duration of the depressive illness and the magnitude of hippocampal atrophy. Thus it is reasonable to hypothesise that hippocampal damage associated with major depressive disorders might compound hippocampal pathology in Alzheimer’s dementia, and in that way facilitate the development and/or accelerate the course of the dementia.

Can a timely and effective treatment of a major depressive disorder reverse this risk? In a study of 38 female outpatients with a history of major depressive disorders, Sheline et al found a significant correlation between reduction in hippocampal volume and the duration of depression that went untreated, but there was no correlation between hippocampal volume loss and time depressed while taking antidepressant medication or with lifetime exposure to antidepressants. Whilst these data are encouraging, more prospective and larger studies need to be done to answer this fundamental question.

Here is the evidence for a bidirectional relation between depression and epilepsy, stroke, and Parkinson’s disease.
Comment

The lady aspirin for cardiovascular disease

Since the 1980s, it has become clear that aspirin is a highly cost-effective and fairly safe component of the acute treatment and secondary prevention of various cardiovascular diseases. However, aspirin in the primary prevention of cardiovascular disease remains controversial. Although four of the five large randomised trials show a reduction in myocardial infarctions, none detected a reduction in cardiovascular mortality or stroke. There is even a suggestion of increased haemorrhagic stroke by the use of prophylactic aspirin in healthy individuals. Moreover, most of the trials included only men or mostly men, and there has been little evidence of the beneficial effect of prophylactic aspirin in women.

Women face the first events of cardiovascular disease later in life than men, with a break-even point occurring at about 65 years of age. Clinical trials in cardiovascular medicine have mainly focused on men because of this age difference in risk. Yet the consequences of cardiovascular events in women seem worse than in men. Therefore it is not clear why there have been so few trials on prevention and treatment of cardiovascular disease which included women, let alone in women exclusively. Specific female conditions, such as menopause, have been studied for prevention of cardiovascular disease, but such research has only dealt with hormone treatment. Although potential benefits have been suggested by many observational and mechanistic studies, postponing menopause with hormone therapy in randomised trials does not reduce the risk of cardiovascular disease in women.

I declare that I have no conflict of interest.

Andres M Kanner
Department of Neurological Sciences, Rush Medical College, Rush Epilepsy Center, Rush University Medical Center, Chicago, IL 60612, USA
andres_M_kanner@rush.edu

Therefore the recently published Womens’ Health Study (WHS) of low-dose aspirin in the primary prevention of cardiovascular disease in women over the age of 45 is welcome. It is not only the largest trial with aspirin so far, it is also the first done in women exclusively. The results have been commented on in the lay press, and are somewhat surprising, because this is the first primary prevention trial with aspirin, that showed a significant reduction in stroke numbers (table).

Myocardial infarction, which has been consistently lowered in the primary prevention observations in men, was not reduced (1.0% in both groups), until after the age of 65 (2.0% for aspirin and 3.0% for placebo, relative risk ratio 0.66 [0.44–0.97], p=0.04). No differences were seen in total (3.1%) or cardiovascular (0.6%) mortality. The overall risk of a cardiovascular event was less than 0.25% a year and strokes were more frequent than myocardial infarction in the women in WHS. For each four myocardial infarctions, five strokes occurred, whereas in the Physicians’ Health Study (PHS), which included men only, for each four myocardial infarctions there were only two strokes. Therefore, the chance of seeing an effect on stroke in WHS was much larger than in PHS.

It is unclear what underlying mechanisms might contribute to the observed sex differences in rates of stroke and myocardial infarction between WHS and PHS. Women tend to have a higher frequency of hypertension, whereas smoking is more common in men. Hypertension is a very substantial risk factor for stroke, although smoking carries a stronger risk for myocardial infarction. In WHS and PHS, which were 16 years apart, hypertension was probably treated better in WHS. Similarly, knowledge about treatment of cardiovascular risk factors in these health-care professionals, which was the study population in WHS, must have been higher than in the male PHS. For instance, statins and angiotensin-converting-enzyme inhibitors, both effective agents against the development and progression of atherosclerosis and with it myocardial infarction, were hardly available in the time of PHS, but would have been used in WHS.

However, the most important risk modifier, of course, is age. Both in PHS and WHS, age is the strongest modulator of the effect of aspirin. In the low-risk women under the age of 65 enrolled in WHS there is no effect of low-dose aspirin on myocardial infarction and the benefit on stroke increases with age. In PHS the same gradient of the aspirin effect was seen with age, but only for myocardial infarction.

Thus women not only present a first cardiovascular event later in life, they also have a higher risk of stroke than myocardial infarction. Because aspirin tends to prevent strokes in women with hypertension, the higher rates of strokes than heart attacks in young women might account for the main the results of WHS, wherein a benefit was seen for stroke, but not myocardial infarction in younger women taking aspirin.

Clearly, the major risk of aspirin is bleeding, even when used in low doses. Thromboxane A₂ production is almost completely inhibited in doses as low as that used in WHS. Low doses of aspirin promote gastric bleeding in existing gastrointestinal ulceration. The most serious fear of bleeding is cerebral bleeding because of aspirin. However, even in the large WHS enrolment including nearly 400 000 person-years, it has not been proven that aspirin causes cerebral bleeding. By contrast, in no previous aspirin study has the risk-benefit ratio of aspirin for gastrointestinal bleeding been so clearly shown. For each stroke prevented, three gastrointestinal bleedings occur. For those needing transfusion, this figure is less than one per stroke prevented. However, stroke is a catastrophic cardiovascular complication and devastating for the patient, family, and society, whereas gastrointestinal bleeding usually can be effectively treated. In this setting, hospitalisation and blood transfusion might be psychologically difficult for a healthy woman taking prophylactic aspirin. Importantly, in WHS there were no fatal excess bleedings because of aspirin use.

There have been only a few trials involving aspirin with daily doses lower than 75 mg. WHS is by far the

<table>
<thead>
<tr>
<th></th>
<th>Aspirin (n=19 934)</th>
<th>Placebo (n=19 942)</th>
<th>Relative risk (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total stroke</td>
<td>221 (1.1%)</td>
<td>266 (1.3%)</td>
<td>0.83 (0.69–0.99)</td>
<td>0.040</td>
</tr>
<tr>
<td>Ischaemic stroke</td>
<td>170 (0.9%)</td>
<td>221 (1.1%)</td>
<td>0.76 (0.63–0.93)</td>
<td>0.099</td>
</tr>
<tr>
<td>Haemorrhagic stroke</td>
<td>51 (0.3%)</td>
<td>41 (0.2%)</td>
<td>1.24 (0.82–1.87)</td>
<td>0.310</td>
</tr>
<tr>
<td>Fatal stroke</td>
<td>23 (0.1%)</td>
<td>22 (0.1%)</td>
<td>1.04 (0.58–1.86)</td>
<td>0.900</td>
</tr>
<tr>
<td>Non-fatal stroke</td>
<td>198 (1.0%)</td>
<td>244 (1.2%)</td>
<td>0.81 (0.67–0.97)</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Data does not equal total number because some cases not classifiable.

Table: Incidence of stroke in the Womens’ Health Study.
largest trial testing a dose of 100 mg every other day or 50 mg daily. It is unlikely that the somewhat disappointing results on the prevention of myocardial infarction are a result of dosing, as correctly pointed out by the WHS authors. The 25% increased risk of gastrointestinal bleeding and 40% of minor bleeding observed in WHS clearly indicates a strong effect of very-low-dose aspirin on platelet function. Also, in the few studies comparing directly two different doses of aspirin, there was no significant evidence that a dose lower than 75 mg is less effective than higher doses as used in PHS. 11 50 mg daily is even lower than the widely used European dose of 80 mg daily, which has been derived from the anti-inflammatory dose used in sick children (known as baby aspirin). The dose used in WHS is even lower (termed lady aspirin) and now proven effective for preventing stroke in low-risk women.

Thus the WHS adds greatly to current knowledge about primary prevention of cardiovascular disease with antiplatelet agents: a clear reduction in myocardial infarction in women over the age of 65 of nearly the same magnitude for men in PHS. 11 50 mg daily is even lower than the widely used European dose of 80 mg daily, which has been derived from the anti-inflammatory dose used in sick children (known as baby aspirin). The dose used in WHS is even lower (termed lady aspirin) and now proven effective for preventing stroke in low-risk women.

Factors, especially hypertension, 12 prophylactic aspirin should be considered, when high blood pressure is well treated. 11

*Freek W A Verheugt, Sidney C Smith Jr
Department of Cardiology, Heartcentre, University Medical Centre, Nijmegen, Netherlands (FWAV); and Center for Cardiovascular Science and Medicine, University of North Carolina, Chapel Hill, North Carolina, USA (SCS)

f.verheugt@cardio.umcn.nl

FWAV has received departmental research funds and consultancy and speaker’s fees from the manufacturers of the mentioned compounds, but declares no conflict with the above trials. SCS has received speaker’s honoraria, travel, and accommodation expenses from the manufacturers of the mentioned compounds, but declares no conflict of interest or involvement with the trials.

5 Grady D, Herrington D, Bittner V, for the HERS Research Group. Cardiovascular disease outcomes during 6·8 years of hormone therapy. JAMA 2002; 288: 49–57.
7 Duenwald M. Aspirin is found to protect women from strokes, not heart attacks. New York Times March 8, 2005; 5F.
Aid organisations are fighting an uphill battle to help victims of Zimbabwe’s disastrous Operation Restore Order which, according to UN envoy Anna Tibaijuka, left an estimated 700 000 people without a home or livelihood and caused chaos and suffering “on an unprecedented scale”.

The clampdown on slum dwellers and street traders—ostensibly to tackle crime—added to the humanitarian nightmares in the southern African nation where more than one-third of the population will soon be dependent on food aid; where HIV/AIDS rates of some 25% are expected to rise; and where life expectancy has plummeted to 33 years.

Although the UN launched urgent appeals for drought-stricken neighbouring countries like Malawi and Mozambique, its efforts to aid victims of Operation Restore Order have been frustrated by the Zimbabwean government.

Plans to launch a flash appeal for Zimbabwe following publication of the report by Tibaijuka, head of the UN housing organisation HABITAT, were thwarted in August. The government disagreed with her figures about the number of people in need of help, saying that only 2000 had been affected by the evictions. It also objected to the standard practice of non-governmental organisations working alongside the UN.

Zimbabwe’s Herald newspaper, which serves as the government mouthpiece in a country which does not tolerate independent journalists, reported on Sept 19 that Mugabe had told UN Secretary-General Kofi Annan that he would not allow the UN to use NGOs in his country. Mugabe has waged a vendetta against many charities, accusing them of being puppets of hostile western governments.

Mugabe’s comments in New York were greeted with anger and dismay by opposition leaders, who said they underlined his “indifference towards the suffering in Zimbabwe”.

“Operation Murambatsvina (Restore Order), and the continued obstinacy of the Mugabe Government on the issue of securing food relief, provides further confirmation that this is a Government consciously waging war against its own people, especially the weak, the poor and the needy”, says Paul Themba Nyathi, spokesman for the Movement for Democratic Change. The party has accused Mugabe of trying to destroy its power base in urban areas and driving people into villages where they can be more easily controlled.

“The man is deluded and Zimbabweans are paying a high price for this delusion. He may still be a hero to a number of African leaders but to the starving people on the ground he is playing ‘Russian Roulette’ with their lives”, adds Nyathi.

Mugabe remained defiant in the face of the international criticism. In a virulent speech to the UN General Assembly, he rounded on HABITAT for remaining silent about Hurricane Rita’s impact in the US city of New Orleans “where a whole community of mainly nonwhites was deliberately abandoned to the ravages of Hurricane Katrina as sacrificial lambs”.

“Where is the Zimbabwe-famous HABITAT, I ask? Why should it maintain ominous silence? For here is real work of the homeless for it! This, indeed, is where it rightly belongs and not anywhere in Zimbabwe!” Mugabe said.

The 81-year-old president, who has ruled since independence in 1980, also flatly rejected UN suggestions that he should provide basic new accommodation as a matter of emergency.

“We have rejected the scandalous demand . . . that we lower our urban housing standards to allow for mud huts, bush latrines, and pit toilets as suitable for the urban people of Zimbabwe and for Africans generally”, Mugabe said. “Nothing could be more
The government claims it will provide funds for new houses, at the expense of cutting education and health programmes. But with the economy in free fall and the country at risk of expulsion from the International Monetary Fund for non-payment of debts, most analysts doubt that the promised housing funds will materialise.

In the meantime, Zimbabweans who lost their homes and source of income in Harare and other targeted cities are struggling to cope. The lucky ones have returned to their villages and been absorbed by local communities. There are reports that others have drifted back to a precarious existence in urban areas, saying they could not survive in the countryside, especially in the face of hostility of local tribal chiefs. Others remain in holding centres and transit camps, where conditions are said to be far worse than the demolished shantytowns.

In a survey released in August, and based on urban areas most affected by Operation Restore Order, ActionAID Southern Africa Partnership Program said that 70% of 23 500 respondents reported losing their shelter; 76% their primary source of income. It said that 22% of households reported that their children had been forced to drop out of school. Some 31% of the households were hosting orphans and 13% had someone who was chronically ill, usually with HIV/AIDS. It said around 15% of surveyed households reportedly had lost access to antiretroviral treatment—although the figures did not include Harare.

In her report, which was based on a 2-week visit to Zimbabwe, Tibajuka said it would take several years to recover from the “untold suffering” inflicted by the demolition campaign which started on May 18.

The vulnerable were hardest hit, said Tibajuka. Many of Zimbabwe’s 1·3 million orphans were affected, including those living in orphanages destroyed during the campaign, orphans left alone because guardians could no longer care for them, and street children rounded up and placed in transit camps, she said.

An estimated 200 000 children under 11 years old were displaced far from their schools and with no means of transport to get to new schools. Many teachers were also displaced. The report said the health consequences were dire. “The combination of overcrowding for evicted people living with friends and hardship for those sleeping outside will have a direct consequence in terms of other communicable diseases like pneumonia and TB”, Tibajuka says.

With an estimated 25% of adult Zimbabweans infected with HIV/AIDS, the report estimates that over 79 500 people with the virus were displaced, disrupting antiretroviral treatment, home-based care and prevention programmes such as condom distribution.

The report cited fears from medical groups that displacements could lead to an increase in HIV across the country and an increase in resistance to treatment with nevirapine. A more aggressive strain of tuberculosis, hitherto confined to Harare, was also expected to spread to other parts of the country.

Home-based care for AIDS sufferers was disrupted in many places, with a “significant” number without adequate shelter or food. Many trained volunteers were displaced, leading to interrupted services for patients not directly affected by the evictions, says the report.

Reports by Amnesty International and other human rights groups echoed similar sentiments.

UN Secretary General Kofi Annan said Operation Murambatsvina was a “catastrophic injustice”. The World Food Programme is currently feeding 1·1 million people, including 900 000 school children. Prospects for the next harvest, due in March or April, look bleak. Chronic shortage of fuel and inputs like pesticides and fertilisers have crippled planting preparations of the maize and tobacco crops.

Zimbabweans grapple daily with shortages of basic necessities and prices of commodities like sugar and bread are soaring. Even relatively affluent Zimbabweans with dollars to buy petrol frequently have to spend the night in queues. The fuel shortages have also affected the country’s ambulance fleet.

In an interview with the Standard newspaper on Sept 18, Zimbabwe Medical Doctors Association president, Takarundu Chinoyoka, complained of the chronic shortage of supplies. He said basic drugs such as MMT, used to treat internal bleeding, antibiotics and protective clothing, were in short supply.

Public-sector doctors went on strike in August to demand huge salary increases to keep up with triple-digit inflation. Following the government rejection of their demands, more than 20 junior and middle-ranking doctors left for UK, USA, Australia, and neighbouring African countries. Only 300 junior doctors, 150 middle ranking, and fewer than 100 specialist doctors remained in the state sector—compared with estimated needs of at least 2500 general and specialist doctors for government central hospitals, according to Chinoyoka.

In the aftermath of Operation Restore Order, Zimbabwe’s blood bank also virtually ran dry as regular donors were unreachable after becoming displaced by the demolitions.

Clare Kapp
In front of an airy, two-storey house, a dozen children, pressing their hands together in the traditional Khmer greeting, crowded around John Tucker, a sandy-haired Texan. Tucker, 57, puts an arm around a slender girl and introduces her.

“When she came to us, she was 7 years old and weighed 15 pounds”, Tucker says. “She had been abandoned by her family. Today, she’s riding her bicycle to school, she weighs over 50 pounds and she’s in second grade. She’s one of our miracles.”

Amid the chatter and laughter, it isn’t easy to see what brought the children to Tucker in the first place: all had AIDS and no one to care for them. When the “miracle girl” arrived she had a CD4-positive T cell count of 39.

At first glance, Cambodia seems an unlikely place for such miracles. When the AIDS epidemic spilled over the border from neighbouring Thailand in the mid-1980s and the adult infection rate shot up to 3%, the country had almost no resources with which to respond. At the time, the country was still reeling from genocidal rule of Khmer Rouge leader Pol Pot, whose reign of terror between 1975 and 1979 left nearly 2 million Cambodians dead. Virtually all of the country’s educated professionals had either been killed or had fled into exile. Of the roughly 1000 doctors who were in Cambodia when Pol Pot came to power, survived by pretending to be a labourer. He was sent to work in the rice paddies in the northwestern province of Battambang.

After the fall of the Pol Pot, the University of Health Sciences reopened with foreign faculty members and Mean Chhi Vun was able to finish his studies. Today, most of the faculty is Cambodian and there are about 3000 physicians in the country.

As head of NCHADS since its founding in February 1998, Mean Chhi Vun has overseen an expansion of AIDS treatment and services and coordinates with international aid agencies and non-governmental organisations, including the programme headed by Tucker.

Tucker, a former financial planner, seems an unlikely person to have come to the rescue of the more than 300 HIV-infected children currently receiving antiretroviral treatment through a programme run by the US-based Catholic Maryknoll Missioners.

He recalls arriving in 2001 as a lay volunteer with Maryknoll: “I couldn’t spell AIDS and I’d never met anybody with AIDS”. At the time, Maryknoll had a programme in Phnom Penh for adults living with HIV and AIDS. “Then people started bringing us children who had been abandoned. They had extended families that were too poor, too afraid or too sick to care for the kids.”

A father of seven and grandfather to 10, Tucker started shelters for the first children. The programme now has 152 children in six group homes and 213 in home-based care, along with a staff of field workers, house mothers, two physicians and a battalion of motorcycle-mounted “DOTS (directly observed treatment, short-course) workers who take medicine to the children’s homes twice a day and watch while the kids take it.

“It would be ludicrous, in a country where the average income is US$15, to leave $20 or $30 worth of medicine in the house and ask (the family) to give it to (the children). They would sell part of it”, Tucker says.

Recognising the extreme poverty of most of the patients, the Maryknoll programme also provides small allowances—an average of $2-50 a week—to some families to help with rent or food costs.

Tucker’s can-do attitude sometimes runs up against unexpected barriers in a country where doctor-patient relationships are traditionally vertical and
Cambodia’s focus on HIV-positive children could mean other children miss out

Patients are not encouraged to ask questions. “You have to learn to listen”, he says, recalling a surgeon who refused to operate when he found out a young patient had AIDS. Tucker finally discovered that the hospital had no protective gear. A donation of two boxes of gloves and 10 pairs of goggles solved the problem and the surgery was performed.

David Pugatch, an assistant professor of paediatrics at Brown University School of Medicine (Providence, Rhode Island, USA) who heads Brown’s paediatric HIV programme, has followed Tucker’s young patients for several years. “The world should hear more about how those kids have done. Those who have survived have done extremely well...reconstituting their immune systems”, he says.

A major challenge is making sure that the children get and take their medicines. “For the children in the Maryknoll programme the directly observed treatment with high active antiretroviral therapy (DOT-HAART) approach worked to ensure medication adherence, and we know that the children who received at least 6 months of DOT-HAART did extremely well”, says Brown. Further study will be needed to assess long-term clinical outcomes and cost-effectiveness of this approach, he added.

The lack of resources remains an obstacle, however. Because large donors balk at funding lifetime medical care, Maryknoll has had to ask individuals, church groups, and other organisations to sponsor children for $45 a month—$22 for the medicine, $13 for blood tests, and $10 for delivery of the medicine.

NCHADS has launched a paediatric AIDS programme that will gradually take over the antiretroviral therapy, but the Maryknoll programme will continue to operate the houses and provide home-based care.

Mean Chhi Vun says he is trying to stretch NCHADS’ scant resources. The system is being decentralised, with staff in each district working closely with non-governmental organisations and village outreach workers. Because labour-intensive DOTS is expensive for large numbers of patients, NCHADS is depending on village health-care workers and peer support. “Compliance is generally high because they know where patients are and when they move”, Mean Chhi Vun says.

NCHADS has succeeded in bringing down the HIV infection rate to 1·9% in 2003, Mean Chhi Vun says. Although he is optimistic, the agency’s director worries about the hurdles that lie ahead. The emergence of drug-resistant virus, in particular, is a concern.

“You can’t treat any infectious diseases without getting resistance, but in developing countries the stakes are much higher, because the more resistance you have, the more your cost-effectiveness plummets”, says Julian Elliott of the National Centre in HIV Epidemiology and Clinical Research at the University of New South Wales in Australia, who is currently a technical adviser to NCHADS in Phnom Penh. “You have to switch people to expensive therapies, but they’re less effective.”

The Cambodian government has allocated US$1 million a year for drugs until 2010, and international aid is expected to enable the country to meet its target of providing antiretroviral therapy to 11 000 people by the end of this year. As of July 30, the number stood at 8350, of whom 582 were children.

No one is sure how many children in the country are HIV positive. Statistically, it could be about 10% of the 120 000 to 170 000 cases nationwide. Tucker has heard figures of 7000 to 8000, but thinks those are high. NCHADS puts the figure at 3000. Whatever the number, long-term funding is a concern.

“The crunch is going to be around 2008 or 2010, when there’s going to be a number of countries whose accumulated cohort of people on ART is going to be quite large”, Elliott says. He expects the decisive factor to be cost-effectiveness, which will be determined by “quality use of antiretrovirals and adherence”.

As he beams at his young charges playing in the front yard of the group home, Tucker says he worries that programmes like his, by focusing on children with AIDS, may be missing the bigger picture. In a country with a total health budget of $18 million, where infant mortality is 89·4 per 1000 livebirths and under-5 mortality is 115 per 1000 livebirths, malnutrition, diarrhoea, and respiratory infections are even bigger threats than AIDS.

“200 children under 5 years die every day in Cambodia from preventable diseases, according to the Millennium Goals”, Tucker adds. “It’s sometimes unfair that kids with AIDS get attention and kids without AIDS don’t, and they die of all these other problems because there isn’t money and resources addressed to them.”

Barbara Fraser
Claiming the right to health

I’ll never forget my first visit to an African upcountry hospital. It was in Uganda in 1993, 7 years after the end of the civil war, in a remote part of the country where no reconstruction seemed to be underway. There was one German doctor and some nurses who seldom turned up. Water was supplied in jerry cans. The radiography machine was powered by a generator for an hour a day, if there was enough fuel. The bathrooms had been completely gutted and did not function, except as aviaries for the finches that made their nests in the porcelain scraps on the floor. Patients shared beds or slept on the floor. The doctor showed us a bed frame with half the springs missing. “We found a way to get this wire”, he said pointing to one of the remaining springs. “The people can bend it themselves.”

The dire state of this hospital, and of Uganda’s health-care system in general, was the product of years of conflict followed by years of fiscal austerity measures, mandated in part by foreign donors. Today, the situation is little improved. Despite a decade of economic growth in Uganda, health indicators—rates of infant, child, and maternal mortality—have not improved since the mid-1990s. This is a disgrace. But should human rights activists be calling the Ugandan government to account?

The Universal Declaration of Human Rights consists of two parts. The first 21 articles deal with civil and political rights, while the remaining nine articles are about economic and social rights. Human rights experts concur about how to protest violations of the first set of rights, but argue about the second. For example, Aryeh Neier and Ken Roth—the past and present directors of Human Rights Watch, respectively—have both written that it is unrealistic to hold governments accountable for violations of social and economic rights (A Neier, Taking Liberties (2003), K Roth, Defending economic, social, and cultural rights Hum Rights Q 2004; 26: 63–73). After all, how can we expect the leaders of a desperately poor country like Uganda to provide health care to all its people? In fact, the International Covenant on Economic, Social and Cultural Rights requires only “progressive realisation” of social and economic rights, meaning that access to health care should improve steadily with economic development. Even so, Neier and Roth are sceptical. Countries like Uganda have many pressing budget concerns, including national defence and investment for economic growth, both of which affect health. So what right do we in the west have to insist that such countries spend their meagre resources on health? Shouldn’t Ugandans decide what Uganda spends its money on? And doesn’t Uganda’s gradual democratic transition, monitored by organisations such as Human Rights Watch, ensure that Ugandans will have the right to do this?

We know it should work that way, but we also know it isn’t so simple. Civil and political rights are threatened in any country where so many people are so desperately poor and vulnerable, no matter how “democratic” such a country may be in theory. Some degree of economic and social well-being seems to be necessary for human rights to flourish, so it is not easy to distinguish cart from horse.

Recently, the contested relation between health and human rights has drawn increasing attention. Human rights experts are taking on such issues as HIV/AIDS, abortion, family planning, and sexual violence. Perspectives on Health and Human Rights contains 30 essays that attempt to create a framework for thinking about this complex field. It is a valuable book, for the guidance it provides and for the questions it raises.

Reading Paul Farmer and Nicole Gastineau’s essay, “Rethinking Health and Human Rights”, on the right to health care for marginalised groups, I could see Neier’s and Roth’s point. Among other examples, Farmer and Gastineau describe a scene from a Russian jail, where prisoners are kept in pretrial detention, often for a year or more, in overcrowded and squalid conditions. Prison health services being what they are, many prisoners develop drug-resistant tuberculosis and die. Farmer argues that the prisoners have a right to specialised treatment for drug-resistant tuberculosis and that “the best way to protect [prisoners’] rights is to cure them of their disease”. But it seems to me that if you removed the civil rights abuses in the prisons, the prisoners wouldn’t have had their lungs scarred by drug-resistant tuberculosis in the first place, and wouldn’t need advanced, expensive treatment.

Indeed, tackling underlying violations of civil and political rights would address many other apparent “right to health” abuses. Domestic violence and rape are human rights abuses, not just because they increase women’s risk of HIV/AIDS, but because they violate their right to freedom from degrading, inhumane treatment. It is important for doctors like Farmer to remind us of the health consequences of poverty and marginalisation, but one can see why Neier and Roth want to keep the focus on civil and political rights, while
the authors emphasise the need for monitoring the maternal mortality rate over time to ensure that it doesn’t worsen, and that it improves with economic growth. Yamin and Maine don’t mention this, but the situation in some countries is alarming. Uganda’s maternal mortality rate has not improved despite economic growth, while Malawi’s maternal death rate has doubled since 1992. There is no excuse for these rates. Although Malawi is poor, it is not poorer than it was 13 years ago; the HIV/AIDS epidemic is severe, but not so severe as to drastically affect rates of maternal mortality. The reasons are clear and unforgivable. In the past 20 years, public investment in health services has plummeted, and many of Malawi’s health workers have moved to developed countries where salaries and working conditions are far better.

The authors emphasise the need for recognising that such violations also have health effects. Alicia Ely Yamin and Deborah Maine take a more straightforward approach in their essay, “Maternal Mortality as a Human Rights Abuse”. 99% of the more than half million maternal deaths that occur annually take place in developing countries, and nearly all could be prevented with fairly simple medical procedures. Economic development itself does not seem to be an independent factor in preventing maternal deaths. However, a functioning health-care system—with a reliable supply chain, a referral system, and adequate human resources—is crucial. Monitoring supplies of sterile gloves, oxygen, and staffing levels is not easy, but Yamin and Maine show how ignoring these mundane matters can lead to human rights abuses.

But McKinnell also tries to paint on a larger canvas. He criticises the world’s inadequate response to HIV/AIDS, and acknowledges that the pharmaceutical industry could have done a lot better. But he is proud of Pfizer’s philanthropy in the form of drug-donation programmes. Many would prefer that he had simply dropped the prices of Pfizer’s drugs.

McKinnell reserves his most severe criticisms for those countries he considers are not paying their fair share of drug-development costs. He singles out Canada and Australia for demanding drug prices below what Pfizer charges their US-based customers. Apparently, we are “free-riding” on the investments made by the US public. He proposes that we must pay more to ensure a future supply of innovative medicines. The USA may pay a larger share of drug-development costs than other countries, but not dramatically more so per capita than Europe. This investment generates wealth for the country and the USA uses its power in multilateral and bilateral trade agreements to protect that intellectual property. There are strong economic arguments for differential pricing of medicines so that less wealthy countries pay less. Surely social welfare and company profits will both be maximised if the products are available in all markets and the prices that are charged reflect local price sensitivities?

A lot of what McKinnell proposes is unoriginal, and the obvious answer to many of the problems he identifies is the introduction of a good system of lowest-cost primary care; it never gets a mention. In suggesting reforms he proposes big business solutions—greater choice, more competition, innovation. The book lacks integration; it reads like an American breakfast menu and provides a similar degree of satisfaction.

David Henry
david.henry@newcastle.edu.au
Historical keywords
Quack

From the Dutch quaken, to prattle, chatter, and perhaps to sound like a duck, quacks were thick on the ground by the 17th century—commonly as “quack-saving physicians”. More or less interchangeable with “empiric”, “mountebank”, and “charlatan”, the label was often deployed by university-educated medical practitioners as a way to distinguish themselves from, and disparage as dangerous, those who advertised products on travelling stages and street corners, or in broadsheets.

In a competitive medical marketplace, there were two key issues between quack vendors and the learned physicians. One was special recognition by the state, which no group had in the UK before the Medical Act of 1858. The public had to judge for themselves who was a proper practitioner. The accusation of quackery, therefore, was a way of doing down one’s commercial opponents. The second issue was the question of what constituted best practice. Up to the 18th century, learned physicians claimed that they provided individualised treatment for their patients by taking into account their constitutions and lifestyles. Quacks, by contrast, sold medicines for specific diseases and for all diseases, and advertised them as suitable for all types of patients. For instance, in the late 17th century Lionel Lockyer claimed that his pill “extracted from the rays of the sun” cured “All or Most part of the known Diseases and Distempers of Man”. Similarly, Anthony Daffy’s “elixir of health” acted on all kinds of diseases and was suitable for “all Ages, Sexes, Complexions and Constitutions”.

Such products were often chemically based and were popular because they acted more immediately than the herbal remedies of learned practitioners. They were also cheaper than doctor’s visits and provided anonymity in the treatment of diseases like syphilis. Remedies such as Daffy’s began to be distributed across England by the late 17th century. Thereafter druggists and industrially based pharmaceutical companies followed the same philosophy of supplying remedies for specific diseases.

Today we have a reversal of who is a quack. Establishment medicine prescribes medicine for diseases rather than treating the individual patient. “Alternative medicine” now fulfills that role and is at times linked to quackery. Perhaps the 20th-century colloquial use of quack, which emerged from Australia and the army to denote any doctor, is a comment on the claims of medicine of whatever kind.

Andrew Wear
Wellcome Trust Centre for the History of Medicine at UCL
a.wear@ucl.ac.uk

Lunch with The Lancet
Douglas Young

When I meet Douglas Young in a quiet Soho restaurant, he is still buoyant from the news that his team at Imperial College, London, UK, have just bagged US$11 million from multi-billionaire Bill Gates to tackle the health problems of the world’s poor. His excitement is understandable—as research grants go, a Gates grant is as close to hitting the jackpot as it gets. Gates’ charitable foundation recently gave $437 million to top-level scientists around the world, including Young, to solve 14 Grand Challenges in public health.

Young’s mission is to investigate latent tuberculosis—a form of Mycobacterium tuberculosis infection harboured by about one-third of the world’s population. Latent tuberculosis is something of a black box; scientists believe it lurks in the lungs, but no-one is certain. The drugs that target the active bacterium are useless against the inert, latent form. In places where people are often co-infected with HIV and M tuberculosis—Africa, for example—latent tuberculosis causes huge problems. Young explains that as the immune system, which suppresses the latent form, becomes ravaged by HIV, the bacterium can become active, speeding up the progression of HIV in the infected person.

Young tells me how he was thrust into a first-hand experience of how infectious diseases devastate the lives of the poor during a stint working on leprosy in Mumbai, India, in the 1970s. He was one of several researchers around the world looking into ways to tackle the disease, including diagnostic tests to distinguish it from tuberculosis. But it dawned on him that although leprosy might have been top of the research zeitgeist at that time, tuberculosis was killing far more people. He switched to working on tuberculosis instead, and became “hooked”. “I can’t imagine working on any other disease”, he says.

Young’s research project is one of many that are geared to achieve the Millennium Development Goals to be met by 2015, one of which is to halt and begin to reverse the incidence of tuberculosis. What will happen to funding for tuberculosis after 2015 is difficult to predict, but Young hopes it doesn’t go the same way as leprosy. In the 1990s, international efforts strove to eliminate leprosy as a global public health problem by 2000. As the new millennium dawned, money for leprosy projects dried up because of a perception that targets had been reached.

Just before we leave the restaurant, I ask what he plans to do when he retires in 8 years. He is refreshingly honest: he plans to give it all up, kick back, and relax. And if, as is more than likely, Young and his team crack the problem of latent tuberculosis, he will have more than earned it.

Priya Shetty
priya4876@hotmail.com
Margot Kruskall

Haematologist who pioneered autologous blood banking. Born on June 13, 1949, in New York City, NY, USA, she died of ovarian cancer in Dover, MA, USA, on Aug 27, 2005, aged 56 years.

Margot Kruskall’s main interest in the early 1980s was autologous blood transfusion. “This was before the AIDS era, and autologous blood really received a lot of lip service from clinicians and very little use, which Margot recognised”, said Harvey Klein, director of the National Institutes of Health Clinical Center’s department of transfusion medicine. “She was one of the first people to really approach it in a scientific way—the mechanisms for collecting blood and storing it, and the effectiveness of using autologous blood.” She “turned it into a science almost single-handedly”, Klein told The Lancet. “She also established what clearly was the premier programme in autologous blood in the USA in Boston” at the Beth Israel Deaconess Medical Center. This programme would serve as a model for others in the USA and abroad.

For her foresight, “she was heralded as the AIDS epidemic occurred”, Klein said. But despite her strong beliefs in autologous donation, Kruskall was scientific and rational about donations from friends and family. “When HIV came about, there was panic everywhere, and patients were afraid to get any blood”, said Justine Carr, who worked with her for 25 years. “The immediate suggestion was to only receive blood from designated family and friends. She was absolute and resolute in that the answer was no, you have no idea what’s going on in their life. I was impressed that she stood by it, and could not be dissuaded.” And as the blood supply became increasingly safe, and autologous donation became less useful and cost effective for the patient, Kruskall was up front about the risks and benefits. “As the data changed, Margot said ‘here’s what the status is in 1995 as compared to 1983’” Klein said. “She was a superb scientist.”

Kruskall graduated from Jackson College at Tufts University, Medford, MA, and from the Medical College of Wisconsin, Milwaukee, WI, and then completed a residency at Mount Auburn Hospital in Cambridge, MA, and a fellowship at Beth Israel Deaconess Medical Center. She spent the rest of her career in Boston at Beth Israel and Harvard Medical School.

In 2002, Kruskall published a paper in The New England Journal of Medicine that made headlines. A woman who needed a kidney transplant had consulted her, and she had found that based on HLA genotyping, two of the woman’s three children were not hers. Kruskall approached Edmund Yunis, of the Dana Farber Cancer Institute, Boston, about the case, and they together determined that the woman was a tetragametic chimera who had probably resulted from separately fertilised XX zygotes, one with HLA haplotypes 1 and 3 and the other with haplotypes 2 and 4. She was very proud of the paper, Yunis told The Lancet, and deserves all the credit for it (N Engl J Med 2002; 346: 1545–52). “She even printed little diplomas and sent them to all the authors”, he said. “Getting it published was a sign of her perseverance.”

Most recently, Kruskall had turned her attention to creating universal red blood cells for transfusion, by removing A or B antigens with a green coffee bean enzyme and making the cells more O-like. She took the work from the bench to early clinical studies along with ZymeQuest, a company created to commercialise the idea.

In the USA, Kruskall was a nationally known figure on blood banking policy. She was “almost invariably a selection when national policies in transfusion were thought about by the federal government, researchers, and policy institutes”, Klein said. And she was a role model for many trainees, particularly because of the way she managed to balance her career with her family. One of those trainees was Beth Shaz, who said “she was a wonderful teacher, one that you always wanted to perform your best for.” In one demonstration of her organisational skills, she put a library of some 2000 papers onto Reference Manager so that she and her colleagues could access them. She used to say that “It’s easy to tackle the sand, hard to tackle the rocks”, said Lynne Uhl, of the division of laboratory and transfusion medicine at Beth Israel Deaconess Medical Center. “Make the primary focus the rocks rather than the sand.”

Kruskall, who was an expert flautist, is survived by her husband, Stephen Kruskall, whom she met in medical school; two daughters, Lauren and Gillian; a son, Peter; and a sister, Denni Day.

Ivan Oransky
ivan-oransky@erols.com
Ciprofloxacin-resistant gonorrhoea in South Africa

Surveillance to monitor susceptibility patterns of Neisseria gonorrhoeae is not a regular feature of the national sexually transmitted infection (STI) control programme in South Africa, and occurs mainly in the context of research settings. Ciprofloxacin is included in the syndromic management package for treatment of potential infection with N gonorrhoeae.

The Prince Cyril Zulu Centre for Communicable Diseases in Durban houses one of the largest STI clinics in the province. We have been testing the susceptibility of N gonorrhoeae isolates from this centre to ciprofloxacin since 1995. Up until 2002, all isolates were susceptible to ciprofloxacin (table). However, reports of treatment failures among patients with genital discharge who were following syndromic management appeared in November, 2003, and coincided with the appearance of ciprofloxacin-resistant N gonorrhoeae isolates (minimum inhibitory concentration [MIC] ≥ 1 mg/L) at a prevalence of 22%.1

The National Department of Health (DoH) in South Africa was informed and a meeting of the STI Treatment Advisory Group to the DoH was convened in January, 2004. The decision was to do susceptibility surveillance at a few selected clinics, including the Prince Cyril Zulu Centre.

259 isolates were obtained from this clinic and tested between January and March, 2003. The prevalence of ciprofloxacin-resistant isolates was 24% (table). This was again reported to the DoH and we expected that the recommendation of ciprofloxacin as first-line treatment for gonorrhoea would be changed. However, ciprofloxacin remained a first-line agent in the syndromic management package for genital discharge. The only adjustment to the guidelines was a change from retreatment with ciprofloxacin to ceftriaxone in patients whose disease failed to respond.

A further susceptibility survey at the same clinic was done in January, 2005. This study revealed that the prevalence of ciprofloxacin-resistant isolates among men presenting with male urethritis syndrome had risen to 42% (table).

Postponing effective treatment in an area with a high prevalence of HIV and N gonorrhoeae has obvious implications in terms of the transmission of both organisms. The initial inappropriate treatment resulted in preferential transmission of the resistant organisms. Non-responding gonococcal disease increases the duration of mucosal inflammation, which in turn increases the likelihood of transmission of HIV.2

The recommendation from WHO is to use drugs that are 95% effective as empirical treatment for the management of STIs.3 However, in KwaZulu Natal, the prevalence of ciprofloxacin-resistant N gonorrhoeae increased from 24% to 42% over a 12-month period, and continues to go unchecked.

We hope that this report expedites an urgent review of the syndromic management guidelines, with a view to replacing ciprofloxacin as first-line treatment in Durban, KwaZulu Natal, is long overdue.

We declare that we have no conflict of interest.

Prashini Moodley,* A Willem Sturm
sturm@ukzn.ac.za
Department of Medical Microbiology, Nelson R Mandela School of Medicine, University of KwaZulu Natal, Private Bag 7, Congella 4013, Durban, South Africa


Prenatal diagnosis

Allan Caine and coauthors (July 9, p 123)1 are to be congratulated on the collection and analysis of a substantial body of data on prenatal karyotypes in women having invasive testing because of a raised risk of Down’s syndrome. However, we have analysed the data in a different way and as a result disagree with their conclusion that all prenatal samples should have a full chromosome analysis.

As researchers at the centre that introduced rapid quantitative fluorescence (QF) PCR prenatal testing into the UK National Health Service,2 we are in complete agreement with Caine and colleagues’ recommendation that karyotyping in addition to QF-PCR testing should be available when clinically indicated—eg, on samples from pregnancies in which the fetus has structural abnormalities detected by ultrasonography, or in which maternal-cell contamination prevents an interpretable result from QF-PCR.

Caine and colleagues’ conclusion from the analysis of their karyotype data is that “substantial numbers of
liveborn children with hitherto preventable mental or physical handicaps” will be born if full karyotype analysis is no longer done on prenatal samples from women referred for increased risk of Down’s syndrome, in line with recent UK National Screening Committee (UKNSC) recommendations. We believe this conclusion to be exaggerated by the inclusion of a so-called “low to high risk” karyotype group in their final calculations of the number of babies likely to be born with congenital abnormalities. These karyotypes include de novo balanced rearrangements, which have a quoted risk of only 6% of phenotypic abnormality (3% above the background general population risk of abnormal phenotype); small marker chromosomes, many of which are benign; and trisomies undetected by rapid testing, many of which will miscarry before reaching term; this group might therefore be better defined as being of “unknown or unpredictable prognosis”.

Such findings cause many counselling difficulties when clinicians cannot advise parents what the outcome of the pregnancy will be, and present the parents with agonising decisions and distress, all of which could be avoided if only targeted testing for trisomies was done, as the UKNSC recommends. Studies that have ascertained the outcome of pregnancies after abnormal results of prenatal diagnosis (data missing from Caine and colleagues’ paper) show that some couples choose to terminate their pregnancies after such equivocal findings, yet many of these terminations may be unnecessary because it is most likely that the baby would be normal. These terminations could be avoided if testing were restricted to the exclusion of the three main autosomal trisomies (13, 18, and 21) by use of rapid aneuploidy detection. In addition, the discussion in Caine and colleagues’ paper takes no account of the likelihood that some fetuses with abnormal karyotypes such as trisomy 16 will miscarry before term.

If this “low to high risk” group is excluded from the data analysis, the total number of high-risk abnormal karyotypes found represents 0.1% of women in this referral group—a figure very much in line with other published data collections. The proportion of women who give birth to babies with congenital abnormalities is likely to be lower than this figure, owing to the spontaneous miscarriage of some abnormal pregnancies. Such a figure may be considered sufficiently low to justify targeted aneuploidy testing alone, to avoid the uncertainties caused by identifying results of unknown prognosis. In addition, targeted testing is in line with WHO criteria for screening programmes.

We therefore disagree with the conclusions of this paper and support the UKNSC recommendations. We declare that we have no conflict of interest.

*Caroline Mackie Ogilvie, Frances Flinter

caroline.ogilvie@genetics.kcl.ac.uk

Genetics Centre, Guy’s and St Thomas’ NHS Foundation Trust, St Thomas Street, London SE1 9RT, UK


Authors’ reply

The points raised by Caroline Mackie Ogilvie and Frances Flinter highlight a fundamental difference in the screening and diagnostic approach to the interpretation of prenatal cytogenetic outcome data. The debate started by our paper focuses on a patient’s reproductive choices following an invasive procedure with its own a-priori risk, and whether these choices should be based on the most comprehensive information available (karyotyping) or restricted to testing for what the screening is designed to detect—ie, trisomy 21 (QF-PCR or fluorescence in-situ hybridisation [FISH]).

Ogilvie and Flinter do not challenge our results or data, but they conclude that our inclusion of the “low to high risk” karyotypes after Down’s screening has inflated the number of phenotypically abnormal outcomes which might survive to term. Our choice of this wide risk definition was deliberate because we recognise that some karyotypes have uncertain prognoses and viabilities, but increasingly, chromosome abnormalities are being better characterised using advanced techniques, and their associated risks quantified more accurately. Ogilvie and Flinter assume that many of the undetected karyotypes will abort spontaneously, be detected by other methods (eg, ultrasound†), or either have no clinical significance or will be associated with outcomes that are difficult to predict antenatally. However, with the exception of Warburton’s study† and in the absence of substantial outcome data based on long-term clinical follow-ups of antenatally detected chromosome abnormalities, we stand by our data because they quantify the numbers and types of karyotypes that will be missed if the UK National Screening Committee’s (UKNSC) recommendations are implemented.

Ogilvie and Flinter use examples to illustrate their case—eg, they quote the 6% risk associated with antenatally ascertained de novo balanced translocations; but do not acknowledge that
new techniques that could soon be used in prenatal diagnoses are beginning to differentiate between genetically balanced and unbalanced translocations of this type. They also ignore the important clinical usefulness of prenatal ascertainment of familial translocations, whereby not only the proband but possibly other family members can be informed.

Ogilvie and Flinter also state that many small supernumerary marker chromosomes are "benign", but do not mention that about 30% of some classes of de novo antenatally ascertained supernumerary marker chromosomes are associated with a significant risk of an abnormal outcome and that others are invariably associated with profound developmental and intellectual handicap. The fundamental difference between these two approaches is that with karyotyping, the information can be considered by the patients and their clinicians and an informed choice made on the basis of the evidence, whereas with rapid aneuploidy testing this choice is withdrawn. We believe that this strategy could be severely detrimental for up to 100 neonates per year in the UK.

We believe that antenatal karyotyping must be retained, but with turnaround times near the 7 days achieved by some UK laboratories. Furthermore, we believe that it is ironic in an era of "patient empowerment" that the UKNSC is adopting a paternalistic "patient empowerment" that the National Screening Committee (UKNSC) recommendation that new antenatal screening programmes for Down’s syndrome should be done by rapid diagnostic tests and not full karyotyping. Experience in many countries has shown that these new rapid methods are of high diagnostic reliability for detecting non-mosaic trisomy 21, 18, and 13. But rapid tests are not designed to detect mosaics and structural chromosomal defects such as translocations and deletions. Prenatal diagnosis is more than just Down’s syndrome screening. Nearly all numeric and structural chromosomal defects up to 3–5 Mb can be detected by standard karyotyping; thus it has been regarded as the diagnostic "gold standard" for the past 30 years. The German Society of Human Genetics decided in a guideline in 1998 to do rapid tests in Germany only as an add-on, followed by standard karyotyping. Similar recommendations were given by the American College of Medical Genetics in 2000.

For the past 6 years, our laboratory has done a quality control assessment of rapid fluorescence in-situ hybridisation (FISH) and PCR diagnostics on behalf of the German Association of Human Cytogeneticists (ACC) and colleagues. FISH had clear limitations in prenatal diagnosis because 51 chromosomally aberrant cases had normal FISH results. These cases are the reasons for keeping the "gold standard" of karyotyping. Its replacement will decrease the diagnostic level in the UK (especially by restricting the rapid test to the search for trisomy 21, 18, and 13 without gonosomal aneuploidies).

If one of the leading countries in prenatal medicine decides to skip karyotyping in favour of a lean antenatal screening programme, this could have an effect on the diagnostic standards for other countries. We wish that the UKNSC decision had been taken after a broad discussion in the EU and worldwide.

There is another reason for being careful about reducing an existing high standard: in the past decade we have learnt a lot about new methods for analysing subtelomere sequences and microdeletions in chromosomes that are not covered by karyotyping. So obviously we have also reached the limits of standard karyotyping.

Will there be new diagnostic techniques available in the near future? Le Caignec and colleagues reanalysed 49 chromosomally normal fetuses with multiple malformations by array-based comparative genomic hybridisation, and identified genomic imbalances that could not be detected by karyotype analyses in more than 10% of cases. This technique will undoubtedly have a major impact on prenatal diagnosis.

We declare that we have no conflict of interest.

*Bernd Eiben, Ralf Glautz
eiben@eurogen.de

Institute of Clinical Genetics Nordrhein, Willy-Brandt-Platz 4, 45127 Essen, Germany


3 Eiben B, Clausen U, Held C, et al. Leitlinien (Guidelines) zum "pränatalen Schnelltest (FISH)" des Berufsverbandes Medizinische Genetik e.V.
Correspondence


One of the issues raised by Allan Caine and colleagues is whether public policy with a goal of eliminating fetuses with certain genetic characteristics amounts in practice to a new negative eugenics—eliminating the genetically undesirable from society. Even though the justification given for such testing and abortion is the right of procreative choice on the part of individual parents, the cumulative effect of those individual decisions is to wipe out certain groups of people within society on the basis of their genetic characteristics. This is a very disturbing issue to face, but we must do so and that requires us to use straightforward language.

However, Caine and colleagues write that “[r]eplacement of full karyotyping with rapid testing for trisomies 13, 18, and 21 after a positive screen for Down’s syndrome will result in substantial numbers of liveborn children with hitherto preventable mental or physical handicaps, and represents a substantial change in the outcome quality of prenatal testing offered to couples in the UK . . . Whether withdrawal of karyotyping from a large proportion of the UK population will have a substantial effect on the incidence of hitherto preventable morbidity and mortality in perinatal and liveborn babies remains to be established.”

“Liveborn children with preventable physical or mental handicaps” should read “liveborn children who should have been aborted.” It is not the handicaps that can be “prevented”, it is the children. The same is true of the statement about “hitherto preventable morbidity and mortality in perinatal and liveborn babies”. And maintaining “outcome quality” means letting as few as possible of these children slip through the screening net.

Prenatal research: a very sensitive field

In July, 2005, stem-cell biologists discussed a proposal to refrain from using the term “embryo” when referring to blastocysts, in order to limit the negative scrutiny induced by this emotive term.1

1 month later, this very fear was realised in the public uproar resulting from the discovery of 351 corpses of fetuses and stillborn babies in a hospital morgue in Paris, which were probably stored for research purposes instead of being cremated according to law.2 Polemics reached such an emotional level that further research on this “material” was threatened.

These two stories epitomise the influence of emotions on ethical debate, especially in the very sensitive field of prenatal medicine. Emotions reveal rules indispensable to social organisation. Morals or “prosocial emotions” function to regulate social behaviours, more often for the long-term interest of a social group than the short-term interests of the individual.1 Emotions arise in us when socially needed values we are attached to are threatened—ie, life protection in this context. Embarrassment, guilt, and shame remind scientists how much human life means to everyone, even in its most early form—and before it is considered so scientifically.

Public concern about such issues is more genuine and stronger when triggered by emotions, and the only way to raise a wider debate about it is through the attraction of media interest. Thus the debate, on leaving the scientific world, allows for many more points of view to emerge. These are often not consistent with scientists’ views, but they reflect legitimate preoccupations. As long as it concerns the whole society and its future (and has deep moral implications), information must go further than scientific knowledge and opinion. David Hume recognised the effect of passion on political questions in his Treatise of human nature.4

However, as in France,1 emotions can negatively seize a debate such that balanced reflection partly or totally disappears. In such situations, a scientist’s role is to provide reason and quietness, and not to hide nor fear scrutiny nor roar with the crowd (unless there is no ethical thought needed). An emotion is not ethical per se, but ethics arise with emotions.

We declare that we have no conflicts of interest.

1 Sebastien Tassy, Guillaume Gorinouc, Pierre Le Coz drsebastientassy@yahoo.fr
Centre for Medical Ethics, Pr Mattei, La Timone Hospital, 264 rue Saint-Pierre, 13385 Marseille cedex 5, France
Eradication of poliomyelitis

The excellent study by Q Sue Huang and colleagues (July 30, p 394) addresses crucial questions about global polio eradication. Several lessons are to be learned from it.

If two cases of vaccine-associated paralytic polio (VAPP) in a population of 3.7 million prompted the switch from live attenuated oral polio vaccine (OPV) to inactivated vaccine (IPV), what should countries with much larger numbers of cases do? India had 181 cases of VAPP in 1999. OPV must be stopped as early as possible after interrupting wild virus transmission, but will it be safe to discontinue OPV without vaccination coverage with IPV?

Huang and colleagues recommend doing studies on vaccine virus persistence in tropical developing countries where transmission is likely to be more intense. Whatever the outcome of such a study, stopping OPV without IPV coverage will be unsafe. Canada, USA, most European countries, and New Zealand have already switched from OPV to IPV and if the success of curtailing vaccine virus circulation was at least in part due to IPV, the same approach will be necessary in developing countries. Indeed, no alternative model is available for them to adopt.

The sewage isolates in New Zealand showed the persistence of vaccine viruses for 4 months after discontinuation of OPV, despite vaccination with IPV. In countries with high birth rates and population density, even such a relatively short duration of persistence might be sufficient for vaccine virus transmission to continue, especially in unvaccinated infants born after stopping OPV. If even one strain continues transmission, it could develop into a circulating vaccine-derived poliovirus (cVDPV), which can undo much of the hard work put in for eradication. The probability of this risk is unknown and may be low, but it is not zero. Without sewage surveillance, it will be detected only after it causes acute flaccid paralysis, which may be a year or more later. By then it would have spread widely and would require widespread vaccination to contain its circulation.

What could be done if this happens? Reintroduction of OPV will expose all newborn, immunologically naive birth cohorts to vaccine viruses. The consequence could be the seeding of more strains leading to cVDPV. A better solution could be to use IPV to interrupt transmission of cVDPV, should the need arise. But we do not know how exactly to do it, especially since campaigns will be necessary.

A better approach is to prevent, as much as possible, the transmission of vaccine viruses, for which high IPV coverage will be necessary before withdrawing OPV. In other words, it would be wiser to introduce IPV while OPV is still in use, and to withdraw OPV only after achieving high IPV coverage.

The source of the type 2 infection with 99.9% homology with Sabin virus, detected in a 10-month girl 19 months after discontinuing OPV, must have been importation. Therefore, no country should stop vaccination against polio while another country continues to use OPV. Although simultaneous stopping of OPV worldwide is ideal, its practicalities need to be worked out. As a deterrent against importation of vaccine viruses, high IPV coverage before stopping OPV would be ideal in developing countries.

I am a member of global, regional, and national committees concerned with immunisation and polio eradication, but opinions presented here are my own.

T J Jacob John
vr tjjohn@sancharnet.in
439 Civil Supplies Godown Lane, Vellore, Tamil Nadu 632 002, India

and changed the supplementary immunisation strategy from nationwide to districts at high risk of cross-border importation.

However, in October, 2003, a previously healthy 33-month-old Cameroonian girl who had received nine OPV doses was paralysed by wild poliovirus type 1 genetically linked to the virus circulating in northern Nigeria, putting an end to 4 polio-free years. In response, Cameroon has reverted to multiple annual rounds of nationwide supplementary immunisation activities. Although 15 other cases have been imported since then, the supplementary immunisation activities have succeeded in preventing a re-establishment of wild poliovirus transmission (ie, circulation of genetically related viruses for more than 6 months after initial importation).

In endemic countries,1 multiple supplementary immunisation activities are needed to increase population immunity levels and surpass the herd immunity threshold for polio. Competent national and international authorities should take appropriate measures to sustain the current political commitment to polio eradication which is indispensible to rid the world of the last pockets of the wild poliovirus.2 Failing in this would imply that the biggest public health programme in history, the Global Polio Eradication Initiative,3 will be derailed.

We declare that we have no conflict of interest.

*Charles Shey Wiysonge, Alfred K Njamnshi, Emmanuel Nomo, Muki S Shey
wiysonge@uctgsh1.uct.ac.za


Paul Webster’s report from India (July 30, p 359)1 makes gloomy reading. Type 1 poliomyelitis had become prevalent, so the usual trivalent vaccine (types 1, 2, and 3) was replaced by a monovalent (type 1) vaccine. Cases of type 3 infection then occurred, and the use of trivalent vaccine is to be restored.

Can we not learn from past experience, especially when analogous situations have arisen with other vaccines? It has long been known that the three types of poliovirus do not confer cross-protection, any more than the different types of influenza virus or rotavirus.2

With bacterial diseases, the same applies. Type 1/2 pertussis infection occurs in, and even kills, children vaccinated with type 1/3 organisms, and vice versa.3 A balanced response to all three antigens is necessary for immunity. Similarly, cholera vaccine must give a balanced response to antigens Inaba, Ogawa, and Bengal; typhoid vaccine likewise to the Vi, O, and H antigens.4 Polio vaccine must contain all three types of the virus.

The other important issue with polio vaccine is raised in the Comment from New Zealand (July 30, p 351):4 the choice between (live) oral poliovirus vaccine (OPV) and inactivated poliovirus vaccine (IPV). But the replacement of OPV with IPV should not be as controversial as Calman MacLennan and Jenny MacLennan imply. Not only is IPV safer than OPV; it is more immunogenic5 and associated with reduced delivery cost.6 Whereas two or three doses of enhanced-potency IPV gave excellent protection in India and Senegal, it took a mean of 10 doses of OPV per child, over a period of 8–9 years, to interrupt transmission of the virus in South America.7

The separate administration of OPV, alongside the injection of vaccines for other childhood diseases, can be obviated by the simple injection of a polyvalent vaccine (IPV, diphtheria, tetanus, pertussis, and Haemophilus influenzae type b) on a three-dose schedule starting at 3 months of age.8–11 I declare that I have no conflict of interest.

Noel W Preston
nw.preston.vi@connectfree.co.uk

Department of Error

MERIT study investigators. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. Lancet 2005; 365: 2091–97.—In table 1 of this Article (June 18), the median admission number per hospital (IQR) for control hospitals should be SB56 (2784–4946).

Efficacy and effectiveness of influenza vaccines in elderly people: a systematic review

T Jefferson, D Rivetti, A Rivetti, M Rudin, C Di Pietrantonj, V Demicheli

Summary

Background Influenza vaccination of elderly individuals is recommended worldwide. Our aim was to review the evidence of efficacy and effectiveness of influenza vaccines in individuals aged 65 years or older.

Methods We searched five electronic databases to December, 2004, in any language, for randomised (n=5), cohort (n=49), and case-control (n=10) studies, assessing efficacy against influenza (reduction in laboratory-confirmed cases) or effectiveness against influenza-like illness (reduction in symptomatic cases). We expressed vaccine efficacy or effectiveness as a proportion, using the formula VE=1–relative risk (RR) or VE*=1–odds ratio (OR). We analysed the following outcomes: influenza, influenza-like illness, hospital admissions, complications, and deaths.

Findings In homes for elderly individuals (with good vaccine match and high viral circulation) the effectiveness of vaccines against influenza-like illness was 23% (95% CI 6–36) and non-significant against influenza (RR 1·04, 0·43–2·51). Well matched vaccines prevented pneumonia (VE 46%, 30–58) and hospital admission (VE 45%, 16–64) for and deaths from influenza or pneumonia (VE 42%, 17–59), and reduced all-cause mortality (VE 60%, 23–79). In elderly individuals living in the community, vaccines were not significantly effective against influenza (RR 0·19, 0·02–2·01), influenza-like illness (RR 1·05, 0·58–1·89), or pneumonia (RR 0·88, 0·64–1·20). Well matched vaccines prevented hospital admission for influenza and pneumonia (VE 26%, 12–38) and all-cause mortality (VE 42%, 24–55). After adjustment for confounders, vaccine performance was improved for admissions to hospital for influenza or pneumonia (VE* 27%, 21–33), respiratory diseases (VE* 22%, 15–28), and cardiac disease (VE* 24%, 18–30), and for all-cause mortality (VE* 47%, 39–54).

Interpretation In long-term care facilities, where vaccination is most effective against complications, the aims of the vaccination campaign are fulfilled, at least in part. However, according to reliable evidence the usefulness of vaccines in the community is modest.

Introduction

Over the past four decades, vaccines have been used to reduce the effects of influenza in elderly individuals. In 2000, 40 of 51 developed or rapidly developing countries recommended vaccination for all individuals aged 60–65 or older,1 and, in 2003, 290 million doses of vaccine were distributed worldwide.2 According to Centres for Disease Control (CDC), the main aim of vaccination in elderly individuals is to reduce the risk of complications in those who are most vulnerable.3 As such, they define two high priority groups—individuals aged 65 years or older, and residents of nursing homes and long-term care facilities.

Two systematic reviews of the effects of influenza vaccines in elderly people have been published.3,4 The first5 was done more than a decade ago, and the second6 has several methodological weaknesses—namely, the exclusion of studies with denominators of less than 30 and pooling of studies of different design—and includes only 15 studies. Our aim was to identify and assess the comparative studies of the efficacy and effectiveness of influenza vaccines in individuals aged 65 years or older. This review is part of two Cochrane reviews that will also include evidence of safety of the vaccines studied and of their effectiveness in carers.7,8
Articles

Procedures
Two reviewers (TJ and DR) independently applied inclusion criteria to all identified and retrieved articles. Three reviewers (TJ, DR, and MR) extracted data from included studies on standard Cochrane Vaccines Field forms. The procedure was supervised and arbitrated by VD. Assessment of methodological quality for randomised controlled trials was done in accord with criteria detailed in the *Cochrane Reviewers’ Handbook*. We assessed studies according to method of randomisation, generation of the allocation sequence, allocation concealment, blinding, and follow-up. We assessed quality of non-randomised studies in relation to the presence of potential confounders, using the appropriate Newcastle-Ottawa Scales (NOS). We used quality at the analysis stage as a means of interpretation of the results. We assigned risk of bias categories on the basis of the number of NOS items judged inadequate in each study: low risk of bias (up to one inadequate item); medium risk of bias (up to three inadequate items); high risk of bias (more than three inadequate items); very high risk of bias (no description of methods). Details of quality assessment and summaries of all included studies are available from TJ.

We assessed the following outcomes: influenza confirmed by viral isolation, serology, or any other type of laboratory testing; influenza-like illness that arose in winter or epidemic periods; pneumonia; admission to hospital for complications associated with influenza-like illness or influenza; and all-cause mortality. We did not consider serological outcome data in the absence of symptoms, since our aim was to assess the evidence of public-health effect of vaccination.

Statistical analysis
We entered extracted data into Cochrane RevMan software. Aggregation of data was dependent on the sensitivity and consistency of definitions of exposure, populations, and outcomes used. When we identified studies as consistent, we did a meta-analysis within each design category.

We grouped reports first according to the setting of the study (community or long-term care facilities) and then by level of viral circulation and vaccine matching. We further stratified by co-administration of pneumococcal polysaccharide vaccine (PPV) and by different types of influenza vaccines. When a study reported data for more than one influenza season or for more than one setting, we considered these separately, creating separate data sets.

We calculated the statistic I² for every pooled estimate to assess the effect on statistical heterogeneity. I² can be interpreted as the proportion of total variation among effect estimates that is due to heterogeneity rather than sampling error, and it is intrinsically independent of the number of studies. When I² is less than 30% there is little concern about statistical heterogeneity. We used random effect models throughout to take account of the between-study variance in our findings.

When possible, we did a quantitative analysis adjusted for confounders if the cohort or case-control studies used the same methods of adjustment (logistic regression) for the same confounders. We constructed a comparison with effect sizes adjusted for the effects of possible known confounders and their standard error, which we derived from the reported CIs and did quantitative analysis with the inverse of the variance. Findings of one case-control study, reporting data stratified by risk factors for influenza, were included by use of the inverse variance, combining stratum-specific effect size and overall effect size.

We summarise efficacy (against influenza) and effectiveness (against influenza-like illness) estimates as relative risk (RR, 95% CI) or odds ratio (OR, 95% CI). Absolute vaccine efficacy (VE) is expressed as a proportion, using the formula VE=1–RR or VE*=1–OR whenever significant. When not significant, we report the relevant RR or OR.

To investigate the causes of heterogeneity we did a further analysis. To assess the effect of viral circulation and vaccine matching on overall heterogeneity, we calculated heterogeneity within each grouping and compared its sum with the overall heterogeneity. We then tested effect size from cohort studies done in long-term care facilities (where the data are more plentiful) stratified by methodological quality of the studies.

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

Results
We included data from 64 studies (96 datasets) in our analyses (figure 1). Half (n=48) the datasets reported A/H3N2 virus circulating, 4% (n=4) B viruses, 1% (n=1) A/H1N1, 1% (n=1) A/H2N2, and 7% (n=7) reported A/H3N2 and A/H1N1 circulating at the same time. The remaining 37% (n=35) of datasets did not provide sufficient information on circulating subtypes. All cohort studies compared the effects of vaccination against no vaccination. In randomised controlled trials, placebo was the comparison. Table 1 shows results of the quality assessment of included studies.

Table 2 shows our main findings. 29 cohort studies in long-term care facilities contributed data to 40 datasets (33 985 observations). The studies were focused, had a short follow-up (mean 3-5 months), and were fairly well resourced. They assessed the effects of vaccines in closed communities. The resident population is described in about half of the included datasets as predominantly aged older than 75 years, with multiple chronic pathologies and...
a high dependency level. However, breakdown of potential confounding factors (such as age, sex, smoking status, and underlying chronic disease) is rarely reported by vaccine exposure, making correction of confounding impossible.

Of the 40 datasets, 29 (6702 observations) were recorded during outbreaks or periods of high viral circulation. In 26, the influenza virus subtype is positively identified (A/H3N2 in 23). The focus of influenza circulation, vaccines did prevent hospital admission for pneumonia (68%), but the study by Deguchi and colleagues35 is at high risk of bias and was set in 301 nursing homes, comprising 22462 elderly individuals, during the non-epidemic 1998–99 season in Japan. The same study has a large weight in the analysis of effectiveness against deaths by influenza and pneumonia (table 2).

With respect to cohort studies of elderly individuals living in the community, we identified 20 (39 datasets). The studies contain more than 3 million observations, mainly obtained using data-linkage from insurance-reimbursement, hospital, or primary-care databases, and report data stratified or adjusted by risk factors and other potential confounders. These studies had a mean follow-up of 5 months. On the basis of this large body of data from different accommodation blocks were analysed. We noted no association (correlation coefficient 0.09) between vaccine coverage and attack rate of influenza-like illness (figure 2).

Efficacy of the vaccines against influenza was tested in only seven datasets (1941 observations) and was not significant (table 2). The effectiveness of the vaccines in preventing complications was assessed in 14 datasets (10 097 observations, with several datasets contributing to more than one outcome); well matched vaccines effectively prevented pneumonia (46%) and hospital admission for influenza or pneumonia (45%; eight datasets), whereas those of poor or unknown match did not (table 2).

Vaccination had a significant effect on the prevention of deaths due to influenza or pneumonia (figure 3), though in the presence of considerable heterogeneity in 20 datasets.31,24,27,38,41,44,48,55–57,59,62,71,73,78 The vaccine was effective if it was a good match, otherwise it was not (table 2). The latter finding is based on a small dataset, however. The effectiveness on all-cause mortality was assessed in only one small study and was significant (60%).

In 11 datasets (27 283 observations), the effects of influenza vaccines in 350 institutional facilities during low viral circulation were assessed. The vaccines prevented influenza-like illnesses (33%), but not influenza (table 2). We identified a few datasets that assessed effectiveness of vaccines in preventing complications. Four briefly-reported datasets from two studies done in situations of low viral circulation and poor vaccine matching report a combined effectiveness of 65% in preventing pneumonia. During periods of low viral circulation, vaccines did prevent hospital admission for pneumonia or influenza (68%), but the study by Deguchi and colleagues35 is at high risk of bias and was set in 301 nursing homes, comprising 22 462 elderly individuals, during the non-epidemic 1998–99 season in Japan. The same study has a large weight in the analysis of effectiveness against deaths by influenza and pneumonia (table 2).

The overall effectiveness of vaccines (VE) against influenza-like illness was 23% when vaccine matching was good, and not significantly different from no vaccination when matching was poor or unknown (table 2). Heterogeneity was high, even within the same influenza season and within the same institution when
### Table 2: Main results

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Vaccine matching</th>
<th>Outcome studied</th>
<th>Number observed</th>
<th>Result</th>
</tr>
</thead>
</table>
| **Cohort studies in long-term care facilities (influenza vaccine vs no vaccination)**
| **Studies done during outbreaks or high viral circulation periods**
| 20–22,24,27,31–33,38,41,44,46,48,49,51,55–59,67,71,73,76,78 | Poor or unknown | Influenza-like illness | 5963 | VE 23% (6–36) |
| 216,17 | Poor or unknown | Influenza-like illness | 919 | RR 0.77 (0.56–1.06) |
| 29 | Good | Pneumonia | 658 | RR 1.04 (0.43–2.51) |
| 2270 | Good | Pneumonia | 482 | VE 46% (30–58) |
| 825,26,60–62 | Poor or unknown | Hospital admission for influenza and pneumonia | 2027 | VE 45% (16–64) |
| 20,24,41,44,48,51,55–57,59,67,73,76,78 | Poor or unknown | Hospital admission for influenza and pneumonia | 814 | RR 0.64 (0.35–1.16) |
| 318,37 | Poor or unknown | Death from influenza and pneumonia | 124 | RR 1.27 (0.07–1.61) |
| 2270 | Good | Death from influenza and pneumonia | 6327 | VE 42% (17–59) |
| 249 | Good | Death from influenza and pneumonia | 1089 | RR 0.34 (0.11–1.10) |
| 13 | Poor or unknown | Death from all causes | 305 | VE 60% (23–79) |
| 742575 | Good (epidemic year) | Hospital admission for influenza and pneumonia | 727776 | VE 26% (12–38) |
| 66 | Poor or unknown | Hospital admission for influenza and pneumonia | 779934 | VE 28% (15–38) |
| 6127 | Good (epidemic year) | Death from all respiratory diseases | 433934 | RR 0.87 (0.67–1.22) |
| 246668 | Poor or unknown | Death from all causes | 426668 | RR 1.32 (1.25–1.39) |
| 2027 | Good | Death from all causes | 404759 | VE 42% (24–55) |
| 397197 | Good | Death from all causes | 68032 | VE 61% (3–84) |
| **Cohort studies in elderly individuals living in the community (influenza vaccine vs no vaccination)**
| Inactivated influenza vaccines in all elderly individuals living in the community**
| Influenza-like illness | 4904 | RR 0.89 (0.69–1.15) |
| 450,73 | Poor or unknown | Death from all causes | 43821 | VE 50% (37–60) |
| 231,71 | Poor or unknown | Death from all causes | 421 | RR 1.27 (0.02–2.12) |
| 820,24,55,59,67,78 | Good (epidemic year) | Hospital admission for influenza and pneumonia | 44823 | VE 61% (43–85) |
| 1624,38,41,44,48,55–57,59,67,73,76,78 | Good (epidemic year) | Hospital admission for influenza and pneumonia | 6127 | VE 42% (17–59) |
| 127 | Poor or unknown | Hospital admission for influenza and pneumonia | 124 | RR 0.87 (0.7–1.04) |
| 592 | Poor or unknown | Death from influenza and pneumonia | 592 | RR 0.23 (0.05–1.03) |
| 117 Good (outbreak) | Death from influenza and pneumonia | 1092 | RR 0.74 (0.53–1.04) |
| 723,35,50,67,73 | Poor or unknown | Death from all causes | 1074 | RR 0.89 (0.69–1.15) |
| 251479 | Poor or unknown | Death from all causes | 251479 | VE 26% (8–40) |
| 25483 | Poor or unknown | Hospital admission for influenza and pneumonia | 25483 | VE 41% (26–53) |
| 23 | Poor or unknown | Hospital admission for influenza and pneumonia | 23 | RR 0.17 (0.02–1.28) |
| 2270 | Poor or unknown | Hospital admission for influenza and pneumonia | 2270 | VE 26% (8–40) |
| 249 | Poor or unknown | Hospital admission for influenza and pneumonia | 249 | VE 26% (8–40) |
| 825,26,60–62 | Poor or unknown | Hospital admission for influenza and pneumonia | 825,26,60–62 | VE 26% (8–40) |
| 305 | Poor or unknown | Death from all causes | 305 | RR 0.87 (0.67–1.22) |
| 846,48,55–57,67,73,78 | Poor or unknown | Death from influenza and pneumonia | 846,48,55–57,67,73,78 | VE 26% (8–40) |
| 2027 | Poor or unknown | Death from influenza and pneumonia | 2027 | VE 26% (8–40) |
| 23 | Poor or unknown | Death from influenza and pneumonia | 23 | RR 0.17 (0.02–1.28) |
| 251479 | Poor or unknown | Hospital admission for influenza and pneumonia | 251479 | VE 26% (8–40) |
| 4958 | Poor or unknown | Hospital admission for influenza and pneumonia | 4958 | VE 32% (14–46) |
| **Case-control studies**
| Inactivated influenza vaccines in elderly individuals living in the community**
| 30729 | Poor or unknown | Influenza-like illness | 4904 | RR 0.89 (0.69–1.15) |
| 30729 | Good (no epidemic year) | Death from all causes | 101691 | VE 50% (37–60) |
| 30729 | Poor or unknown | Death from all causes | 43821 | VE 50% (37–60) |
| 30729 | Poor or unknown | Death from all causes | 43821 | VE 50% (37–60) |
| 30729 | Poor or unknown | Death from all causes | 43821 | VE 50% (37–60) |
| **Randomised controlled trials (inactivated vaccines vs placebo)**
| Healthy | Good (outbreak) | Influenza-like illness | 2047 | VE 43% (21–58) |
| 397197 | Poor or unknown | Influenza | 2217 | VE 58% (34–73) |

Evidence, we divided our analysis into five separate comparisons (table 2).

Our first comparison showed that in elderly individuals living in the community (>1 million observations in 18 datasets from 15 studies), inactivated influenza vaccines were not effective against influenza-like illness, influenza, or pneumonia. No comparison provided enough data for stratification by viral circulation and vaccine matching. Eight datasets (779 934 observations; 727 776 in six datasets with good vaccine matching) addressed vaccine effectiveness against hospital admissions for influenza or pneumonia; well matched vaccines prevented hospital admissions for these illnesses (26%), but not for cardiac disease.60–62 Death from respiratory disease was not significantly affected.
In the second comparison, we assessed the effectiveness of inactivated influenza vaccines in elderly individuals living in the community and at risk of complications associated with influenza. Seven datasets from six studies were relevant. The only significant effect was that for deaths from all causes (table 2).

In the third comparison, we looked at the effectiveness of inactivated influenza vaccines in all healthy elderly individuals living in the community. In this stratum, six studies with seven datasets contributed several hundred thousand observations, but most outcomes were only assessed by one study. The only notable results are the vaccines effectiveness in preventing hospital admission for influenza or pneumonia (50%) and a lack of effect on all-cause mortality (table 2).

In the fourth comparison, we assessed the effectiveness of inactivated influenza vaccines in all elderly individuals living in the community after adjustment for confounders (>1 million observations from 19 datasets from seven studies). The datasets were from several consecutive influenza seasons. Most of the studies included in this analysis used data linkage and adjusted their OR calculations to allow for the effect of confounding of several variables (sex, age, smoking, comorbidities). The vaccines significantly affected all-cause mortality (47%), and hospital admission for influenza or pneumonia (27%; figure 4), respiratory diseases (22%), and cardiac disease (24%).

Finally, we assessed the effectiveness of virosomal influenza vaccines in elderly individuals living in the community. Such vaccines had no effect on all-cause mortality, but, with a vaccine with a good match, prevented influenza-like illnesses and hospital admission during a year of low viral circulation. Addition of PPV did not significantly improve the effectiveness of influenza vaccines.

We identified ten case-control studies (12 datasets). Six datasets from five studies assessed the effects of inactivated influenza vaccines on elderly individuals living in the community, four datasets from three studies looked at the co-administration of inactivated influenza with PPV in elderly individuals living in the community, and one study assessed the effect of influenza vaccine with adjuvant PPV on institutionalised elderly individuals. Since all datasets adjusted their ORs for likely confounding factors, we structured our analysis on five strata, further subdividing each analysis by viral circulation and vaccine matching whenever possible.

Before adjustment, inactivated influenza vaccines were associated with an increased risk of admission for any respiratory disease, and did not prevent hospital admission for influenza and pneumonia in elderly individuals living in the community, or affect mortality from influenza and pneumonia, though this conclusion is based on a dataset of 1092 observations (table 2). After adjustment, however, the vaccines did reduce the risk of...
death from influenza and pneumonia (26%), and prevent admission for influenza and pneumonia and for all respiratory diseases (29%). Similarly, before adjustment, inactivated influenza and concomitant PPV in individuals living in the community did not prevent hospital admission for influenza and pneumonia, whereas after adjustment it did.40,65,66

We identified five randomised controlled trials published over four decades and including just over 5000 observations. Given their heterogeneous nature by vaccines tested (monovalent, trivalent, live, or inactivated aerosol vaccines), setting, follow-up, and outcome definition, no firm conclusions can be drawn from this body of evidence. Two trials had adequate randomisation and allocation concealment,19,45 and one trial had adequate measures to prevent attrition bias.72 The results of the most recent trial19 are difficult to interpret because of the presence of selection bias. Based on the results of a meta-analysis of the two trials, inactivated vaccines were more effective than placebo against influenza-like disease

<table>
<thead>
<tr>
<th>Subcategory and study</th>
<th>Treatment (n/N)</th>
<th>Control (n/N)</th>
<th>RR (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outbreak and vaccine matching circulating strains</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fevery, 197543</td>
<td>3/154</td>
<td>1/63</td>
<td>1.23 (0.13–11.58)</td>
<td>1.94</td>
</tr>
<tr>
<td>Horman, 198648</td>
<td>5/100</td>
<td>3/59</td>
<td>0.98 (0.24–3.97)</td>
<td>4.62</td>
</tr>
<tr>
<td>Saah 2, 198673</td>
<td>3/244</td>
<td>8/214</td>
<td>0.33 (0.09–1.22)</td>
<td>5.12</td>
</tr>
<tr>
<td>Fyson 2, 198339</td>
<td>4/121</td>
<td>5/224</td>
<td>0.56 (0.15–2.06)</td>
<td>5.19</td>
</tr>
<tr>
<td>Goodman, 198244</td>
<td>0/36</td>
<td>9/84</td>
<td>0.12 (0.01–2.02)</td>
<td>1.25</td>
</tr>
<tr>
<td>Patriarca 1, 198547</td>
<td>6/548</td>
<td>21/470</td>
<td>0.25 (0.10–0.60)</td>
<td>9.43</td>
</tr>
<tr>
<td>Strassburg, 198646</td>
<td>4/85</td>
<td>3/39</td>
<td>0.39 (0.10–1.59)</td>
<td>4.55</td>
</tr>
<tr>
<td>Carter 1, 199045</td>
<td>2/95</td>
<td>0/46</td>
<td>2.42 (0.12–49.40)</td>
<td>1.10</td>
</tr>
<tr>
<td>Carter 2, 199044</td>
<td>0/30</td>
<td>1/55</td>
<td>0.60 (0.01–13.44)</td>
<td>1.00</td>
</tr>
<tr>
<td>Carter 3, 199044</td>
<td>3/332</td>
<td>2/126</td>
<td>0.57 (0.10–3.37)</td>
<td>2.99</td>
</tr>
<tr>
<td>Medleyjohn 1, 198775</td>
<td>1/36</td>
<td>3/32</td>
<td>0.18 (0.02–1.58)</td>
<td>2.02</td>
</tr>
<tr>
<td>Taylor, 199244</td>
<td>0/65</td>
<td>1/52</td>
<td>0.38 (0.02–9.20)</td>
<td>0.99</td>
</tr>
<tr>
<td>Morens, 199547</td>
<td>6/36</td>
<td>0/3</td>
<td>1.41 (0.10–20.60)</td>
<td>1.38</td>
</tr>
<tr>
<td>Muto, 200145</td>
<td>60/1726</td>
<td>28/623</td>
<td>0.77 (0.50–1.20)</td>
<td>21.58</td>
</tr>
<tr>
<td>Murayama, 199944</td>
<td>0/60</td>
<td>1/68</td>
<td>0.38 (0.02–9.09)</td>
<td>0.99</td>
</tr>
<tr>
<td>Subtotal</td>
<td>100/3884</td>
<td>86/234</td>
<td>0.58 (0.41–0.83)</td>
<td>65.30</td>
</tr>
<tr>
<td>Test for heterogeneity: χ²=15.62, p=0.41, I²=4.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z=3.02, p=0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Outbreak and matching absent or unknown | | | | |
| Ruden, 197444 | 2/204 | 13/192 | 0.14 (0.03–0.63) | 4.39 |
| Saah 1, 198673 | 2/219 | 12/234 | 0.18 (0.04–0.79) | 4.14 |
| Arroyo, 198444 | 2/266 | 6/94 | 0.81 (0.04–14.74) | 1.18 |
| Subtotal | 9/561 | 31/258 | 0.34 (0.11–10.2) | 13.38 |
| Test for heterogeneity: χ²=4.92, p=0.18, I²=39.0% | | | | |
| Test for overall effect: Z=1.92, p=0.05 | | | | |

| No outbreak and vaccine matching circulating strains | | | | |
| Patriarca 2, 198547 | 2/339 | 4/119 | 0.18 (0.03–0.95) | 3.30 |
| Carniti, 199444 | 2/169 | 1/73 | 0.86 (0.08–9.38) | 1.73 |
| Deguchi, 200145 | 1/10736 | 5/11723 | 0.22 (0.03–1.37) | 2.11 |
| Subtotal | 5/11247 | 10/11915 | 0.27 (0.09–1.87) | 7.13 |
| Test for heterogeneity: χ²=1.20, p=0.55, I²=0% | | | | |
| Test for overall effect: Z=2.20, p=0.03 | | | | |

| No outbreak and matching absent or unknown | | | | |
| Howells 1, 197544 | 1/134 | 15/356 | 0.18 (0.02–1.33) | 2.37 |
| Howells 2, 197544 | 3/123 | 22/267 | 0.30 (0.09–0.97) | 6.08 |
| Howells 3, 197544 | 0/181 | 11/287 | 0.07 (0.00–1.15) | 1.25 |
| Saah 3, 198673 | 3/225 | 5/226 | 0.60 (0.15–2.49) | 4.48 |
| Subtotal | 7/665 | 53/1136 | 0.30 (0.14–0.67) | 14.18 |
| Test for heterogeneity: χ²=2.44, p=0.63, I²=0% | | | | |
| Test for overall effect: Z=2.92, p=0.003 | | | | |

| Total | 121/16357 | 180/15822 | 0.46 (0.33–0.63) | 100.00 |
| Total events: 121 (treatment), 180 (Control) | | | | |
| Test for heterogeneity: χ²=29.31, p=0.30, I²=11.3% | | | | |
| Test for overall effect: Z=4.79, p<0.0001 | | | | |

Figure 3: Influenza vaccines compared with no vaccination for prevention of deaths caused by influenza or pneumonia in residents of long-term care facilities by level of viral circulation and quality of vaccine matching
Numbers after names of authors indicate different datasets.
in conditions of high viral circulation among elderly individuals living in the community (43%). The vaccines were also 58% effective against influenza (table 2).

Of the 15 main comparisons with 61 outcome combinations, we noted in a subsequent analysis that seven comparisons with 20 outcome combinations had an I² of greater than 30%, and that the heterogeneity of these studies could be explained by grouping by viral circulation and vaccine matching (table 3).

Discussion

Our findings show that, according to reliable evidence, the effectiveness of trivalent inactivated influenza vaccines in elderly individuals is modest, irrespective of setting, outcome, population, and study design.

In view of the known variability of incidence and effect of influenza, we constructed a large number of comparisons and strata to reduce to a minimum possible heterogeneity between studies and to aid comparability. Despite our attempts we noted significant residual between-studies heterogeneity that could be explained only in part by different study designs, methodological quality, settings, viral circulation, vaccine types and matching, age, population types, and risk factors. We think the residual heterogeneity could be the result of the unpredictable nature of the spread of influenza and influenza-like illness and the bias caused by the non-randomised nature of our evidence base. The findings of the cohort studies that we included are likely to have been affected to a varying degree by selection bias; differential uptake of influenza vaccines is linked to several factors (anxiety over unwanted effects, disease threat perception, societal and economic conditions, education, health status) and hence to outcome. Indeed, one cohort study,43 had real difficulties in achieving high coverage in those most at need. Differential vaccine uptake and the resulting selection bias is the most likely explanation for the high effectiveness of influenza vaccines in preventing deaths from all causes. A further example of the potential effect of such bias is the apparently counterintuitive effectiveness of the vaccines in elderly individuals living in the community. In this population, the vaccines are apparently ineffective in the prevention of influenza, influenza-like illness, pneumonia, hospital admissions, or deaths from any respiratory disease, but are effective in the prevention of hospital admission for influenza and pneumonia and in the prevention of deaths from all causes. That such differences are the result of a baseline imbalance in health status and other systematic differences in the two groups of participants cannot be discounted. Evidence from randomised controlled trials, in which bias is reduced to a
non-recipients. The difficulties of achieving good systematic differences between vaccine recipients and however, is modest, irrespective of adjustment for possibility.

suggests that control through vaccination is a greater effect on its complications. This finding effect on cases of influenza-like illness, but have a gradient of effectiveness, in which vaccines have little effects as a result of study design, it is possible to detect remaining heterogeneity and an overestimation of the with similar viral exposure and risk levels. Despite a more consistent than that in the community: older, several locations, the resident population is usually linked to age, sex, and health status, and a low effect on influenza-like illness throughout our studies, reporting observations from a few seasons, but that it should be taken on the basis of all available evidence. The conclusions drawn from studies done in individuals who live in long-term care facilities are different to those drawn from studies in individuals who live in the community. Whereas studies done in residents of care homes often indicate the inevitably improvised nature of efforts to study the effect of vaccines during an epidemic often concurrently in several locations, the resident population is usually more consistent than that in the community: older, with similar viral exposure and risk levels. Despite a remaining heterogeneity and an overestimation of the effects as a result of study design, it is possible to detect a gradient of effectiveness, in which vaccines have little effect on cases of influenza-like illness, but have greater effect on its complications. This finding suggests that control through vaccination is a possibility.

The effectiveness of vaccines in the community, however, is modest, irrespective of adjustment for systematic differences between vaccine recipients and non-recipients. The difficulties of achieving good coverage in those who most need it, or the diluting effect on vaccines for influenza of other agents circulating in the community (causing influenza-like illness, clinically indistinguishable from influenza), might be to blame. We noted empirical proof of both, with differential vaccine uptake among the same population linked to age, sex, and health status, and a low effect on influenza-like illness throughout our datasets, even in periods of supposedly high influenza viral circulation when the proportion of cases of influenza-like illness caused by influenza and the possible benefits of vaccination are highest.

On the basis of these observations, we believe efforts should be concentrated on achieving high vaccination coverage in long-term care facilities coupled with a systematic assessment of the effect of such a policy. One possible way to improve this strategy might involve the vaccination of carers in an effort to reduce transmission. The effect of vaccination of high risk groups should also be further assessed.

Finally, investment in the development of better vaccines than available at present should be linked to better knowledge of the causes and patterns of influenza-like illnesses in different communities. This partnership could lead to the inception of a more comprehensive and perhaps more effective strategy for the control of acute respiratory infections, relying on several preventive interventions that take into account the multi-agent nature of influenza-like illness and its context (such as personal hygiene, and provision of electricity and adequate food, water, and sanitation).
References


33 D’Alessio DJ, Cox PM Jr, Dick EC. Failure of inactivated influenza vaccine to protect an aged population. JAMA 1969; 210: 485–89.


Incidence of adamantane resistance among influenza A (H3N2) viruses isolated worldwide from 1994 to 2005: a cause for concern

Rick A Bright, Marie-jo Medina, Xiyan Xu, Gilda Perez-Oronoz, Teresa R Wallis, Xiaohong M Davis, Laura Povinelli, Nancy J Cox, Alexander I Klimov

Summary
Background Adamantanes have been used to treat influenza A virus infections for many years. Studies have shown a low incidence of resistance to these drugs among circulating influenza viruses; however, their use is rising worldwide and drug resistance has been reported among influenza A (H5N1) viruses isolated from poultry and human beings in Asia. We sought to assess adamantane resistance among influenza A viruses isolated during the past decade from countries participating in WHO’s global influenza surveillance network.

Methods We analysed data for influenza field isolates that were obtained worldwide and submitted to the WHO Collaborating Center for Influenza at the US Centers for Disease Control and Prevention between Oct 1, 1994, and Mar 31, 2005. We used pyrosequencing, confirmatory sequence analysis, and phenotypic testing to detect drug resistance among circulating influenza A H3N2 (n=6524), H1N1 (n=589), and H1N2 (n=83) viruses.

Findings More than 7000 influenza A field isolates were screened for specific aminoacid substitutions in the M2 gene known to confer drug resistance. During the decade of surveillance a significant increase in drug resistance was noted, from 0·4% in 1994–1995 to 12·3% in 2003–2004. This increase in the proportion of resistant viruses was weighted heavily by those obtained from Asia with 61% of resistant viruses isolated since 2003 being from people in Asia.

Interpretation Our data raise concerns about the appropriate use of adamantanes and draw attention to the importance of tracking the emergence and spread of drug-resistant influenza A viruses.

Introduction
Adamantane derivatives, such as amantadine and rimantadine, have been used successfully for the prevention and treatment of influenza A virus infection for more than 30 years.1–3 These drugs, known as M2 channel blockers, inhibit influenza A virus replication by blocking the M2 protein ion channel thereby preventing fusion of the virus and host-cell membranes and the release of viral RNA into the cytoplasm of infected cells.4 The prophylactic effect of these drugs varies between 80% and 90%, and the drugs can reduce the duration of illness by about 1·5 days if given within 48 h of infection.1–7

Human and animal studies have shown the frequent occurrence of amantadine-resistant influenza viruses after exposure to the drug, and drug-resistant viruses can be transmitted from one person to another without apparent loss of pathogenicity.19–21 Additionally, complete cross-resistance between amantadine and rimantadine has been shown.1 The genetic basis for resistance to these drugs has been well characterised and is associated with an aminoacid substitution at position 26, 27, 30, 31, or 34 in the transmembrane region of the M2 protein.12,13 Most drug-resistant influenza viruses contain one of these aminoacid changes, but variants with dual mutations have also been described.19 Resistance to adamantanes has been reported predominately in people in semi-closed settings (eg, nursing home facilities, paediatric wards, and family households) where antiviral treatment was used.8,12–20

Since 1991, drug-susceptibility surveillance has been undertaken routinely in the characterisation of influenza virus isolates submitted to the WHO Collaborating Center for Influenza at the US Centers for Disease Control and Prevention. Previous surveillance studies have identified a low incidence of resistance to amantadine and rimantadine (approximately 1%) among circulating influenza viruses. However, 10 years have elapsed since the last comprehensive global study of resistance to these drugs was published.21–24 Furthermore, influenza A (H5N1) viruses isolated from both human beings and avian sources in southeast Asia since 2003 have an S31N aminoacid substitution in the M2 protein and, thus, are resistant to amantadine and rimantadine.25–28 In this study we report results of a surveillance study for resistance to adamantanes among circulating influenza viruses collected worldwide between Oct 1, 1994, and Mar 31, 2005.

Methods
Viruses Worldwide data for influenza field isolates that were obtained and submitted to the WHO Collaborating Center for Influenza at the US Centers for Disease

Lancet 2005; 366: 1175–81
Published online September 22, 2005
DOI:10.1016/S0140-6736(05)67382-2
See Comment page 1139
Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Influenza Branch, Atlanta, GA, USA
(R A Bright PhD, M-j Medina MS, X Xu MD, G Perez-Oronoz MSPH, T R Wallis MS, X M Davis PhD, N J Cox PhD, A I Klimov PhD); and Wisconsin State Laboratory of Hygiene, Madison, WI, USA (L Povinelli PhD)
Correspondence to: Dr Rick Bright rbright@cdc.gov
Control and Prevention (CDC) between Oct 1, 1994, and Mar 31, 2005, were included in this study. Viruses were identified from individuals of various age, sex, and geographic groups around the world without any demographic specification or bias from the CDC. The selection of samples submitted or screened for drug resistance was random and therefore there was no intentional or systematic bias in the sample selection process. The sampling method remained consistent throughout the study period and only the number of isolates submitted in any given year by any given country varied. In total, 6524 H3N2, 589 H1N1, and 83 H1N2 viruses were screened for a resistant genotype to amantadine and rimantadine. Each year presented represents the 12-month period from the previous October to the following September to include an entire influenza season in the northern hemisphere (eg, 1995 includes isolates gathered from Oct 1, 1994, to Sept 30, 1995).

Procedures were propagated in embryonated hen eggs or Madin-Darby Canine Kidney (MDCK) cells. Virus haemagglutinin subtypes were identified by a haemagglutination inhibition assay and neuraminidase subtypes were identified by a neuraminidase-specific PCR assay.27 We extracted viral RNA from 100 µL of virus culture using a total nucleic acid extraction kit with a MagNaPure instrument (Roche Diagnostics, Mannheim, Germany) according to the manufacturer’s protocol. Reverse-transcription-PCR (RT-PCR) for cDNA synthesis and PCR for DNA amplification was undertaken in a 96-well PCR plate with a one-step RT-PCR kit (Qiagen, Hilden, Germany) and 10 mM amplification primers, M2F-5’-CAGATGCAGCGATTCAGTG and MR-5’-Biotin-AGTA GAAACAAGGTAGTTTTTTACTC, in a 25 µL volume containing 2 µL of total viral RNA from a subset of viruses. cDNA products were purified with the QiAquick PCR purification procedure (Qiagen, Hilden, Germany). We did sequencing reactions with the ABI BigDye 3.1 Terminator Cycle Sequencing kit (Applied Biosystems, Warrington, UK) with primers F1, F584, R656, and R1027; products were resolved on an ABI 3100 Auto Sequencer (Applied Biosystems, Warrington, UK). DNA sequence analysis was done with DNASTar programs. Viruses were screened for a resistant genotype to adamantane and rimantadine within the M2 protein. Sensitive and resistant viruses with known M2 gene sequences were used as controls.

We ascertained biological susceptibility to rimantadine for a subset of 22 viruses (investigators were unaware of the identity of the viruses) to validate drug-resistant phenotypes to genotypes reported in this study. In brief, monolayers of MDCK cells in a 24-well tissue culture plate were pretreated with 0, 0.2, 2.0, or 20 µg/mL rimantadine in 100 µL of Dulbecco’s modified eagle medium supplemented with 0.2% bovine serum albumin, 25 mM HEPES, penicillin, and streptomycin for 30 min at 37°C, plus 5% CO2. Viruses were diluted in a ratio of 1:2 or 1:20 in 100 µL of culture medium containing 2 µg/mL TPCK-Trypsin and added to cell monolayers. Plates were then wrapped in plastic film and centrifuged at 700 g at ambient temperature for 45 min. We then removed the plastic and added 300 µL of culture medium containing the respective concentration of rimantadine to each well. Plates were incubated at 37°C with 5% CO2 for 36 h. We assessed virus replication by measuring haemagglutinin titres of the supernatant, which were represented as the inverse of the highest dilution of virus agglutinating a 0-1% suspension of turkey red blood cells. Sensitive and
resistant viruses with known M2 gene sequences were used as controls.

**Statistical analysis**

We used poisson regressions to test for a significant increase in rates of resistance across the study period and to test the trend of the overall rate of resistance for all countries combined. A p value of less than 0·05 was deemed statistically significant and 95% CIs were determined for resistance rates to account for varying year-to-year sample size.

**Results**

In all, 6524 influenza A (H3N2) viruses were screened for specific mutations known to correlate with resistance to amantadine and rimantadine (table 1). All M2 sequences were typical for human influenza viruses and different from sequences known for avian or swine viruses. Overall, 392 (6%) H3N2 viruses contained an aminoacid substitution in the M2 protein that correlates with drug resistance. Additionally, an RFLP strategy known to identify specific mutations of the M2 gene that correlate with resistance was used for a blinded subset of viruses.

Data from both confirmation methods correlated 100% with resistance was used for a blinded subset of viruses. Statistical analysis showed a substantially rising percentage of drug-resistant H3N2 viruses isolated from the US and specific countries in Asia, including China and Hong Kong, Taiwan, and South Korea (table 1, figure). During 1995–2004, rates increased significantly for China (p < 0·0001), Hong Kong (p < 0·0001), South Korea (p < 0·0001), and all Asia combined (p < 0·0001). The trend for resistance phenotypes of a subset of 22 viruses to further validate the pyrosequencing method. We assessed viruses for their ability to replicate in MDCK cells in the absence or presence of rimantadine. Phenotypic results from the biological assay of both sensitive and resistant viruses correlated 100% with virus genotypes identified by pyrosequencing (table 2).

Geographic analysis showed a substantially rising percentage of drug-resistant H3N2 viruses isolated from the US and specific countries in Asia, including China and Hong Kong, Taiwan, and South Korea (table 1, figure). During 1995–2004, rates increased significantly for China (p < 0·0001), Hong Kong (p < 0·0001), South Korea (p < 0·0004), and all Asia combined (p < 0·0001). For Japan, however, the overall rate change for this period was not significant. For the USA, the rate change during 1995–2005 was significant (p < 0·0001). The trend for overall data showed a significant increase (p < 0·0001). Significant rate changes were also noted for a model with data for between 1995 and 2004 from five important countries or regions: China, Hong Kong, Japan, South Korea, and the USA (p < 0·0001), and rates among these countries were also significantly different (p < 0·0001).

There were substantial increases (spikes) in rates for China, Hong Kong, South Korea, USA, and overall in

### Table 1: Frequency of resistant H3N2 viruses by geographic origin and year of isolation

<table>
<thead>
<tr>
<th>Year</th>
<th>North America</th>
<th>Asia</th>
<th>Africa*</th>
<th>Europe†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>1996</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>1997</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>1998</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>1999</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>2000</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>2001</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>2002</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>2003</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>2004</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
<tr>
<td>2005</td>
<td>0/04 0/04 0/03 0/04 1/01 (0·0%)</td>
<td>0/04 2/65 (3·1%)</td>
<td>0/04 0/04 0/05 1/00 (10·0%)</td>
<td>0/06 0/05 0/8</td>
</tr>
</tbody>
</table>

Data are number resistant/number tested and percentage resistant when greater than 0. *Egypt, Madagascar, Senegal, South Africa. †Bangladesh, Guam, Kyrgyzstan, Macau, Malaysia, Nepal, Philippines, Oman, Qatar, Saudi Arabia, Singapore, Syria, Thailand, Turkey, Vietnam. ||Blagovia, Czechoslovakia, Cyprus, France, Germany, Greece, Hungary, Italy, Lithuania, Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden, Switzerland, Ukraine, UK. §Australia, Fiji, New Caledonia, New Zealand, Solomon Island. ||Argentina, Bahamas, Bolivia, Brazil, Chile, Columbia, Costa Rica, Cuba, Ecuador, El Salvador, French Guyana, Guadeloupe, Guatemala, Jamaica, Martinique, Nicaragua, Panama, Paraguay, Peru, Trinidad, Uruguay, Venezuela.
the past 2–3 years (table 1, figure). Trends of rate changes were not significant before the spikes occurred, which indicate the existence of such spikes and their significant effects on the overall trends (table 3). For example, the frequency of resistant viruses began increasing in 2000 in China, with a significant spike between 2002 and 2003. This trend of having spikes was followed in 2003 by an increasing frequency in Taiwan, Hong Kong, and South Korea. By comparison, the frequency of resistant viruses obtained from the USA was consistently near 2% between 1999 and 2004 (table 1, figure). However, the frequency of resistant viruses gathered in the USA during the first 6 months of the 2005 influenza season increased significantly to 15%. One or more drug-resistant virus was identified in 2004 in 11 other countries: Mexico (5/26, 19%), Canada (6/20, 30%), Argentina (2/65, 3%), Japan (2/46, 4%), Thailand (2/64, 3%), France (1/14, 7%), Italy (1/7, 14%), Russia (2/25, 8%), Paraguay (1/32, 3%), Brazil (1/14, 7%), and Peru (12/39, 31%; webtable). Detection of drug-resistant viruses from many of these countries was rare before 2004.

In the USA, drug-resistant viruses were obtained from individuals of various age, sex, and geographic groups. Since 2004, case histories for 92 patients with drug-resistant isolates were examined to determine use of amantadine or rimantadine before virus collection. Of those, travel histories of 18 patients were examined to ascertain recent travel to specific countries with high frequencies of drug-resistant influenza viruses. We showed that two nursing home residents had been on antiviral therapy for several days before virus collection, two patients had travelled to Asia immediately before onset of illness, and two siblings had a grandparent who returned home ill from a trip to the Philippines. None of the other cases had documented adamantane treatment or a history of recent travel to Asia.

From 1998 to 2004, 589 influenza A (H1N1) viruses were obtained worldwide and screened for aminoacid substitutions in the M2 gene associated with drug resistance to adamantanes. Of these viruses, two (0·3%) were identified with resistance-conferring substitutions (V27A, collected in the UK, and G34E, collected in China). Additionally, of 83 influenza A (H1N2) viruses screened, one that was isolated in the USA contained an A30T substitution associated with adamantane resistance.

Discussion
This study reveals an alarming increase in the incidence of amantadine-resistant and rimantadine-resistant H3N2 influenza A viruses over the past decade. Our study, which assessed more than 7000 influenza A viruses obtained worldwide, is the largest and most comprehensive report on adamantane resistance to date. The last major surveillance study by Ziegler and co-workers23 showed a drug-resistance frequency of 0·3%...
among H1N1 and H3N2 viruses and identified only 16 drug-resistant viruses internationally over a 4-year period. This low reported incidence accords with findings from previous studies.21,22,24 These reports had shown that isolation of drug-resistant influenza viruses was a rare event, with a frequency of 1–2%.

We have shown an escalating trend in circulating drug-resistant viruses in recent years. This trend was first observed among viruses isolated in Asia starting from 2000 and then in other regions of the world, including North America in 2004. Viruses collected in 2004 from South Korea, Taiwan, Hong Kong, and China show drug-resistance frequencies of 15%, 23%, 70%, and 74%, respectively. The first 6 months (October, 2004, to March, 2005) of the 2005 influenza season in the USA have already shown a significant increase in the incidence of adamantane resistance from 2% in 2004 to 15% in 2005 among isolates obtained from a comparable geographic and demographic distribution. Of particular note from the 10-year study period is the fact that of all resistant viruses, more than 84% were identified since the 2003 influenza season, indicating a distinct spike in drug resistance among circulating H3N2 influenza viruses throughout the world. This trend was recorded with no fundamental changes in the surveillance methods over the study period other than an expansion in the number of viruses being submitted over time.

Up to 30% of individuals who receive amantadine or rimantadine for the treatment of influenza virus infection can excrete viruses resistant to these drugs.8,10–17 Clinical and epidemiological data that indicate whether resistant isolates were recovered as a result of adamantane treatment or exposure were not available from countries outside the USA making it difficult to fully assess the effect of adamantane use on the observed trend. However, of 92 US patients from whom drug-resistant viruses were isolated since 2004, only two (2%), who were from nursing home facilities, were known to have
received adamantane treatment before virus collection. None of the other drug-resistant isolates obtained in the USA since 2004 were from patients documented to be on adamantane therapy, residents from semi-closed settings, nor in contact with patients on antiviral therapy.

The trend of a rising incidence of resistant viruses isolated from Asian countries compared with the lower incidence in the Americas during the period studied might be a consequence of differences in the procurement of amantadine and rimantadine in these countries. In the USA, adamantanes are licensed as anti-influenza drugs that are prescribed for influenza or influenza-like illness only by licensed physicians. However, in China, Russia, and some other countries, amantadine or rimantadine are available in over-the-counter formulations and are included in various cold remedies that do not need a prescription.9,10 Additionally, adamantanes are now broadly available as generic drugs, making it difficult to obtain or compare quantitative data on their distribution and clinical use. The lack of these data and their effect on increased resistance to these drugs is a limitation of our study.

Alternatively, drug-resistant mutations in the M2 gene could have occurred spontaneously, as was previously reported for some influenza H1N1 viruses (A/PR/8/34, A/WSN/33) that were resistant to adamantanes before these drugs were developed. Such a spontaneous mutation in the M2 gene may have been accompanied by the virus antigenic drift that allowed the widespread of recent influenza H3 viruses over the world. Such an assumption could explain the fact that the high increase in the proportion of drug-resistant isolates was observed for influenza A (H3) but not for influenza A (H1) viruses. The possibility that further genetic and antigenic evolution of influenza A (H3N2) viruses could result in the disappearance of this mutation in the M2 protein, reverting back to the drug-sensitive phenotype, should not be excluded.

Epidemic and pandemic strains of influenza often have been identified first in Asia.9 Many influenza experts have speculated that variants of influenza A (H5N1) viruses, which are currently circulating in avian species throughout southeast Asia and causing severe disease and mortality in human beings, have potential to cause the next influenza pandemic.11–16 We have shown that since 2003 all human and most avian influenza A (H5N1) isolates tested are resistant to amantadine and rimantadine.17,18 Data reported here on the high incidence of drug-resistant H3N2 viruses circulating in the same geographic region as H5N1 viruses suggest that greater caution is needed in the use of amantadine and rimantadine. With the increasing rates of resistance shown here, amantadine and rimantadine will probably no longer be effective for treatment or prophylaxis in the event of a pandemic outbreak of influenza. Our findings have broad implications for agencies and governments planning to stockpile these drugs as a front-line arsenal against epidemic or pandemic strains of influenza. With the establishment and increasing incidence of amantadine-resistant and rimantadine-resistant viruses throughout the world, careful assessment of the ability to rely on such drugs for treatment and prophylaxis before and during a pandemic or seasonal epidemic is warranted.

Our data raise concerns about the increasing incidence of adamantane-resistant influenza A viruses circulating throughout the world and draw attention to the importance of tracking the emergence and worldwide spread of drug-resistant viruses.

References


Increased risk of incident HIV during pregnancy in Rakai, Uganda: a prospective study

Ronald H Gray, Xianbin Li, Godfrey Kigozi, David Serwadda, Heena Brahmbhatt, Fred Wabwire-Mangen, Fred Nalugoda, Mohamed Kiddugavu, Nelson Sewankambo, Thomas C Quinn, Steven J Reynolds, Maria J Wawer

Summary

Background  HIV acquisition is significantly higher during pregnancy than in the postpartum period. We did a prospective study to estimate HIV incidence rates during pregnancy and lactation.

Methods  We assessed 2188 HIV-negative sexually active women with 2625 exposure intervals during pregnancy and 2887 intervals during breastfeeding, and 8473 non-pregnant and non-lactating women with 24 258 exposure intervals. Outcomes were HIV incidence rates per 100 person years and incidence rate ratios estimated by Poisson multivariate regression, with the non-pregnant or non-lactating women as the reference group. We also assessed the husbands of the married women to study male risk behaviours.

Findings  HIV incidence rates were 2·3 per 100 person years during pregnancy, 1·3 per 100 person years during breastfeeding, and 1·1 per 100 person years in the non-pregnant and non-lactating women. The adjusted incidence rate ratios were 2·16 (95% CI 1·39–3·37) during pregnancy and 1·16 (0·82–1·63) during breastfeeding. Pregnant women and their male partners reported significantly fewer external sexual partners than did the other groups. In married pregnant women who had a sexual relationship with their male spouses, the HIV incidence rate ratio was 1·36 (0·63–2·93). In married pregnant women in HIV-discordant relationships (ie, with HIV-positive men) the incidence rate ratio was 1·76 (0·62–4·03).

Interpretation  The risk of HIV acquisition rises during pregnancy. This change is unlikely to be due to sexual risk behaviours, but might be attributable to hormonal changes affecting the genital tract mucosa or immune responses. HIV prevention efforts are needed during pregnancy to protect mothers and their infants.

Introduction  An observational study in Malawi showed that rates of HIV acquisition are significantly higher during pregnancy than in the postpartum period.1,2 and data from Rakai, Uganda, suggested similar trends, but differentials in HIV incidence between pregnant and postpartum women were not significant.3,4 However, neither study adjusted for behavioural factors that can affect HIV acquisition. Women might be at increased risk of HIV during pregnancy because of their own or their partner’s sexual behaviours, or because of the physiological changes during gestation. Irrespective of the mechanism of heightened risk, these findings have important implications for HIV prevention, both to protect mothers from primary HIV infection during pregnancy and to potentially avoid mother-to-child HIV transmission, which can be increased by the rise in HIV-1 viraemia associated with recent maternal infection.5

To address these issues we assessed the incidence rate of HIV during pregnancy and lactation, and compared this to the incidence rate of HIV during periods of non-pregnancy and non-lactation using longitudinal data from a community cohort in Rakai, Uganda. We postulated that if the incidence rate of HIV acquisition during pregnancy remained raised after adjustment for characteristics and behaviours, physiological changes of pregnancy might be responsible for any differences in incidence rates recorded during pregnancy, lactation, and non-pregnancy and non-lactation periods.

Methods  

Study population  In this analysis we assessed HIV acquisition in three populations: a) all sexually active women; b) married women linked to (ie, having sex with) their husbands, among whom we could assess both the woman’s and her male partner’s self-reported sexual risk behaviours; and c) married women who were in HIV-discordant relationships with HIV-infected men, and thus had known high-risk exposures. We controlled for the characteristics and behaviours reported by the women and their husbands to assess whether the incidence of HIV during pregnancy was higher than HIV incidence during other periods.

The Rakai Community Cohort Study was established in 1994 for a community randomised trial of the control of sexually transmitted diseases for prevention of HIV (1994–99), and has continued annual surveillance thereafter.6 Censuses and surveys were done at 10–12 month intervals for all consenting adults aged 15–49 years who were resident in rural communities in Rakai District, Uganda. Interviews were undertaken to ascertain sociodemographic characteristics, sexual risk behaviours, and health status, including symptoms of genital ulcer disease. Serum samples were obtained for HIV
testing. The predominant mode of HIV transmission in this population is penile-vaginal intercourse. Women provided written informed consent at each study visit, and the study was approved by institutional review boards in Uganda and the USA.

Procedures
Pregnant women were identified by interview and physical examination, and for women who were unsure of their pregnancy status or whose last menstrual period occurred more than a month before the survey, pregnancy status was confirmed by a urinary human chorionic gonadotropin (hCG) test. During 1994–99, mothers were followed up postpartum,3 and those who were HIV-negative at the time their pregnancy was identified provided a postpartum blood sample to assess HIV acquisition during pregnancy. Those who were breastfeeding and who were HIV-negative after delivery were asked to provide a follow-up blood sample to detect HIV acquisition during lactation. Women who were initially HIV-negative and who were neither pregnant nor breastfeeding provided a comparison group for assessment of HIV risk.

To ascertain HIV acquisition in early and late pregnancy we also assessed women whose length of gestation at identification of pregnancy was less than the median and women whose length of gestation was more than the median.

Since risk of HIV in women might be associated with behaviours and HIV status of male partners, we identified husbands of married index women. We estimated mean number of male-reported sexual partners (inclusive of the index woman), and the proportion of men reporting more than one sexual partner during the interval of their wife’s exposure risk. The incidence of HIV acquisition by the index women and the incidence rate ratio of HIV were estimated after adjustment for male and female reported sexual behaviours. Additionally, we assessed the proportion of initially HIV-negative women who were in an HIV-discordant relationship with a HIV-positive male spouse, since these women are at very high risk of HIV infection. We also estimated rates of transmission per coital act in these HIV-discordant couples.

HIV status was assessed by two separate enzyme immunosorbent assays (Vironostika HIV, Organon Teknika, Charlotte, NC and Cambridge Biotech, Worcester, MA, USA) with western blot confirmation of discordant enzyme immunosorbent assay results and HIV seroconverters (HIV WB, Bio-Merieux-Vitek, St Louis, MS, USA). HIV-1 viral load in adults and infants was ascertained by RT-RNA PCR with Roche Amplicor 1.5 (Roche, Molecular Systems, Branchburg, NJ, USA).

Statistical analyses
The unit of analysis was a woman-interval of exposure to heterosexual HIV acquisition during pregnancy, breastfeeding, or neither state. For pregnant women, the exposure interval was from a negative HIV test at the time of pregnancy identification to the time of a repeat HIV test in the immediate postpartum period (an average exposure of 0·38 person years or 4·6 months per pregnancy). For breastfeeding women, the exposure interval was from the time of a negative postpartum HIV test to the next blood draw (an average of 1·05 person years per woman during lactation). For non-pregnant and non-lactating women, the exposure interval was between two successive HIV tests at about 10–12 month intervals (an average of 0·97 person years per woman interval). The exposure intervals during pregnancy were shorter than those in the other states because the average observation time between identification of pregnancy and postpartum HIV testing was 4·1 months per woman.

The HIV incidence rate per 100 person years was estimated for each pregnancy and lactation exposure state and we assumed that HIV infection occurred at the midpoint of the follow-up interval. We assessed HIV incidence rates by sociodemographic characteristics and behaviours for each interval of exposure in the three pregnancy and lactation states. Sociodemographic characteristics included age, marital status, and education; and sexual risk behaviours included the number of sex partners and condom use in the past year. Unadjusted and adjusted incidence rate ratios and 95% CIs were estimated by Poisson regression. Variables that were significantly associated with HIV acquisition in univariate analyses at p<0·15 were included in multivariate Poisson regression models to estimate the adjusted incidence rate ratios of HIV acquisition associated with pregnancy or lactation, relative to intervals in non-pregnant and non-lactating women (reference group). The covariates included in adjusted analyses were age, marital status, education, multiple sex partners, genital ulcer disease, and condom use. In analyses of women with linked male partners, models were adjusted for both male and female reported behaviours. An individual woman could contribute multiple periods of exposure to each state (eg, having two or more pregnancies and breastfeeding episodes, and multiple intervals of non-pregnancy and non-lactation), so robust standard errors were estimated with generalised estimation equation methods to adjust for correlated data.

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

Results
We identified 3134 pregnancies in which mothers were HIV-negative at the time their pregnancy was identified. There were 3031 breastfeeding women who were HIV-
negative after delivery and who provided a follow-up blood sample to detect HIV-acquisition during lactation during 1994–99. Additionally, there were 30545 follow-up intervals in women who were initially HIV-negative and who were neither pregnant nor breastfeeding. In the 3134 pregnancies, 509 women (16%) reported no sexual activity during pregnancy, leaving 2625 pregnancies to sexually active women at risk of heterosexual HIV-acquisition. These pregnancies occurred in 2188 women; 1765 women (81%) contributed information on only one pregnancy, 409 (19%) contributed information on two pregnancies, and 14 (1%) provided data for three pregnancies between 1994 and 1999.

Of 3041 intervals of breastfeeding in 2183 HIV-negative women, 144 (5%) were in women who had no sexual relationships during lactation, leaving 2887 intervals of breastfeeding in sexually active women from which to ascertain HIV incidence. Of the 2183 sexually active HIV-negative women, 1721 (60%) contributed one interval of observation, 273 (13%) contributed two, 145 (7%) contributed three, 36 (2%) contributed four, seven (0.3%) contributed five, and one woman (0.05%) was observed for six follow-up intervals. There were 30545 follow-up intervals for HIV-negative women who were neither pregnant nor lactating, and, among these, there were 6287 (21%) intervals in women with no sexual activity, leaving 24258 intervals of potential HIV risk. These intervals were recorded in 8473 women, of whom 2847 (34%) contributed one, 1817 (21%) contributed two intervals, 1135 (13%) contributed three, 788 (9%) contributed four, 758 (9%) contributed five, 609 (7%) contributed six, 376 (4%) contributed seven, and 143 (2%) contributed eight intervals of observation.

There were 997 person years of observations in 2625 pregnancies in women who were sexually active during their pregnancy. The 2887 intervals in sexually active breastfeeding women provided 3043 person years of exposure. Women who were neither pregnant nor breastfeeding provided 24 258 intervals and 24 161 person years of observation.

Of 2625 pregnant women, 2391 (91%) were currently married and 1240 (52%) had a husband who provided information on his sexual behaviours during his wife’s pregnancy. Similarly, in the 2887 intervals among breastfeeding women, 2627 (91%) women were currently married and we identified 1378 husbands (53%) who provided information of their behaviours. Of the 24258 non-pregnant and non-lactating women-intervals, 17772 occurred in currently married women (73%), and we identified 8338 linked husbands (47%). The difference in the proportion of linked husbands between the three exposure groups was highly significant ($p<0.0001$) and indicates lower rates of spousal absence during pregnancy and breastfeeding than at other times.

Table 1 shows the sociodemographic and behavioural characteristics by exposure group. Women contributing intervals of pregnancy or lactation were significantly younger, more frequently married, and less educated than the non-pregnant and non-lactating group. During pregnancy women reported a lower proportion of multiple sexual partners (1%) than did breastfeeding women (2%, $p=0.002$) or non-pregnant and non-lactating women (3%, $p<0.0001$). Condom use was also less frequent in pregnant women (104/2625, 4%) than in breastfeeding women (265/2887, 9%, $p<0.0001$) or the non-pregnant and non-lactating women (2377/24 258; 10%; 9%, $p<0.0001$). Symptoms of genital ulcer disease within the past 6 months were less common in pregnant women (145/2627, 5%) than in non-pregnant and non-lactating women (144/2887, 5%, $p=0.02$). Information on receipt of medical injections was available for a subsample of respondents since questions on injections were added late in the study, and was 44% (448/1019) in pregnant women, 45% (180/397) in those breastfeeding, and 44% (3105/7036) in the non-pregnant and non-lactating women. These differences were not significant ($p=0.9$).

Table 2 shows HIV incidence rates and the unadjusted incidence rate ratios of HIV acquisition. The HIV incidence rate was higher during intervals of pregnancy than during breastfeeding, or exposure intervals with no pregnancy or lactation. Relative to the non-pregnant and non-breastfeeding women, the incidence rate ratios of HIV acquisition. The HIV incidence rate ratio of HIV acquisition was also significantly higher during pregnancy than during breastfeeding (1.76, 95% CI 1.05–2.94).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Pregnant</th>
<th>Lactating</th>
<th>Non-pregnant/ non-lactating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2625</td>
<td>2887</td>
<td>24,258</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–19</td>
<td>481 (18%)</td>
<td>406 (14%)</td>
<td>2669 (10%)</td>
</tr>
<tr>
<td>20–29</td>
<td>1533 (58%)</td>
<td>1826 (63%)</td>
<td>1422 (6%)</td>
</tr>
<tr>
<td>&gt;30</td>
<td>611 (23%)*</td>
<td>655 (23%)*</td>
<td>11,611 (48%)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>121 (5%)</td>
<td>130 (5%)</td>
<td>2651 (11%)</td>
</tr>
<tr>
<td>Currently married</td>
<td>2391 (91%)</td>
<td>2627 (91%)</td>
<td>17,772 (73%)</td>
</tr>
<tr>
<td>Previously married</td>
<td>113 (4%)*</td>
<td>130 (5%)*</td>
<td>3835 (16%)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>290 (11%)</td>
<td>280 (10%)</td>
<td>2921 (13%)</td>
</tr>
<tr>
<td>Primary</td>
<td>1847 (70%)</td>
<td>2083 (72%)</td>
<td>16,976 (66%)</td>
</tr>
<tr>
<td>Secondary or higher</td>
<td>488 (19%)*</td>
<td>524 (18%)*</td>
<td>5500 (21%)</td>
</tr>
<tr>
<td>Sex partners in past year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2600 (99%)</td>
<td>2831 (98%)</td>
<td>23,615 (97%)</td>
</tr>
<tr>
<td>&gt;2</td>
<td>25 (1%)*</td>
<td>56 (2%)*</td>
<td>643 (3%)</td>
</tr>
<tr>
<td>Condom use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>2521 (96%)</td>
<td>2622 (91%)</td>
<td>21,881 (90%)</td>
</tr>
<tr>
<td>Irregular</td>
<td>88 (3%)</td>
<td>219 (8%)</td>
<td>1422 (6%)</td>
</tr>
<tr>
<td>Consistent</td>
<td>16 (1%)*</td>
<td>46 (2%)*</td>
<td>955 (4%)</td>
</tr>
<tr>
<td>GUD present</td>
<td>335 (5%)*</td>
<td>218 (8%)</td>
<td>1649 (7%)</td>
</tr>
<tr>
<td>No GUD</td>
<td>2483 (95%)</td>
<td>2660 (93%)</td>
<td>22,582 (93%)</td>
</tr>
</tbody>
</table>

Data are number of intervals (%). *comparison of pregnant and lactating versus non-pregnant/non-lactating women. $p<0.0001$, $p<0.05$. 

Table 1: Characteristics of sexually active women during pregnancy and lactation, and during non-pregnant and non-lactating intervals
HIV incidence rates during pregnancy were significantly greater than in the non-pregnant and non-lactating group for virtually all strata of covariates. For example, the HIV incidence rate was significantly higher in pregnant women than in non-pregnant and non-lactating women aged 15–19 years, those never married, previously married women, and those reporting no condom use or inconsistent condom use. By contrast, HIV incidence rates during breastfeeding were not significantly different from those in non-pregnant and non-lactating women for any covariate strata.

With respect to other covariates associated with HIV, incidence rates were highest in adolescent women and those aged 20–29 years, those previously married and never married, those reporting multiple sex partners during the period of exposure, and in women with symptoms of genital ulceration. HIV incidence rates were higher in women with higher education in the pregnant and breastfeeding groups than in the non-pregnant and non-lactating group. Tests for interactions between education and HIV acquisition in pregnant and breastfeeding women were not significant. HIV incidence rates were lower in non-condom users than condom users, but interpretation was constrained by small sample sizes in the pregnant and breastfeeding groups, and by the fact that female condom use in this population is strongly correlated with multiple sexual partners.

We assessed HIV incidence rates in pregnant women identified early and late in gestation. The median length of gestation at the time a woman was identified as pregnant was 5 months. In 1257 women the length of gestation at identification of pregnancy was less than the median with 715·8 person years, an average of 6·8 months exposure per pregnancy and in 1368 women the length of gestation was more than the median with 280·8 person years per pregnancy, an average exposure of 2·5 months per woman. Of those whose pregnancy was identified earlier than 5 months, there were 17 seroconversions over 715·8 person years, with an incidence rate of 2·4 per 100 person years. Of those identified in later pregnancy, there were six seroconversions over 280·8 person years, an incidence rate of 2·1 per 100 person years.

Table 2: HIV incidence among pregnant or lactating women and women who were neither pregnant nor lactating

<table>
<thead>
<tr>
<th>Marital status</th>
<th>Pregnant</th>
<th>Lactating</th>
<th>Non-pregnant/non-lactating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>23/997 (2·3)</td>
<td>40/3043 (1·3)</td>
<td>275/24 161 (1·1)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–19</td>
<td>7/191 (3·7)</td>
<td>6/412 (1·5)</td>
<td>31/2393 (1·3)</td>
</tr>
<tr>
<td>20–29</td>
<td>13/582 (2·2)</td>
<td>27/1931 (1·4)</td>
<td>137/10 183 (1·1)</td>
</tr>
<tr>
<td>30–39</td>
<td>3/273 (1·3)</td>
<td>7/700 (1·0)</td>
<td>107/11 590 (0·9)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No education</td>
<td>5/50 (10·0)</td>
<td>4/291 (1·4)</td>
<td>31/2491 (1·2)</td>
</tr>
<tr>
<td>Primary</td>
<td>15/914 (1·6)</td>
<td>32/2766 (1·2)</td>
<td>176/17 530 (1·0)</td>
</tr>
<tr>
<td>Secondary or higher</td>
<td>3/39 (7·7)</td>
<td>5/141 (3·6)</td>
<td>67/3941 (1·7)</td>
</tr>
<tr>
<td>Sex partners in past year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23/986 (2·3)</td>
<td>38/2983 (1·3)</td>
<td>247/23 507 (1·1)</td>
</tr>
<tr>
<td>2</td>
<td>10/10 (100)</td>
<td>2/61 (3·3)</td>
<td>28/658 (4·3)</td>
</tr>
<tr>
<td>Condom use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>19/950 (2·0)</td>
<td>35/2762 (1·3)</td>
<td>235/21 718 (1·1)</td>
</tr>
<tr>
<td>Irregular</td>
<td>4/32 (12·5)</td>
<td>4/232 (1·7)</td>
<td>24/1470 (1·6)</td>
</tr>
<tr>
<td>Always</td>
<td>0/5</td>
<td>1/49 (2·0)</td>
<td>16/975 (1·6)</td>
</tr>
<tr>
<td>GUD present</td>
<td>4/52 (7·7)</td>
<td>5/240 (2·1)</td>
<td>39/696 (2·3)</td>
</tr>
<tr>
<td>No GUD</td>
<td>19/945 (2·0)</td>
<td>35/2803 (1·3)</td>
<td>234/22 441 (1·0)</td>
</tr>
</tbody>
</table>

*Risk ratio of HIV relative to non-pregnant or non-lactating women. GUD=genitourinary disease.

Table 3: Male reported sexual behaviours and HIV acquisition in married women linked to their male partners

<table>
<thead>
<tr>
<th>Marital status</th>
<th>Pregnant</th>
<th>Lactating</th>
<th>Non-pregnant/non-lactating</th>
</tr>
</thead>
<tbody>
<tr>
<td>External partners reported by husbands</td>
<td>43/1240 (34%)</td>
<td>545/1378 (40%)</td>
<td>327/8 138 (39%)</td>
</tr>
<tr>
<td>HIV incidence in married women*</td>
<td>7/483 (1·4%)</td>
<td>15/1396 (1·1%)</td>
<td>87/8 147 (1·1%)</td>
</tr>
</tbody>
</table>

*Incident cases/person years (incidence per 100 person years).
Pregnant | Lactating | Neither pregnant nor lactating
---|---|---
Husband HIV positive* | 105/1183 (9%) | 117/1221 (9%) | 754/7851 (10%)
HIV viral load in HIV positive husbands† | 105 (4.11) | 117 (4.18) | 754 (4.2)
Transmission from HIV positive male‡ | 0/40 (15.0) | 11/113 (9.6) | 80/729 (8.3)
Transmissions per coital act from HIV positive husbands§ | 1/204 (0.013) | 5/10-406 (0.0009) | 50/59-331 (0.0007)

*Data are Number/N (%). †Data are number log10 (HIV viral load capsid/mL). ‡Data are incident cases/person years (incidence rate per 100 person years). §Data are incident cases/coital acts (transmission rate per coital act).

Table 4: HIV acquisition among women in discordant relationships with an HIV-infected male partner

280-8 person years, with an incidence rate of 2.1 per 100 person years (incidence rate ratio 1.11, 95% CI 0.42–3.44).

Poisson multivariate models were constructed to estimate the adjusted rate ratios of HIV acquisition associated with intervals of pregnancy or lactation, relative to intervals with neither pregnancy nor breastfeeding, and included covariates that were significantly associated with HIV risk in univariate analyses (age, marital status, education, multiple sex partners, genital ulcer disease, and condom use). Relative to the non-pregnant and non-lactating group, the adjusted incidence rate ratio of HIV acquisition during pregnancy was 2.16 (95% CI 1.39–3.37), which is close to the unadjusted estimate of 2.03 (table 2), suggesting minimum confounding by these covariates. For breastfeeding, the adjusted rate ratio of HIV acquisition was 1.16 (95% CI 1.07–1.26). The adjusted incidence rate ratio of HIV acquisition during pregnancy relative to breastfeeding women was 1.82 (1.09–3.05). Because there were major disparities between the three groups in the proportion of women aged over 30 years (table 1), we also undertook stratified analyses confined to women aged 15–29 years. In this age-stratified analysis, the multivariate adjusted incidence rate ratio of HIV acquisition for pregnant women relative to non-pregnant and non-lactating women was 2.15 (1.33–3.47). In breastfeeding women, the 15–29 years age-stratified incidence rate ratio was 1.14 (0.77–1.68).

Other covariates significantly associated with HIV risk in multivariate analyses included age groups 15–19 years and 20–29 years, previously married status (ie, separated, divorced, or widowed), women reporting multiple sex partners (2.98, 2.00–4.46), and symptoms of genital ulcer disease (2.07, 1.52–2.81).

Table 3 shows the sexual behaviours reported by the husbands of married women during the intervals of female exposure to HIV risk. Husbands of the pregnant women reported a lower mean number of sex partners (1·48 [SD 0·86]) than did husbands of the breastfeeding women (1·57 [0·95]) and of the non-pregnant and non-lactating women (1·57 [0·98], p=0.007). Also, the proportion of husbands of pregnant women reporting more than one sex partner (36%) was lower than multiple partners reported by husbands of breastfeeding women (40%), and husbands of non-pregnant and non-lactating women (39%). These married women had a higher HIV incidence during pregnancy than the non-pregnant or non-breastfeeding women, but this difference was not significant (adjusted incidence rate ratio 1·36, 95% CI 0·63–2·93).

As shown in table 4, the proportions of uninfected women married to HIV-infected men were similar in the three exposure groups (p=0·50), and HIV viral loads of the infected male partners were also similar between the three strata (p=0·50). In women married to HIV-infected men, the HIV incidence rate was higher during pregnancy than incidence rates estimated in either breastfeeding or non-pregnant and non-lactating women, although these differences were not significant (incidence rate ratio 1·76, 95% CI 0·62–4·03). The mean monthly frequency of intercourse was lower during pregnancy (6·7 acts per month) than during breastfeeding (7·5 acts per month) and during non-pregnant and non-lactating intervals (8·0 per month; p<0·05). Therefore, we also estimated the rate of HIV acquisition per coital act, which was higher during pregnancy than in the non-pregnant and non-lactating group (incidence rate ratio 1·42, 95% CI 0·37–3·82).

Discussion

Women had a significantly heightened risk of HIV acquisition during pregnancy when compared with women who were breastfeeding or were neither pregnant nor breastfeeding, and the excess risk of HIV acquisition during pregnancy remained significant after adjustment for sociodemographic and behavioural factors. These findings are lent support by workers in Malawi, who reported a 2·19 fold higher rate of HIV incidence during pregnancy (incidence rate 7·9 per 100 person years) than in the postpartum period (3·6 per 100 person years). Investigators in Rwanda also reported higher HIV incidence rates during the early postpartum period than at later time intervals. Our findings lend support to these earlier observations, but, unlike these previous studies, we were able to carefully control for sexual risk factors, including those of the woman’s primary partner.

We assessed sexual behaviours reported by the women and their husbands, and showed that pregnant women were significantly less likely to report multiple sexual partners than were non-pregnant and non-lactating women, and in married couples the husbands of pregnant women reported significantly fewer sexual partners than husbands of non-pregnant and non-lactating women. Although there could be misreporting of sexual behaviours, the results are unlikely to differ between the three exposure groups, so both female and male sexual behaviours are unlikely to account for the excess risk of HIV during pregnancy. Also, behavioural factors should have less effect on HIV risk in HIV-
discordant couples, in whom rate of HIV acquisition during pregnancy was higher than at other times although this difference was not significant. We also examined other potential exposures and showed that use of medical injections did not differ between the three groups of women. Therefore, by a process of elimination, we conclude that behavioural factors are unlikely to explain why the HIV incidence rate is increased during pregnancy, and we speculate that biological factors might have a role.

The high levels of oestrogen and progesterone during pregnancy can affect a woman’s susceptibility to HIV infection by inducing structural changes in the genital tract mucosa, or by immunological effects. Pregnancy causes increased ectopy because of hyperplasia of the columnar epithelium and glands, hyperaemia, and stromal oedema, which could also increase susceptibility to HIV. Upper genital tract epithelial cells express HIV-1 co-receptors, which are under hormonal regulation, and these cells can be productively infected with HIV. In HIV-negative women, combined oral contraceptives can upregulate cervical HIV-1 CCR5 co-receptors, which might also increase susceptibility to HIV infection. Hormonal contraception has been associated with increased risks of HIV acquisition in some but not all epidemiological studies. In macaques, topical progesterone increases susceptibility to simian immunodeficiency virus (SIV) challenge, whereas systemic or topical estriol protects against SIV, so the net effect of high oestrogen and progesterone concentrations in pregnancy on susceptibility to HIV cannot be determined. The fetal trophoblast is thought to induce stimulation of CD4+ T-helper cells and suppression of cytotoxic natural killer cells, which might increase susceptibility. Also, hormonal changes during the luteal phase of the menstrual cycle are associated with increased mucosal lymphoid aggregates and suppression of cytotoxic T-lymphocyte activity. In summary, immunological changes during pregnancy could increase susceptibility to HIV infection in pregnant women, but the evidence is inconclusive.

A strength of this study is that pregnant women were HIV-negative at the time their pregnancy was detected, and by retesting shortly after delivery we could determine the interval of HIV exposure risk with some precision. Infections during pregnancy might have been missed if women were in the window period at the time of the HIV test during pregnancy or postpartum. However, such misclassification would also affect detection of incident infections in the breastfeeding and the non-pregnant and non-lactating women, so differential misclassification of the timing of HIV infection is unlikely to have biased our estimates of HIV acquisition. Another strength of this study was that we controlled for sexual behaviours reported by the women and their husbands, and examined other potential exposures, such as medical injections. There could be differential misreporting of sexual behaviours during pregnancy, breastfeeding, or at other times. Self-reported risk behaviours, such as multiple sex partners, were associated with raised HIV incidence in the breastfeeding and in the non-pregnant and non-breastfeeding women, but too few pregnant women reported multiple partnerships to assess effects in this small subgroup. Also, other Rakai studies have shown consistency in the frequency of intercourse reported by partners within couples, and in HIV-discordant couples sequencing of the HIV-1 viral genome of transmitted virus has confirmed self-reported monogamy.

Because early HIV infection is associated with increased HIV-1 viraemia, incident maternal infections might increase mother-to-child HIV transmission, although no effect of maternal serconversion was reported in a small study of 16 Thai children. Our study did not have sufficient power to address this issue.

We cannot generalise our findings beyond the Rakai setting and it would be important for other investigators to verify our results. However, if women are at increased risk of HIV acquisition during pregnancy, our findings present a public-health problem, both for the mother and possibly for her unborn infant. We believe that it would be prudent to warn women of this potential risk of HIV acquisition during pregnancy, and to promote safe sex (ie, monogamy and condom use), or sexual abstinence where feasible. This action could also protect women from other sexually transmitted infections, which have adverse effects on pregnancy outcome. Another important consideration is that 20–25% of African women are exposed to the risk of unintended pregnancy but are not using contraception. Thus, reduction of HIV risk by prevention of unintended pregnancy adds to the compelling rationale for provision of family planning services, especially where the risk of HIV exposure is high.

Contributors
R H Gray, G Kigozi, D Serwadda, F Wabwire-Mangen, F Nalugoda, M Kiddugavu, N Sewankambo, and M J Wawer contributed to original data collection, analyses and preparation of this report; X Li contributed to creation of datasets and analysis; H Brahmbhatt contributed to the analysis of mother-to-child HIV transmission and preparation of this report; S J Reynolds contributed to data analysis and interpretation; and T C Quinn contributed to laboratory assays and preparation of this report.

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

Acknowledgments
This work was supported by grants RO1A134826 and RO1A134265 from the National Institute of Allergy and Infectious Diseases; grant 5P30HD06826 from the National Institute of Child Health and Development; grant HM1F5686/CDFDA#1214 from the Henry M Jackson Foundation for the Advancement of Military Medicine and the United States Department of Defense, and The Johns Hopkins University, Bill and Melinda Gates Institute of Population and Reproductive Health. We thank James Shelton for his helpful comments on the manuscript.
References


Post-surgery adjuvant therapy with intermediate doses of interferon alfa 2b versus observation in patients with stage IIb/III melanoma (EORTC 18952): randomised controlled trial

Alexander M M Eggermont, Stefan Suciu, Rona MacKie, Wlodzimierz Ruka, Alessandro Testori, Wim Kruit, Cornelis J A Punt, Michelle Delauney, Francois Sales, Gerard Groenewegen, Dirk J Ruiter, Izabella Jagiello, Konstantin Stoitchkov, Ulrich Keilholz, Danielle Lienard, for the EORTC Melanoma Group*

Summary
Background Individuals affected by melanoma with thick primary tumours or regional node involvement have a poor outlook, with only 30–50% alive at 5 years. High-dose and low-dose interferon alfa have been assessed for the treatment of these patients, with the former having considerable toxicity and a consistent effect on disease-free survival, but not on overall survival, and the latter no consistent effect on either. Our aim was, therefore, to assess the effect of two regimens of interferon of intermediate dose versus observation alone on distant metastasis-free interval (DMFI) and overall survival in such patients.

Methods We did a randomised controlled trial in 1388 patients who had had a thick primary tumour (thickness ≥4 mm) resected (stage IIb) or regional lymph node metastases dissected (stage III) and had been assigned to 13 months (n=553) or 25 months (n=556) of treatment with subcutaneous interferon alfa 2b, or observation (n=279). Treatment comprised 4 weeks of 10 million units (MU) of interferon alfa (5 days per week) followed by either 10 MU three times a week for 1 year or 5 MU three times a week for 2 years, to a total dose of 1760 MU. Our primary endpoint was DMFI. Analyses were by intent to treat.

Findings After a median follow-up of 4–65 years, we had recorded 760 distant metastases and 681 deaths. At 4–5 years, the 25-month interferon group showed a 7·2% increase in rate of DMFI (hazard ratio 0·93, 97·5% CI 0·66–1·03) and a 5·4% improvement in overall survival. The 13-month interferon group showed a 3·2% increase in rate of DMFI at 4–5 years (0·97, 0·75–1·26) and no extension of overall survival. Toxicity was acceptable, with 18% (195 of 1076) of patients going off study because of toxicity or as a result of refusal of treatment because of side-effects.

Interpretation Interferon alfa as used in the regimens studied does not improve outcome for patients with stage IIb/III melanomas, and cannot be recommended. With respect to efficacy of the drug, duration of treatment seemed more important than dose, and should be assessed in future trials.

Introduction There is little evidence that non-surgical therapy confers any survival benefits over observation alone in individuals with metastatic melanoma. Furthermore, polychemotherapy and biochemotherapy, despite higher response rates, are no more effective than single-agent treatment with dacarbazine, indicating that an effective adjuvant therapy for melanoma remains to be identified.

The efficacy of interferon alfa has been widely assessed in individuals at intermediate risk for relapse (stage II: primary tumour ≥1·5 mm, clinically node negative) and in those at high risk for relapse (stage IIb: primary tumour ≥4·0 mm, node negative; stage III: any primary tumour, node positive), despite findings of phase III trials in patients with stage IV disease showing no survival benefit. Dose and duration of treatment are thought to be key factors that affect efficacy.

The findings of one phase III trial, which assessed treatment with high-dose interferon alfa for 1 year, indicate a positive effect of the drug on both disease-free survival and overall survival. The effect on overall survival was small, however, not sustained over time, and has not been confirmed by the results of a second trial. Preliminary findings from another trial of high-dose interferon alfa showed a significant effect on both disease-free survival and overall survival after 1·3 years, but after 2·2 years follow-up the effect on overall survival had diminished. There is, therefore, no firm or consistent evidence of a lasting effect for high-dose interferon alfa.

Various trials of adjuvant therapy with low-dose interferon alfa have also been done with no indication of a survival benefit, thought the results of some show a slight positive effect on disease-free survival others do not. We chose to study intermediate doses of interferon alfa at a time when only high-dose interferon was showing...
any effect on survival, and when many judged the toxicity of treatment to be prohibitive. Without any guidance about a dose-efficacy relation in stage IV cancer, the aim of the EORTC (European Organisation for Research and Treatment of Cancer) trial 18952 was to assess the effect of intermediate doses of interferon alfa versus observation alone in patients with stage IIb/III melanoma to try to identify a threshold concentration of activity and any effect of treatment duration.

**Methods**

**Patients**

Between May, 1996, and June, 2000, we recruited to a randomised controlled trial patients aged 18–75 years who had had curative dissection of regional lymph node metastases (melanoma stage III: primary tumour of any size [T], node positive [N],) or who had had curative dissection of regional lymph node metastases (melanoma stage III: primary tumour of any size [T], node positive [N],) in this trial, N, comprised patients with non-enlarged microscopically involved lymph nodes on sentinel node biopsy, and N, those with palpable tumour-involved nodes. We recruited patients from 85 institutions in 22 countries (webtable). We did not include patients with mucosal or ocular melanoma, those previously treated with systemic drugs for melanoma, those with other malignant diseases (other than basal cell carcinoma, in-situ cervical cancer), autoimmune disease, uncontrolled infections, cardiopulmonary disease, liver or renal disease, or those taking corticosteroids.

All patients provided written informed consent, and the EORTC protocol review committee and the ethical
review boards of the participating centres approved the protocol of the study.

Procedures
We resected the primary tumours of patients with a free margin of 1 cm or more, and regional lymph node dissections had to contain more than five nodes (inguinal), more than ten nodes (axillary), or more than 15 nodes (neck). Adjuvant therapy was started within 10 weeks of definitive surgery. The diagnosis of all patients was reviewed by the EORTC melanoma pathology review board. Case report forms were sent to the EORTC data centre every 3 months in year 1, every 4 months in year 2, every 6 months in years 3–5, and yearly thereafter. Initial staging included chest radiograph and either ultrasound of regional lymph nodes and abdomen or a CT scan.

We randomly assigned patients to one of three groups in a 2:2:1 fashion: 13 months or 25 months of treatment with subcutaneous interferon alfa 2b (Intron A, Schering-Plough, Kenilworth, New Jersey, USA), or observation. Treatment comprised 4 weeks of 10 million units (MU) of interferon (5 days per week), followed by either 10 MU three times a week for 1 year or 5 MU three times a week for 2 years to a total dose of 1760 MU in both treatment groups. Randomisation was done centrally from the EORTC data centre in Brussels, with the minimisation technique. Patients were stratified by centre, sex, site of tumour (head, neck, or trunk; limb; unknown), tumour staging (T4N0; TanyN1; TanyN2), and, for the latter stages, the number of positive lymph nodes (n=1; n=2–4; n>5; unknown), and the Breslow thickness (<1 mm; 1·0–1·99 mm; 2·0–2·99 mm; 3·0–3·99 mm; ≥4 mm; unknown).

Our primary endpoint was distant metastases free interval (DMFI) after randomisation, defined as time from randomisation to appearance of distant metastases. We censored the follow-up of patients who did not develop metastases and who died from causes other than melanoma at the date of death, and of those still alive at the latest date of follow-up. Our secondary endpoints were distant metastases free survival (DMFS)—defined as the time from randomisation to occurrence of distant metastases or death, whatever the cause—duration of survival (time from randomisation until death, whatever the cause), and toxicity. It is noteworthy that we included as events for DMFS analysis patients who developed metastases (also included in DMFI analysis) and those who did not develop distant metastases and who died from causes other than melanoma.

Statistical analysis
We calculated that 1000 patients (200 in observation group, and 400 patients in each treatment group) would need to be enrolled to detect a 15% difference (40% vs 55%) at the 5% significance level in DMFI at 3 years between each treatment group and the observation group. However, 6 months before the projected end of accrual, the results of the ECOG (Eastern Cooperative Oncology Group) trial 1690 were published, showing that 1 year of treatment with high-dose interferon alfa had no survival advantage compared with observation and that the 5-year disease-free survival rate advantage was about 10%. We therefore decided to increase the study population to 1400 patients. For a two-group comparison, a total of 471 patients (ie, 177 + 294) followed up until distant metastases would allow the detection with an 80% statistical power of a 10–5% difference in the 4-year DMFI rates (40% vs 50–5%), corresponding to a hazard ratio (HR) of 0·746 in case of an exponential distribution. Therefore, for the two comparisons (10 MU interferon three times a week after induction for 1 year vs observation, and 5 MU three times a week after induction for 2 years vs observation) a total of 765 events (distant metastases) were needed to do the final analysis.

We calculated actuarial curves for the three time-to-event endpoints—DMFI, DMFS, and survival—with the Kaplan-Meier technique, and obtained the standard errors (SE) of the estimates with the Greenwood formula. We tested the difference between curves for significance with the two-tailed log-rank test.

### Table 1: Baseline characteristics

<table>
<thead>
<tr>
<th></th>
<th>13-month interferon alfa (n=553)</th>
<th>25-month interferon alfa (n=556)</th>
<th>Observation (n=279)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>312 (56%)</td>
<td>308 (55%)</td>
<td>152 (54%)</td>
</tr>
<tr>
<td>Female</td>
<td>241 (44%)</td>
<td>248 (45%)</td>
<td>127 (46%)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>49 (37–74)</td>
<td>50 (36–75)</td>
<td>47 (20–75)</td>
</tr>
<tr>
<td>&lt;35</td>
<td>105 (19%)</td>
<td>88 (16%)</td>
<td>48 (17%)</td>
</tr>
<tr>
<td>35–49</td>
<td>188 (34%)</td>
<td>188 (34%)</td>
<td>106 (38%)</td>
</tr>
<tr>
<td>50–64</td>
<td>192 (35%)</td>
<td>196 (35%)</td>
<td>83 (30%)</td>
</tr>
<tr>
<td>≥65</td>
<td>68 (12%)</td>
<td>84 (15%)</td>
<td>42 (15%)</td>
</tr>
<tr>
<td><strong>Stage of disease</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ib</td>
<td>141 (25%)</td>
<td>142 (26%)</td>
<td>73 (26%)</td>
</tr>
<tr>
<td>II (N1)</td>
<td>141 (25%)</td>
<td>144 (26%)</td>
<td>68 (24%)</td>
</tr>
<tr>
<td>II (N2)</td>
<td>271 (49%)</td>
<td>270 (49%)</td>
<td>138 (49%)</td>
</tr>
<tr>
<td><strong>Number of positive lymph nodes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>144 (26%)</td>
<td>150 (27%)</td>
<td>73 (26%)</td>
</tr>
<tr>
<td>1</td>
<td>187 (34%)</td>
<td>191 (34%)</td>
<td>103 (37%)</td>
</tr>
<tr>
<td>2–4</td>
<td>161 (29%)</td>
<td>152 (27%)</td>
<td>76 (27%)</td>
</tr>
<tr>
<td>≥5</td>
<td>61 (11%)</td>
<td>63 (11%)</td>
<td>27 (10%)</td>
</tr>
<tr>
<td><strong>Breslow thickness (mm)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>29 (5%)</td>
<td>33 (6%)</td>
<td>35 (5%)</td>
</tr>
<tr>
<td>1–1·99</td>
<td>86 (16%)</td>
<td>90 (16%)</td>
<td>41 (15%)</td>
</tr>
<tr>
<td>2·3·99</td>
<td>115 (21%)</td>
<td>113 (20%)</td>
<td>60 (22%)</td>
</tr>
<tr>
<td>≥4</td>
<td>254 (46%)</td>
<td>254 (46%)</td>
<td>131 (47%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>69 (12%)</td>
<td>66 (12%)</td>
<td>32 (11%)</td>
</tr>
<tr>
<td><strong>Location of primary tumour</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limb</td>
<td>241 (44%)</td>
<td>247 (44%)</td>
<td>124 (44%)</td>
</tr>
<tr>
<td>Head, neck, or trunk</td>
<td>287 (52%)</td>
<td>288 (52%)</td>
<td>142 (51%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>25 (5%)</td>
<td>21 (4%)</td>
<td>15 (5%)</td>
</tr>
<tr>
<td><strong>Ulceration of primary tumour</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>285 (52%)</td>
<td>263 (47%)</td>
<td>135 (48%)</td>
</tr>
<tr>
<td>Yes</td>
<td>165 (30%)</td>
<td>205 (37%)</td>
<td>106 (38%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>103 (19%)</td>
<td>88 (16%)</td>
<td>38 (14%)</td>
</tr>
</tbody>
</table>

Data are number (%) unless otherwise indicated.
estimated the cumulative incidence of distant metastases and the incidence of death without distant metastases, and their corresponding SE, with competing risk methods.\(^9\) We used the Gray test\(^{15}\) to compare these incidences.

We used the Cox’s proportional hazards model to obtain the estimate and the 97·5% CI of the HR of the instantaneous event rate in each treatment group versus that in the observation group, adjusting for possible confounding factors. The Wald test has been used to ascertain the prognostic importance of every variable included in the model.\(^7\) We tested the prognostic interaction between variables by including products of variables in the model. For the pair-wise comparisons, we judged a p value of 0·025 or less significant; for the subgroup analyses according to the initial stage, we judged a p value of 0·01 or less significant. All analyses were done with SAS version 8.2 (SAS Institute, Cary, NC, USA), and were by intent to treat. The database, located at the EORTC data centre, was frozen in May, 2004.

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

Results
We enrolled and randomly assigned 1388 patients (figure 1). In the 13-month and 25-month treatment groups, 388 of 553 (70%) and 382 of 556 (69%) patients, respectively, received at least three-quarters of the required dose of interferon alfa during the treatment period they were on study—ie, until 13 months or 25 months from randomisation, or until the moment distant metastases appeared.

Out of 1388 patients, ten were identified as ineligible by the centre after randomisation and were taken off the protocol (figure 1). The trial co-ordinator identified another nine patients who had been incorrectly enrolled. In total, 19 patients (1%) were judged ineligible (11 in the 13-month interferon group, seven in the 25-month interferon group, and one in the observation group): one because of previous treatment, ten because of incorrect stage allocation, and eight for other reasons.

Table 1 shows the baseline characteristics. About a quarter of patients in each of the groups had melanoma stage IIb, a quarter stage III (N1), and half stage III (N2). Overall the baseline characteristics of the groups did not differ. The median age of all patients was 49 years. At the time of final analysis, the median follow-up was 4·65 years (range 1–7 years). 53 patients (3%) were lost to follow-up (figure 1).

With respect to our primary endpoint, 760 of 1388 patients developed distant metastases; 309 in the 13-month treatment group, 287 in the 25-month treatment group, and 164 in the observation group (figure 2). At 4·5 years’ follow-up, using the Kaplan-Meier method, the estimated rates of DMFI were similar across the three groups (table 2). Stage of disease at randomisation was an important prognostic factor. The estimates of the 4·5 year DMFI rates in stage IIb, stage III (N\(_1\)), and stage III (N\(_2\)) cancers were 60·9%, 52·7%, and 31·2%, respectively (table 3). With the Cox model the treatment differences, adjusted for the initial stage, remained practically unchanged: for 13-month interferon alfa versus observation the HR was 0·94 (97·5% CI

### Table 1: Baseline characteristics

<table>
<thead>
<tr>
<th></th>
<th>13-month interferon alfa</th>
<th>25-month interferon alfa</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>388</td>
<td>382</td>
<td>190</td>
</tr>
<tr>
<td>Age (years)</td>
<td>49.2 (2.8)</td>
<td>49.3 (2.8)</td>
<td>49.1 (2.8)</td>
</tr>
<tr>
<td>Gender (men)</td>
<td>57%</td>
<td>56%</td>
<td>56%</td>
</tr>
<tr>
<td>T stage (I)</td>
<td>23%</td>
<td>22%</td>
<td>25%</td>
</tr>
<tr>
<td>T stage (II)</td>
<td>30%</td>
<td>31%</td>
<td>30%</td>
</tr>
<tr>
<td>T stage (III)</td>
<td>47%</td>
<td>47%</td>
<td>45%</td>
</tr>
</tbody>
</table>

### Table 2: Treatment comparison by endpoint

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>13-month interferon alfa</th>
<th>25-month interferon alfa</th>
<th>Observation</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMFI</td>
<td>43.2% (2.2)</td>
<td>42.7% (2.2)</td>
<td>40.0% (3.1)</td>
<td>0.93 (0.75–1.16)</td>
</tr>
<tr>
<td>Cumulative rate of distant metastases*</td>
<td>56.8% (2.2)</td>
<td>52.8% (2.2)</td>
<td>60.0% (3.1)</td>
<td>0.95 (0.77–1.18)</td>
</tr>
<tr>
<td>DMFS</td>
<td>42.3% (2.2)</td>
<td>46.1% (2.2)</td>
<td>40.0% (3.1)</td>
<td>p=0.44</td>
</tr>
<tr>
<td>Incidence of distant metastases†</td>
<td>56.6% (2.2)</td>
<td>52.4% (2.2)</td>
<td>60.0% (3.1)</td>
<td>0.97 (0.77–1.21)</td>
</tr>
</tbody>
</table>

*100% DMFI rate at 4·5 years. †Cumulative incidences calculated with competing risk methods.
0.76–1.17; p=0.52) and for 25-month interferon alfa versus observation the HR was 0.83 (0.66–1.03; p=0.05).

After further adjustment for additional prognostic factors for DMFI (table 3)—ie, by sex, tumour site, ulceration—the treatment difference between the 25-month interferon alfa group and observation and the test for interaction between treatment and stage were borderline significant.

Since the results of these exploratory analyses suggested that the higher the stage the lower the difference between treatment with 25-month interferon alfa and observation, we did subgroup analyses by tumour stage. In the stage Ib subgroup, adjustment for sex, tumour site, and ulceration led to a significant treatment difference between the 25-month interferon alfa group and observation (HR 0.54) (table 3). The 217 patients with stage III (N1) melanomas had a very similar outcome to the 120 patients with two to four positive lymph nodes (HR 1.04), whereas the 15 patients with five or more nodes had the worst prognosis (HR 4.2, p=0.029). Using the Cox model, the comparison of 25-month interferon alfa and observation, adjusted for number of lymph nodes, sex, location of tumour, and ulceration led to estimated HRs of 0.66 and 0.90 in patients with stage III (N1) and stage III (N2) melanomas, respectively (table 3). For the comparison of 13-month interferon alfa and observation, the increase of the HRs, according to stage, was less pronounced and not significant.

A combined variable, taking into account stage and number of lymph nodes, was of greater prognostic importance than stage alone. Adjustment for this variable in a Cox model did not greatly affect the difference in treatment effect noted between the 13-month interferon group and the observation group (table 3). However, the difference noted for the 25-month interferon alfa group versus the observation group increased (table 3). In the multivariate analysis, the importance than stage alone. Adjustment for this combined variable, taking into account stage and number of lymph nodes, was of greater prognostic importance than stage alone. Adjustment for this variable in a Cox model did not greatly affect the difference in treatment effect noted between the 13-month interferon group and the observation group (table 3). However, the difference noted for the 25-month interferon alfa group versus the observation group increased (table 3). In the multivariate analysis, the interaction between this combined variable and treatment was not significant (p=0.18).

The combined variable, taking into account stage and number of lymph nodes, was of greater prognostic importance than stage alone. Adjustment for this variable in a Cox model did not greatly affect the difference in treatment effect noted between the 13-month interferon group and the observation group (table 3). However, the difference noted for the 25-month interferon alfa group versus the observation group increased (table 3). In the multivariate analysis, the interaction between this combined variable and treatment was not significant (p=0.18).

### Table 3: Effect of treatment on DMFI, DMFS, and survival by group after adjustment for stage of melanoma or for stage and number of lymph nodes

<table>
<thead>
<tr>
<th></th>
<th>DMFI Overall outcome at 4.5 years</th>
<th>DMFS Overall outcome at 4.5 years</th>
<th>Duration of survival Overall outcome at 4.5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13-month interferon alfa vs observation</td>
<td>25-month interferon alfa vs observation</td>
<td>13-month interferon alfa vs observation vs observation</td>
</tr>
<tr>
<td>All patients*</td>
<td>p=0.58</td>
<td>0.81 (0.65–1.01)</td>
<td>0.83 (0.67–1.03)</td>
</tr>
<tr>
<td>Treatment comparison adjusted for stage (n=356)</td>
<td>0.95 (0.76–1.18)</td>
<td>0.97 (0.78–1.20)</td>
<td>0.83 (0.67–1.03)</td>
</tr>
<tr>
<td>Interaction between treatment and lymph nodes‡</td>
<td>p=0.33</td>
<td>p=0.31</td>
<td>p=0.12</td>
</tr>
<tr>
<td>Treatment comparison by subgroup</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage IIb† (n=356)</td>
<td>0.78 (0.44–1.37)</td>
<td>0.76 (0.44–1.33)</td>
<td>0.58 (0.33–1.03)</td>
</tr>
<tr>
<td>Stage III (N1)‡ (n=353)</td>
<td>0.89 (0.52–1.54)</td>
<td>0.96 (0.56–1.64)</td>
<td>0.69 (0.40–1.20)</td>
</tr>
<tr>
<td>Stage III (N2)‡ (n=679)</td>
<td>1.01 (0.73–1.40)</td>
<td>1.02 (0.73–1.41)</td>
<td>0.92 (0.66–1.27)</td>
</tr>
<tr>
<td>All patients‡</td>
<td>0.94 (0.76–1.17)</td>
<td>0.96 (0.77–1.19)</td>
<td>0.82 (0.66–1.02)</td>
</tr>
<tr>
<td>Interaction between treatment and lymph nodes‡</td>
<td>p=0.72</td>
<td>p=0.72</td>
<td>p=0.25</td>
</tr>
</tbody>
</table>

Data are based on Cox model and presented in cells as HR (97.5% CI for overall adjusted comparisons and 99% CI for subgroup analyses); p value. *Adjusted for sex, tumour site, ulceration, and stage (0=IIb, 1=N1, 2=N2). †Adjusted for sex, tumour site, and ulceration. ‡Adjusted for sex, tumour site, and ulceration, and stage/number of positive lymph nodes (0=IIb, 1=N1 and 1–4 positive lymph nodes, 2=N2 and 1 positive lymph node, 3=N2 and 5 positive lymph nodes), 11 patients with in-transit metastases at randomisation excluded from these analyses.

---

**Figure 3: Kaplan-Meier curve of DMFS by group**

Articles

www.thelancet.com Vol 366 October 1, 2005 1193
The treatment comparisons remained unchanged when ineligible patients were excluded (data not shown), and when those who refused the observation-only intervention or did not start treatment with interferon alfa (in the treatment group) just after randomisation were excluded. In the 25-month interferon alfa group, among the seven patients with stage IIb melanomas who did not receive interferon, two (29%) relapsed and died.

As well as DMFI we assessed DMFS, time to first metastases, or death from other causes. In addition to the 760 patients who developed distant metastases, 17 patients died free of distant metastases: seven in the 13-month interferon alfa group, nine in the 25-month treatment group, and one in the observation group. The details of the causes of death are given later.

Figure 3 shows the DMFS by group. The 4.5-year estimate rates did not differ significantly across the groups (table 2). With the Cox model, the treatment differences adjusted for the initial stage of melanoma did not differ greatly from those shown in table 2: for 13-month interferon alfa versus observation the HR was 0.96 (0.77–1.21; p=0.64), and for 25-month interferon alfa versus observation the HR was 0.85 (0.68–1.05; p=0.08). Additional adjustment by sex, and ulceration and site of primary tumour, gave an HR of 0.83 (p=0.05; table 3). As for DMFI, we did subgroup analyses by stage of cancer for DMFS. As shown in table 3, the largest difference was noted between patients with stage IIb melanomas enrolled to the 25-month interferon alfa intervention and those enrolled to the observation group. Adjustments by stage and number of lymph nodes gave similar results (table 3).

Since distant metastases and death due to other causes are competing risks, we compared the three groups for each type of risk (table 2). Neither the 4.5-year cumulative incidence rate of distant metastases nor of death due to other causes differed significantly between groups. For every treatment group, the sum of the cumulative incidences of distant metastases and death due to other causes equals the complement of the DMFS rate (eg, in the 25-month interferon alfa group, 52.4%+1.6%=54.0%=100%–46.1%; table 2). The cumulative rates of distant metastases at 4.5 years (inverse of Kaplan-Meier curves) are only slightly higher than the cumulative incidences of distant metastases (eg, in the 25-month interferon alfa group, 52.8% is greater than 52.4%). 681 of 1388 patients died, 278 in the 13-month and 257 in the 25-month interferon alfa groups, and 146 in the observation group (figure 4; table 2). With the Cox model, we stratified the treatment differences for the initial stage, though results did not differ greatly: for 13-month interferon alfa versus observation, the HR was 0.97 (0.77–1.22; p=0.77), and for 25-month interferon alfa versus observation the HR was 0.85 (0.68–1.07; p=0.12). For consistency with the presentations of the two previous endpoints, subgroup analyses were done by cancer stage. As shown in table 3, in individuals with stage IIb melanomas, the difference between 25-month interferon alfa and observation was borderline significant.

We constructed a score based on a Cox model in which tumour stage, number of involved lymph nodes, sex, body site, and ulceration status were retained as the most important prognostic factors for DMFS. Considering 70 and 150 as cut-off points of this score, we formed three groups of 430 (score ≤150), 625 (score >70–149), and 322 (score >149) patients. We noted no significant increase in the HRs for 25-month interferon versus observation obtained in these three groups: DMFI 0.81, 0.76, 0.89 (p=0.64); DMFS 0.83, 0.79, 0.89 (p=0.74); and survival 0.74, 0.82, 0.95 (p=0.36).

Intermediate doses of interferon alfa are associated with side-effects (table 4). Treatment was stopped or interrupted because of side-effects in 87 of 539 (16%) patients in the 13-month interferon alfa group and in 108 of 539 (20%) patients in the 25-month interferon alfa group. Among the 195 patients who went off-study...
because of toxicity or refusal to continue, 131 (67%) had grade 3-4 toxicities (influenza-like syndrome [46%], gastrointestinal toxicity [39%], neurological toxicity [9%], liver [4%], local, or other [6%]). 50 (26%) had grade 2 toxicities (influenza-like syndrome [22%], gastrointestinal toxicity [0.5%], neurological toxicity [3%], local, or other [0.5%]), and the remaining 14 (7%) patients had other toxicities at a lower grade.

There were no deaths that could be directly related to treatment. Of the 17 patients who died because of causes other than metastatic melanoma, one died in the observation group because of postoperative complications after non-melanoma related surgery, seven died in the 13-month interferon alfa group (two from cardiovascular disease [CVD], five from other cancers), and nine died in the 25-month interferon alfa group (eight of CVD and one from other cancer).

Discussion

Our findings indicate that the adjuvant use of intermediate doses of interferon alfa over 1 or 2 years for the treatment of stage IIb/III melanoma has no effect on DMFI or overall survival. Patients treated with the lower dose of interferon alfa for 2 years had slightly better DMFI and overall survival rates than those who received no treatment, but this difference was not significant. Treatment at the higher dose for the shorter period had a negligible effect. These results are in keeping with the findings of a meta-analysis16 of interferon trials in melanoma, which show some benefit on disease-free survival but none on overall survival, and no clear dose-response effect.

Our results show that the effect of adjuvant interferon alfa is greater at an earlier stage of disease, with 25 months of treatment with interferon alfa having a borderline significant effect on stage IIb, some effect on stage III (N.), and no effect on patients with stage III disease with palpable nodal involvement (N.). This observation is in line with the previously reported effect of adjuvant therapy on disease-free survival with interferon alfa at a dose of 3 MU in three trials in Europe,1,2 and confirms the previously reported lack of response to low-dose treatment in stage III disease.1,10-11 In EORTC 18871,12 ultra low-dose treatment had no effect on either disease-free or overall survival.

Our data suggest that both duration of treatment and tumour load are important variables in the potential efficacy of therapy with interferon alfa, so another study (EORTC trial 18991) has been set up to which 1256 stage III patients have been randomised to 5 years of pegylated interferon or observation. This trial has completed accrual, but final results are not expected until 2006.

Grade 3-4 toxicities led to interruption or stopping of treatment in just over a tenth of patients in this trial, substantially less than the reported occurrence of grade 3-4 events in more than three-quarters of patients treated on high-dose interferon.1 Nevertheless, intermediate-dose interferon alfa causes enough symptoms of constitutional toxicity to make a substantial proportion of patients sick enough to refuse further treatment. Mood changes are often cited in connection with therapy and were observed also at the low doses of interferon used in this study. It is noteworthy, however, that signs of depression arose also in the observation group and could be part of the disease state. There were no cardiovascular deaths in the observation group, however, whereas there were two and eight such deaths in the 13-month and 25-month treatment groups, respectively. A possible association between interferon alfa and heart disease has not been reported previously and should be assessed when an individual patient meta-analysis of all trials of adjuvant interferon alfa is undertaken.

The findings of the meta-analysis by Wheatley and colleagues,16 those of the systematic review by Lens and Dawes,17 and the results of an analysis of the pooled data of four trials in high-density interferon5 all show an effect on disease-free survival but none on overall survival. The most prevalent current practice position in Europe, not to consider high-dose interferon as standard of care, is justified in view of the absence of consistent evidence of a long-term overall survival benefit and the high toxicity associated with high-dose interferon.18 Our results also indicate no benefit on DMFS or on overall survival in melanoma patients with stage IIb or stage III melanoma of intermediate doses of interferon given for either 13 or 25 months. They do though suggest that longer duration of therapy in early-stage disease could be beneficial.

Contributors

A M M Eggermont, S Suciu, and R MacKie wrote the report, with help from the other authors. A M M Eggermont coordinated the study and S Suciu did the statistical analyses. K Stoitchkov was the EORTC Melanoma Group coordinating physician, I Jagiello the data manager, and D J Ruter the pathologist. W Ruka, A Testori, W Krut, C J A Punt, M Delauney, F Sales, G Groenewegen, R MacKie, K Stoitchkov, U Keilholz, and D Lienard recruited patients.

Conflict of interest

We declare that we have no conflict of interest.

EORTC melanoma study group

Universiteit Gent, Gent, Belgium (V Cocquyt), Universitair Ziekenhuis Antwerpen, Edegem, Belgium (Prof J-B Vermorken), Cliniques Universitaires St Luc UCL, Brussels, Belgium (Prof Y Humblet), Hôpital Universitaire Erasme, Bruxelles, Belgium (T Velu), U Z Gasthuisberg, Leuven, Belgium (J Thomas), Hôpital Bichat, Paris, France (Prof B Crickx), Hôpital St Louis, Paris, France (J-Ollivaud), Centre Claudius Regaud, Toulouse, France (C Chevreau), Centre Antoine Lacassagne, Nice, France (Prof A Thys), Institut Gustave Roussy, Villejuif, France (M-F Avril), Institut Bergonie, Bordeaux, France (A Ravaud), Hospices Civils de Strasbourg, Strasbourg, France (C Chartier), Hôpital Cochin, Paris, France (J Gorin), Hôpital Saint Cloud, Saint Cloud, France (C Vilmer-Renaud), Centre Eugene Marquis, Rennes, France (T Lesimple), The Netherlands Cancer Institute AVL, Amsterdam, Netherlands (O Nieweg), Vrije Universiteit Medisch Centrum AZVU, Amsterdam, Netherlands (P Van Leeuwen), Leiden University Medical Centre, Leiden, Netherlands (S Osanto), Academisch Ziekenhuis Groningen, Groningen, Netherlands (H-J Hoekstra), Hospital Clinico y Provincial de Barcelona, Barcelona, Spain (T Castel Rodo), CHR Metz-Thionville, Thionville, France (F Truchetet), Turk University Central Hospital, Turk, Finland
Articles

(1) Pehamberger H, Soyer P, Steiner A, et al. Adjuvant interferon-
alpha in patients with regional node metastases from cutaneous

(2) Kirkwood JM, Strawderman MH, Ernstoff MS, Smith TJ.
Dacarbazine in metastatic melanoma: what have we learned in

(3) Kleeberg UR, Suciu S, Bröcker EB, et al. Final results of the
EORTC 18871/DKG 80-1 randomized phase III trial: rIFN-alpha2b
significantly prolongs relapse-free and overall survival
compared with the GM2-KLH/QS-21 vaccine in patients with

(4) Kirkwood JM, Manola J, Ibrahim J, Sondak V, Ernstoff MS, Rau O. A pooled analysis of eastern cooperative oncology group
and intergroup trials of adjuvant high-dose interferon for

a-2a as adjuvant therapy in resected primary melanoma thicker


(7) Cameron DA, Cornbleet MC, Mackie RM, et al. Adjuvant interferon

(8) Cascinelli N, Belli F, MacKie RM, Santinami M, Bufraino R,
Morabito A. Effect of long-term adjuvant therapy with interferon
alpha-2a in patients with regional node metastases from cutaneous

(9) Hancock BW, Wheatley K, Harris S, et al. Adjuvant interferon in
high-risk melanoma: the AIM HIGH Study—United Kingdom
Coordinating Committee on Cancer Research randomized study of
adjuvant low-dose extended-duration interferon Alfa-2a in high-risk

(10) Kleeberg UR, Suciu S, Bröcker EB, et al. Final results of the
EORTC 18871/DKG 80-1 randomized phase III trial: rIFN-alpha2b
vs rIFN-gamma vs Iscador M vs observation after surgery in

Acknowledgments

We thank St Jude Children’s Research Hospital for provision of an SAS
macro that allowed for the computation of the cumulative incidences of
distant metastases and of death without distant metastases, due to other
causes, and the corresponding treatment comparisons (Gray test). We
also thank the previous data managers of this study: A Kirkpatrick,
C Molabu. The National Cancer Institute funded the study (grant
codes numbers SU10 CA14188–26 through SU10 CA11488–35). The contents
of this report are solely the responsibility of the authors and do not
represent the official views of the National Cancer Institute (Bethesda,
MD, USA). Schering-Plough International also provided an educational
grant, and interferon alfa 2b free of charge.

References

1 Eggermont AMM, Kirkwood JM. Re-evaluating the role of dacarbazine in metastatic melanoma: what have we learned in

2 Eggermont AMM. Role of interferon-alpha in malignant
melanoma remains to be determined. Eur J Cancer 2001; 37:
2147–53.

3 Kirkwood JM, Strawderman MH, Ernstoff MS, Smith TJ.
Borden EC, Blum RH. Interferon-a2b adjuvant therapy of high-risk
resected cutaneous melanoma: the Eastern Cooperative Oncology

intergroup trial E1690/S9111/C9190. J Clin Oncol 2000; 18:
2444–59.

alfa-2b significantly prolongs relapse-free and overall survival
compared with the GM2-KLH/QS-21 vaccine in patients with

6 Kirkwood JM, Manola J, Ibrahim J, Sondak V, Ernstoff MS,
Rao U. A pooled analysis of eastern cooperative oncology group
and intergroup trials of adjuvant high-dose interferon for

7 Grob JJ, Dreno B, Chastang C, et al. Randomised trial of interferon
a-2a as adjuvant therapy in resected primary melanoma thicker


9 Cameron DA, Cornbleet MC, Mackie RM, et al. Adjuvant interferon

10 Cascinelli N, Belli F, MacKie RM, Santinami M, Bufraino R,
Morabito A. Effect of long-term adjuvant therapy with interferon
alpha-2a in patients with regional node metastases from cutaneous

11 Hancock BW, Wheatley K, Harris S, et al. Adjuvant interferon in
high-risk melanoma: the AIM HIGH Study—United Kingdom
Coordinating Committee on Cancer Research randomized study of
adjuvant low-dose extended-duration interferon Alfa-2a in high-risk
Excess bodyweight is the sixth most important risk factor contributing to the overall burden of disease worldwide. 1·1 billion adults and 10% of children are now classified as overweight or obese. Average life expectancy is already diminished; the main adverse consequences are cardiovascular disease, type 2 diabetes, and several cancers. The complex pathological processes reflect environmental and genetic interactions, and individuals from disadvantaged communities seem to have greater risks than more affluent individuals partly because of fetal and postnatal imprinting. Obesity, with its array of comorbidities, necessitates careful clinical assessment to identify underlying factors and to allow coherent management. The epidemic reflects progressive secular and age-related decreases in physical activity, together with substantial dietary changes with passive over-consumption of energy despite the neurobiological processes controlling food intake. Effective long-term weight loss depends on permanent changes in dietary quality, energy intake, and activity. Neither the medical management nor the societal preventive challenges are currently being met.

Hippocrates wrote “Corpulence is not only a disease itself, but the harbinger of others”, recognising that obesity is a medical disorder that also leads to many comorbidities. This association is profoundly important for the affected individuals, but the associated morbidity is also economically damaging for society. The number of deaths per year attributable to obesity is roughly 30 000 in the UK and ten times that in the USA, where obesity is set to overtake smoking in 2005 as the main preventable cause of illness and premature death.

WHO describes obesity as one of the most blatantly visible, yet most neglected, public-health problems that threatens to overwhelm both more and less developed countries. The problems of overweight and obesity have achieved global recognition only during the past 10 years, in contrast to underweight, malnutrition, and infectious diseases, which have always dominated thinking. WHO now accepts a body-mass index (BMI) in infectious diseases, which have always dominated thinking. WHO now accepts a body-mass index (BMI) in appropriate population samples, measured BMI in appropriate population samples, and individuals from disadvantaged communities seem to have greater risks than more affluent individuals partly because of fetal and postnatal imprinting. Obesity, with its array of comorbidities, necessitates careful clinical assessment to identify underlying factors and to allow coherent management. The epidemic reflects progressive secular and age-related decreases in physical activity, together with substantial dietary changes with passive over-consumption of energy despite the neurobiological processes controlling food intake. Effective long-term weight loss depends on permanent changes in dietary quality, energy intake, and activity. Neither the medical management nor the societal preventive challenges are currently being met.

Search strategy and selection criteria
Studies of interest were identified by systematic searches of MEDLINE and EMBASE for all 191 countries of the world with the keywords “BMI”, and “obesity” each paired with “cardiovascular disease”, “hyperlipidaemia”, “cholesterol”, “stroke”, “ischaemic heart disease”, “osteoarthritis,” “diabetes mellitus type 2”, “cerebrovascular disease”, and in combination with each country’s name. We contacted WHO Regional Officers for help with searches and governments and individuals in searches for unpublished data. Cochrane reviews, meta-analyses, and other systematic reviews were preferentially used.
Thus, the UK Government now estimates that a BMI of 25·0 kg/m² decreases the life expectancy of English men by 2 years and, given the progressive epidemic of obesity, the effect will increase to 5 years by 2050. What is not yet confirmed, however, is whether intentional weight loss in obese individuals prolongs life as well as reducing risks. Preliminary evidence suggests a 30–40% reduction in diabetes-related mortality with moderate (less than 10% of bodyweight weight loss). People with newly diagnosed diabetes who lost 10 kg in their first year of management were found to have gained a further 4 years of life.

Disease burden from excess weight in adults

Detailed estimates of the years of ill health and lives lost between the ages of 30 years and 75 years because of excess weight are shown for the subregions of the world in figure 2. These predictions are based on detailed estimates of the prevalence of various disorders and deaths from them, the prevalence of high BMI according to age, and the proportion of the disease burden attributable to the excess weight. Cardiovascular disease dominates, followed by diabetes and some cancers, especially in women. Again, the burden of disease is high in eastern Europe and Latin America, but the Asian countries have a surprisingly high burden in view of their lower obesity rates. This finding relates to the higher absolute risk of diabetes and probably cardiovascular disease among Asian, Hispanic, and perhaps African populations, partly because they are more prone to abdominal obesity with its excess risks.

Fat distribution

Many of the comorbidities of obesity are reflected in the so-called metabolic syndrome, originally defined arbitrarily by WHO on the basis of insulin resistance with other features of obesity or pragmatically in the USA on the basis of three of five features: large waist circumference, abnormal concentrations of triglycerides, HDL cholesterol, and fasting glucose, and hypertension. Lower waist circumference cut-off points for Asian populations have been used in Asian analyses of the metabolic syndrome, but now the International Diabetes Federation has proposed a universal system in which an ethnically specific waist circumference is the first requirement with abnormalities in two of the other four (triglycerides, HDL cholesterol, fasting blood glucose, and blood pressure) as in the latest criteria. The INTERHEART findings, from 52 countries on the predictive importance of the waist/hip ratio rather than waist or BMI measures alone, imply that the early emphasis on waist/hip ratios might have to be reapplied even though the waist measurement is simpler to use in clinical practice.

Currently, up to 30% of middle-aged people in more developed countries have several features of the metabolic syndrome. The prevalence is as high as 60% among individuals in the seventh decade of life. Only an estimated 30% of adults have no features at all. Insulin resistance is induced by fat deposited intracellularly and the secretory products of the expanded adipocyte mass, which is the body’s most prolific endocrine organ. These products include cytokines such as interleukins 1 and 6 and tumour necrosis factor α. The latter also has a paracrine suppressive effect on the secretion of adiponectin, a powerful insulin sensitisser which is secreted less as the adipocyte mass expands. The infiltration of fat into

Figure 1: Prevalence of obesity worldwide by age and sex

Derived from James and colleagues, where the full list of countries included in each subregion is the same as in the main WHO analysis.

older. Thus, the UK Government now estimates that a BMI of 25·0 kg/m² decreases the life expectancy of English men by 2 years and, given the progressive epidemic of obesity, the effect will increase to 5 years by 2050. What is not yet confirmed, however, is whether intentional weight loss in obese individuals prolongs life as well as reducing risks. Preliminary evidence suggests a 30–40% reduction in diabetes-related mortality with moderate (less than 10% of bodyweight weight loss). People with newly diagnosed diabetes who lost 10 kg in their first year of management were found to have gained a further 4 years of life.

www.thelancet.com Vol 366 October 1, 2005
the pancreatic islet cells amplifies the age-related decline in the islets’ capacity to maintain the increased insulin output demanded by insulin resistance, so glucose intolerance and premature type 2 diabetes readily develop.

Abdominal obesity accentuates the problem, probably because of the unusually high influx of portal fatty acids, cytokines, and hormones into the liver from omental adipocytes that normally are almost devoid of fat. The resulting distortion of hepatic metabolism includes increased synthesis of apolipoprotein B and VLDL and greater release of insulin to enter the general circulation. The effects of cytokines on the peripheral tissues with increased intracellular lipid also lower cellular insulin sensitivity; the surge in lipids promotes proliferation of the vasa vasorum of the arterial media and apoptosis by the medial macrophages, with a further release of cytokines. These changes help to explain the role of obesity as a promoter of intracellular inflammatory pathophysiological processes by inflammatory mechanisms with resulting arterial damage.28

**Hypertension**

The risk of hypertension is up to five times higher among obese people than among those of normal weight,29 the variability in response reflecting differential genetic susceptibility as well as dietary factors. Up to two-thirds of cases of hypertension are linked to excess weight,30 and cross-sectional population surveys31 suggest that more than 85% of hypertension arises in individuals with BMI values above 25 kg/m². The increase in blood pressure with excess weight gain arises partly because of the release from adipocytes of angiotensinogen (a precursor of angiotensin that has well-known effects on blood pressure), an increase in blood volume associated with the greater body mass, and in response to a rise in blood viscosity. The change in blood viscosity is induced by the release of profibrinogen and plasminogen activator inhibitor 1 from adipocytes with a fall in plasminogen activator.32

Diets conducive to weight gain independently amplify blood pressure. Dietary fats, especially saturated fats, induce a rise in systolic and diastolic blood pressures as well as hypercholesterolaemia, as shown in the Dietary
Approaches to Stop Hypertension (DASH) trials.33 Energy-dense diets rich in fats and refined sugars promote weight gain,14,15 and high sugar intakes also induce increases in blood pressure of 6-9 mm Hg (systolic) and 5-3 mm Hg (diastolic).36 Energy density is reduced by higher intake of fruit and vegetables, which the DASH trial also showed lowered blood pressure. The challenge, therefore, is to assess the contribution of weight gain as distinct from that attributable to dietary factors including salt.37 Data from the DASH trial (table) suggest that blood pressure can be lowered independently of weight change, especially in people with hypertension, and the overall effect is equivalent to that achieved with a reasonably potent blood-pressure-lowering drug. In adults in North Karelia, Finland, during a 15-year period in which vegetable consumption trebled there was a substantial fall in intake of total fat and saturated fat accompanied by a 15% decrease in total serum cholesterol concentrations38 and a substantial decline in salt intake. These changes were accompanied by an increase, rather than a decrease, in the average BMI of the population. The study was not randomized but the influence of these intervention studies on volunteers and an unwell population suggests that dietary changes are more important than weight loss in lowering blood pressure especially in people with hypertension.39 Avenell and colleagues’ Cochrane analysis39 could not distinguish between the effect of weight loss per se and the accompanying changes in diet leading to the weight loss.

**Coronary artery disease and strokes**

Dyslipidaemia progressively develops as BMI increases from 21 kg/m² with a rise in proatheromatous, dense, small-particle-sized LDL. This change increases the risk of coronary heart disease by 3-6 times. With low HDL concentrations, as well as high concentrations of triglycerides, CHD risk increases.40 The combined effect of dietary saturated and trans fatty acids on plasma lipids is amplified by the lack of n-3 long-chain fatty acids, which have complex competitive effects with the more pervasive n-6 polyunsaturated fatty acids on prostanooid synthesis, cellular function, and thrombosis.41 There is also an interaction with abdominal obesity; the influence of abdominal weight gain and external sources of infection as well as endogenous inflammation on the development of the metabolic syndrome, dyslipidaemia, and diabetes can now be quantified.42 Findings from the Asia-Pacific Cohort Collaboration Study involving 26–33 cohorts and more than 300 000 adults followed up for almost 7 years found for each unit change in BMI a 9% difference in ischaemic-heart-disease events and a change of about 8% in hypertensive deaths and ischaemic strokes.3

A surprising finding, given the clear relation with hypertensive deaths, was the closer relation of BMI to ischaemic than to haemorrhagic stroke. This finding could reflect the importance of weight-independent dietary factors such as salt in determining death rates from haemorrhagic strokes.

Left-ventricular hypertrophy occurs in 70% of women with both obesity and hypertension, and around 14% of cases of heart failure in women (11% in men) are attributable to obesity.43 The effect of obesity on heart function is probably due to a combination of factors including hypertension, dyslipidaemia, diabetes mellitus, increased fat mass and left-ventricular mass, endothelial dysfunction, and atherosclerosis.

These epidemiological inferences are paralleled by intervention studies, which have shown that weight loss improves the lipid profile as well as hypertension. Extensive Cochrane analyses44 suggest that a weight loss of 10 kg will induce a reduction in total cholesterol concentration of about 0·25 mmol/L (about 5%). Again, however, a distinction should be made between the immediate effect of weight loss and the longer-term effects of maintaining a lower weight by eating an appropriate diet. The observed hazards of weight loss in people with existing heart failure44 show the probable importance of further losses of lean body mass, including cardiac muscle, in older patients who have already replaced much of their lean body mass with fat.

**Diabetes**

The relation between obesity and type 2 diabetes is so close that Sims and co-workers coined the term “diabesity” in the 1970s, when they showed that in young men with no family history of diabetes who were overfed for 6 months BMI increased to 28·0 kg/m² and there were reversible rises in fasting concentrations of insulin, glucose, and triglycerides, and impaired glucose tolerance.45 Stevens and colleagues46 showed that around 90% of individuals who develop type 2 diabetes have BMI higher than 23·0 kg/m², the risk of diabetes being greatly increased by early weight gain, especially in childhood and in people with a family history of diabetes, with abdominal obesity, and whose mothers who had gestational diabetes. In Japan, the risk of diabetes is three times higher than 23·0 kg/m², the risk of diabetes being greatly increased by early weight gain, especially in childhood and in people with a family history of diabetes, with abdominal obesity, and whose mothers who had gestational diabetes. In Japan, the risk of diabetes is three times higher than

<table>
<thead>
<tr>
<th>Change in systolic blood pressure (mm Hg)</th>
<th>Change in diastolic blood pressure (mm Hg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normotensive</td>
<td>Hypertensive</td>
</tr>
<tr>
<td>Increase fruit and vegetables by 200 g/day</td>
<td>-0.8</td>
</tr>
<tr>
<td>Decrease fat intake by 10% of energy</td>
<td>-2.7</td>
</tr>
<tr>
<td>Decrease daily salt intake from 15 g to 4 g</td>
<td>-1.6</td>
</tr>
<tr>
<td>Total weight-independent dietary benefit</td>
<td>-7.1</td>
</tr>
<tr>
<td>Increase intake of free sugars to 152 g/day</td>
<td>6.8</td>
</tr>
<tr>
<td>Reduce weight</td>
<td>Average -6.11</td>
</tr>
</tbody>
</table>

Table: Measured and estimated changes in blood pressure from weight-independent dietary changes in relation to weight loss-induced changes in adults with normal or high blood pressure

*In adults with BMI >25 kg/m², †Data from the DASH trial (table) with some values inferred from the graphs. The sucrose effect is taken from Raben and colleagues’ paper34 and the effect of weight loss from Avenell and colleagues’ systematic review.43 ‡For loss of 10% of bodyweight. §For 10 kg loss.
Imprinting of metabolic control in fetal life and early childhood

Worldwide analysis of diabetes shows that four of the five countries with the most cases are in Asia,14 with the risks of diabetes increasing from very low BMI. Clinicians must therefore be more proactive and alert to the possibility of impaired glucose tolerance and diabetes even at BMI values of around 23·0 kg/m² when there is even slight abdominal obesity. In the most deprived areas of India, 14% of adults have diabetes and further 18% have glucose intolerance; on the basis of Chinese studies, the latter group have an annual probability of 11% of developing diabetes within 5 years.15

These ethnic differences seem to be imprinted by generations of fetal and postnatal malnutrition, combined with recent rapid childhood weight gain, which seems particularly conducive to the development of insulin resistance and the metabolic syndrome.16 In experimental models, poor maternal feeding leads to profound changes in development of hepatic and pancreatic tissue in the offspring with altered expression of various metabolic pathways and changes in telomere length accompanied by substantial differences in the lifespan of the progeny.17,18 These seem to be clear parallels with the concept of a disjunction in nutritional experience having profound effects in both animals and people.

Nevertheless, the development of diabetes is substantially preventable in both white and Asian people by small weight losses with dietary change and moderate exercise.19,20,21,22

Respiratory effects

People with pre-existing respiratory disease can be severely handicapped by weight gain: resting metabolic rates and movement costs are higher, but the physical effect of thoracic and abdominal fat restricts vital capacity and can be severely debilitating. Respiratory complications such as atelectasis and infection readily occur after anaesthesia. Whether obesity specifically induces bronchospasm is less clear, but overweight patients with asthma are further burdened, and their clinical condition can become evident only after weight gain, perhaps induced by steroids. The mechanical effects of bulky fatty tissue around the neck induce an obstruction to breathing, particularly during sleep, leading to sleep apnoea. A neck circumference of 43·0 cm or more in men or 40·5 cm or more in women is associated with episodes of disrupted breathing, recurring up to 30 times a night. Observers describe loud snoring, followed by a pause of 10 s or longer in breathing, then a loud grunt and resumption of normal respiration. About 3% of middle-aged people in more developed countries are affected, with a male to female ratio of four to one.23 Sleep apnoea can lead to pulmonary hypertension, right heart failure, drug-resistant hypertension, stroke, and arrhythmias, but the main risk is accidents caused by daytime somnolence, for example when driving.

Cancers and reproductive abnormalities

Obesity is one of the most important known preventable causes of cancer. About 10% of all cancer deaths among non-smokers are related to obesity. The WHO International Agency for Research on Cancer24 estimated that overweight and inactivity account for a quarter to a third of cancers of the breast, colon, endometrium, kidney, and oesophagus. The underlying mechanisms are difficult to define. Acid reflux, due to abdominal bulk, contributes to oesophageal cancer, and colon cancer has been linked to hyperinsulinism. Breast cancer seems to be related to the abnormally high concentrations of free oestrogen in postmenopausal obese women caused by peripheral conversion of sex hormones in adipose tissue by aromatase, together with a fall in the concentrations of plasma sex-steroid-binding globulin. These changes probably also explain the propensity to endometrial cancer and could be relevant to the suggested link between overweight and prostate cancer.

The excess oestrogen concentration also interferes with the feedback regulation of the hypothalamo-pituitary axis, disrupting normal reproductive function and causing irregular, commonly anovulatory cycles; the greater the degree of obesity, the more profound the effect on ovarian function, and obesity probably now accounts for 6% of primary infertility.25 In men, similar changes are now recognised as leading to impotence and increasing infertility, with abdominal obesity again a particular risk.26

These disturbances in sex hormones are also commonly accompanied in women by hirsutism and development of the polycystic ovary syndrome, characterised by substantial insulin resistance, androgen production from oestrogens partly resulting from the greater adipocyte aromatase activity in obesity. This poorly defined syndrome27 responds to weight loss28 and changes in intake of essential fatty acids,29 however, as well as treatments for the insulin resistance.

Obesity-related changes in hormone concentrations adversely affect pregnancy; the risk of admission to hospital is four to seven times higher than for non-obese women. The US Surgeon General’s obesity report30 highlighted an increase of three to ten times in the risk of pre-eclampsia, more common gestational diabetes, difficulties in labour and delivery, and higher rates of caesarean deliveries with more maternal and infant deaths. Infants are at greater risk of neural-tube defects and macrosomia.

Arthritis

That obesity leads to joint pain and arthritis of the knees and hips is not surprising, but the involvement of the carpometacarpal joints of the hand31 implies a metabolic
There is a strong connection between gallbladder disease, especially gallstones, and obesity, due to supersaturation of bile with cholesterol. In women, the risk is three times higher with BMI of 32·0 kg/m² or above and seven times higher with BMI of 45·0 kg/m² or above than in those with lower BMI. There is a particular risk in patients who lose weight rapidly; gallstone formation after bariatric surgery has been reported to affect 38% of patients.

**Psychological features of obesity**

Obesity was a sign of wealth and wellbeing in the past and still is in many parts of Africa, particularly since the HIV epidemic began. Care is needed to distinguish the social from the pathophysiological consequences of weight gain. In affluent societies and many Asian countries, slenderness is now the ideal, so individuals gaining weight, especially women, feel increasingly unacceptable and become anxious and depressed and can develop obsessive behaviours as they attempt to deal with their excess weight. Discrimination is rampant; obese individuals are less acceptable marriage partners, are handicapped in job promotions, and earn less.

In US women, obesity increases the risk of being diagnosed with major depression by 37%, whereas obese men have a 37% lower risk of depression than men of normal weight. In men, underweight is associated with significantly higher risks of depression and suicide, although whether the association is causal, or whether depressed men smoke more heavily, for example, is unclear.

Two eating disorders are linked with both depression and obesity: binge eating disorder (a subgroup of bulimia nervosa) and night eating syndrome (panel 1). These disorders affect a substantial proportion of patients attending obesity clinics; recognition of the characteristics is important, because psychological assessment and counselling are essential.

**Weight gain despite good physiological control of intake**

Despite the obesity epidemic, individuals have extraordinarily fine control of their food intake on a weekly if not daily basis. Although there are unpredictable variations in daily intake in response to social events, and smaller fluctuations in energy output from changes in physical activity, body energy stores remain fairly constant. A weight gain of 0·5–1·0 kg in a year amounts to 3500–7000 kcal (14·6–29·3 MJ), implying an error in the regulation of food intake of less than 0·5%. Average daily consumption therefore in the short, medium, and long term. These mechanisms are surpassed only by the multi-

---

**Panel 1: Criteria for common eating disorders**

**Binge eating syndrome**
- Large meals, eaten rapidly, without control
- Three or more of: rapid eating, solitary or secretive eating, eating despite fullness, eating without hunger, self disgust, guilt, depression
- Striking distress while eating
- If vomiting is part of the disorder, classify as bulimia
- No compensatory features—eg, excess exercise, purging or fasting
- > 2 days/week for 6 months

**Night eating syndrome**
- Evening hyperphagia; >50% of daily intake after evening meal
- Guilt, tension, and anxiety while eating
- Frequent waking and more eating
- Morning anorexia
- Consumption of sugars and other carbohydrates at inappropriate times
- Persistent for longer than 2 months

---

contribution. Hyperuricaemia and gout are well-recognised features of both weight gain and the metabolic syndrome.

**Non-alcoholic steatohepatitis**

The prevalence of non-alcoholic steatohepatitis is increasing rapidly in more developed countries as part of the obesity epidemic. It is set to become one of the most common causes of end-stage liver failure in more developed countries, because it progresses from benign fatty changes to cirrhosis, portal hypertension, and hepatocellular carcinoma. The changes in liver histology seen in alcoholic disease are also typical of non-alcoholic steatohepatitis, but the aetiological factors are obesity, diabetes, hyperlipidaemia, and hypertension. The disorder is generally asymptomatic, although some patients describe tiredness and abdominal discomfort; hepatomegaly occurs in up to 75% of patients, but other signs of liver disease are rare. The finding of raised concentrations of γ-glutamyl transpeptidase and alanine aminotransferase, and to a lesser extent aspartate aminotransferase and alkaline phosphatase, can be the first indication of non-alcoholic fatty liver disease; these findings can be associated with an abnormally echogenic white or bright appearance of the liver on ultrasonography. CT and MRI can show gross hepatic steatosis, but liver biopsy is the gold-standard diagnostic test, revealing identical features to alcoholic liver disease. The prevalence of non-alcoholic steatohepatitis in the general population is between 2% and 9%. 50% of patients with the disorder develop fibrosis and 30% cirrhosis, and 3% will develop liver failure and need transplantation.
faceted processes that come into play when people are deprived of food. Those processes enable deliberately overfed young but not older volunteers to return spontaneously to normal bodyweights over a few months.88

In the light of this surprisingly good control of energy balance, why do people become overweight and obese as they become middle aged? Genetic factors are well recognised to influence who gains weight, and the magnitude of weight gain, as shown by overfeeding studies in twins.89 Statistical analyses suggest that 50% or more of the variation between individuals in BMI has a genetic basis,90 but these effects are dominated by polygenic environmental interactions that reflect many genetic influences affecting spontaneous physical activity, twitchiness, basal metabolic rate, propensity to synthesise diurnally lean rather than fat tissues, and appetitive behaviour. Monogenic mutations, of leptin secretion or receptor activity for example, are very rare but single base changes in the gene for the hypothalamic melanocortin receptor, normally involved in appetite suppression, explain about 5% of obesity in children if it is severe early in life.91

These genetic influences cannot explain the population’s public-health problem of obesity. The adult phase of weight gain (figure 1) corresponds to a substantial fall in leisure-time sports for men. Women tend to gain weight once they cohabit and begin to share meals with men, who have intrinsically higher energy needs and commonly take more exercise.92 Oral contraceptives could provide further physiological and social conditions conducive to weight gain; repeated pregnancies certainly do so.93 The well-documented progressive fall in physical activity with age that the less effective mechanisms downregulating food intake are under severe strain as energy needs decline. Before major changes occurred in use of cars, mechanical aids, television, and computers in the 1960s to 1980s, the fall in total energy output from age 25 years to 75 years in the Baltimore ageing study amounted in men to 1200 kcal (5.02 MJ) per day. To avoid any gain in body energy would therefore have required a progressive fall in intake of about 270 kcal (1.13 MJ) daily, each decade, throughout adult life. Now the environment is deliberately designed to promote inactivity, even children are sedentary, especially when both parents work and they are confined indoors or at school.

Accompanying the documented secular and age-related declines in physical activity are changes in food habits that might originally have been responses to reduced energy needs. Social historians describe the three or four large meals a day taken by hard-working people, amounting to 3000–4500 kcal (12.5–18.8 MJ) per day to cope with physical demands at work and in the home. As working conditions and household aids apparently improved, meals became smaller; breakfast was omitted or reduced, sandwiches or single, smaller courses were eaten in the middle of the day, and the evening meal became the main meal of the day. In the 1960s and 1970s in more developed countries rates of overweight and obesity were of little immediate concern; however, by 1983 the potential public-health problem of obesity was being highlighted.94 What can explain the huge rise in obesity rates?

Physical inactivity

Many studies have shown the relation between sedentary lifestyle and weight gain, but reliable direct measures of physical activity are only just emerging.95 Nevertheless, the secular decline in physical activity is obvious. Morris and colleagues showed more than 50 years ago that vigorous exercise was crucial to cardiovascular health, but highly sedentary adults now derive benefit from even slight exertion.96 Exercise has many benefits, from psychological to physical, independent of its contribution to weight stability. However, the recent emphasis on weight maintenance has highlighted the importance of total energy output—60–90 min per day of walking.97 10 000 steps monitored on a pedometer, or 15 000 steps in individuals attempting to maintain weight loss. Such activity is difficult nowadays without redesigning cities to necessitate more walking and spontaneous movement. Gyms tend to be attended by more affluent and motivated individuals. Physical activity is helpful in weight loss, and essential for limiting the progressive decline in lean tissues with age, but its main importance in bodyweight is in maintaining rather than increasing a 5–10% weight loss.

Changes in daily food intake patterns

Short-term regulation of food intake is readily overcome by sudden increases in the energy density of food, for example by fat-rich evening meals that allow no compensatory adjustments until the next day.98 Sugar-rich drinks also circumvent the meal-based regulation of appetite.99 Foods with higher energy density—those rich in fats, extracted sugars, and refined starches—are unwittingly consumed in greater amounts, the density rather than the macronutrient content being the determinant of intake.100–103 Nevertheless, the urge to eat sugary and salty foods is driven by selective taste buds and neuronal projections to the limbic pleasure centres, and the combination of the fats and sugars, rare and precious in our early evolution, is especially alluring. When displayed in larger portions, the visual impact of foods with higher energy density dominates appetitive regulation in adults and children older than about 4 years so they consume more.104 Food companies have long known the commercial benefits of promoting larger portion sizes. Given the fixed energy requirements of a population, the only ways to promote sales involved provision of products with higher content of fats, sugars, and salt, in larger portions, making them available everywhere, and promoting drinking and eating on the move since this distracts the normal appetite regulatory responses.
Eating outside the home also restricts the ability to control the composition and quantity of food. Targeting of children from infancy to generate brand loyalty, which distorts dietary patterns, and expansion of sales to the huge potential markets of less developed countries are the only means seen by food companies for maintaining expansion, profits, and shareholder value. Contrary to initial estimates that the dominant factor precipitating the obesity epidemic in the UK was a decline in physical activity rather than excessive intake, recent evidence from secular trends and obesity rates in 36 countries shows rising intakes. Similarly, national studies of BMI of different groups show that intake is now the dominant determinant with lower physical activity following, rather than preceding, weight gain in some cases. Thus, a decline in activity was probably a particular feature of the 1960s to 1980s, but the transformation of our food habits in response to intense industry competition is now the main amplifier of the epidemic.

**Drugs**

An increasing number of drugs are now being documented as causing weight gain (panel 2).

**Assessment and management**

Despite the plethora of diet books and heavily promoted schemes for effortless and rapid weight loss, the escalating epidemic of obesity shows the failure of these approaches. The medical issue is now how to help transform patients’ lives in the long-term when they are constantly distracted and disheartened by the claims for miracle cures. Patients need to create a micro-environment as a buffer against the all-pervading toxic environment, and the greater the genetic contribution to the individual’s obesity the more abnormal their micro-environment must be.

Although there is plenty of evidence clearly proving the relation between obesity and disease, this relation is rarely apparent to affected individuals. An obese person’s health might not be as obviously compromised as that of someone with asthma or chronic pain, unless comorbidities have already developed. Most people are unaware of the underlying development of the sinister early signs of the metabolic syndrome, which helps to explain the lack of motivation for change of many obese individuals. Motivation depends on the acceptance and recognition that obesity is a medical disorder; since many clinicians do not, this is asking a lot of a patient. Recognition depends on improving the patient’s understanding, which also involves increased public awareness of obesity in a medical context and therefore depends on more coherent views being set out by government, the medical profession, schools, and the media as well as by the food, advertising, and retail industries. Until that happens, clinicians have to tackle the obesity problem one person at a time.

Many obese individuals are already being monitored in chronic disease clinics, but the preliminary assessment of the patient’s excess weight is commonly neglected. Immediate and complete assessment need not be undertaken in a busy time-constrained clinic by a stressed clinician but should be arranged for an early date. The assessment environment needs to be appropriate, friendly, and unthreatening with large enough chairs and suitable equipment such as large blood-pressure cuffs at hand. A full history should be taken, with particular attention to symptoms of comorbidities, such as sleep apnoea, that might be unrecognised. Emphasis should be given to a family history of diabetes, including gestational diabetes, and cardiovascular disease as well as the obesity itself. Successful and unsuccessful attempts at weight loss, a social history including work and leisure activities, and the availability of a support network is as important for long-term care as enquiries about smoking and alcohol intake. Motivation should be assessed because it is essential for a favourable outcome and can be encouraged in different ways. A new symptom or other triggers, such as the arrival of a baby or grandchild or the death or illness of a friend or relative, can precipitate a determination to cope with long-term weight management. The estimated proportion of people who are sufficiently motivated to accept treatment is believed to be less than 20%, and in many cases treatment is essential but needs to be set out as a facet of comorbidity management in patients who deny their weight problem. Efficient use of resources is to focus on individuals who are most motivated. Lack of motivation is a massive barrier to change. However, the presence of motivation is powerful and should be harnessed by continuing support, encouragement, and follow-up by a weight-management team, which needs to be developed for effective long-term care.

Clinical examination should be undertaken; height, weight, and waist circumference should be measured.

---

**Panel 2: Drugs that can cause weight gain**

Antipsychotics, especially olanzapine  
Antidepressants: tricyclics, selective serotonin-reuptake inhibitors, monoamine oxidase inhibitors, mirtazepine, lithium  
Corticosteroids  
Oral contraceptive and progestagenic compounds  
β blockers  
Oral hypoglycaemic agents: glitazones (peripheral rather than visceral gain), sulphonylureas  
Insulin  
Anticonvulsants: phenytoin, sodium valproate  
Antihistamines: many antihistamines, though weight gain is greater with older agents  
Pizotifen, used as a prophylactic migraine treatment
and BMI calculated. Hypertension should be excluded with the use of a large arm cuff for blood-pressure measurement. Simple investigations should be done to identify markers of the metabolic syndrome and comorbidities, to provide a baseline for future readings to map improvements, and to show to patients that their blood tests indicate no reason, hormonal or otherwise, why they should not lose weight.

Measurement of blood glucose after overnight fasting is essential; if the concentration is raised, further tests for diabetes mellitus will be needed, including glucose tolerance testing, measurement of haemoglobin A1C, and screening for microalbuminuria. Measurement of blood concentrations of lipids, particularly triglycerides and HDL cholesterol and total cholesterol, allows an objective calculation of the probability of cardiovascular events.36–48 Care is needed when extrapolating risk scores to patients with diabetes, for whom the UK Prospective Diabetes Study scoring system is more appropriate.49 The validity of these scoring systems in non-white ethnic groups (eg, south Asians100) is still uncertain, and adjustments to the scoring system might be needed.101

The value of simple measures of physical fitness rather than formal exercise testing should also be considered.102 Non-alcoholic steatohepatitis should be assessed by liver function tests and renal function by measurement of plasma urea and electrolytes. Thyroid function should be tested to exclude myxoedema and electrocardiography undertaken to detect possible left-ventricular hypertrophy. Other tests will depend on the individual as dictated by history and initial assessment. Chest radiographs might be appropriate, as well as screening for obesity-related cancers, hormone profiling in suspected polycystic ovary syndrome or infertility, and measurement of uric acid concentrations in serum in gout. Disorders such as sleep apnoea should be carefully investigated; the cardinal symptoms are too readily assigned simply to excess weight.

When the patient’s weight is stable, physical activity should be recorded, preferably with a pedometer, and intake assessed by mean of a systematic food diary—many more reliable guide to food habits than history taking. A non-judgmental approach is crucial to helping and negotiating with patients their options for long-term change. Symptom control for a related disorder and negotiating with patients their options for long-taking. A non-judgmental approach is crucial to helping far more reliable guide to food habits than history intake assessed by mean of a systematic food diary—a

**Dietary management**

Management of the diet is much neglected by doctors and even misinterpreted by dietitians if energy intake is based on dietary history. The weight conscious and the obese systematically underestimate intake. Intake is better predicted by estimation of the patient’s energy expenditure from their sex, age, weight, and crude classification of exercise patterns.104 This approach together with an individualised diet with an energy deficit of 500–600 kcal (2·09–2·51 MJ) is almost universally used in longer-term trials and has been identified in Cochrane analyses as one of the best options. A lower energy intake triggers the drive to eat, and a standard diet of 1000 kcal or 1200 kcal (4·18 MJ or 5·02 MJ) puts heavier patients under greater physiological stress.

Dietary quality is important;41 about 20% protein restricts the recognised inevitable loss of about 25% lean tissue that accompanies fat loss and helps satiety. Dietary benefits are amplified by daily intake of 400–600 g vegetables and fruits, with less than 20% fat, adequate n-3 fatty acids but the lowest possible amount of saturated fatty acids, less than 5% sugar, and fibre-rich carbohydrates; such diets also have lower energy density and greater bulk, which further improves satiety. Explicit guidance on transferring to a low-energy-density diet can double the quantity of food eaten and still achieve the energy deficit needed.

Patients are helped by avoidance of eating or drinking on their feet or while watching TV, thereby improving cognitive control of intake. Calorie counting is tedious and not very effective39 because few patients, let alone their doctors, know their true energy requirements. Monitoring with a simple diary the portion sizes, cooking habits, and the bulk of family purchases of vegetable oils, sugar, soft drinks, fast foods, and alcohol provides important insights for both the patient and the management team.

Lately, very strict diets such as the low-carbohydrate Atkins diet have become popular. They have been shown to have good effects on blood lipid concentrations, blood pressure, and glucose control. These effects are, however, generally short lived and not superior to standard approaches over the longer term.105,106–108 The degree of weight loss strongly depends on the ability of patients to adhere to their diets,109 and the more restrictive the regimen the greater the demand for intense discipline in the face of an intense physiological desire to eat. Meal-replacement therapy, in which two meals are replaced by a standard low-energy drink or meal during weight loss and one during weight maintenance, can succeed for some patients,110 but the recognised longer-term benefits of a low-energy-dense diet rich in the appropriate foods and nutrients are compromised. As with all dietary trials for weight loss and maintenance, the outcomes in terms of the main causes of death are still awaited, although the benefits of appropriate dietary interventions for delaying the onset of type 2 diabetes and improving the main contributors to cardiovascular ill-health are clear.

**Pharmacotherapy and surgery**

Objections to pharmacotherapy linger, stimulated by memories of cocktails of diuretics, thyroid extract, and amphetamines combined with barbiturates. These
concerns were fuelled by the withdrawal of fenfluramine and mixtures of ephedrine and caffeine,10,11 which has led to a rigorous demand for evidence of efficacy when obesity drugs are evaluated. The only agents currently accepted by most regulatory agencies on the basis of extensive data are orlistat and sibutramine; rimonabant is undergoing evaluation.112 These drugs in general increase by three to four times the proportion of patients achieving at least 5% weight loss at 1 year. They have other beneficial effects on blood lipid concentrations, blood pressure, and insulin resistance, which may exceed expected for the degree of weight loss achieved. However, these additional effects vary depending on each drug’s particular mode of action. The effects of orlistat and sibutramine have been dealt with extensively elsewhere.39,108,113–115

The criteria of the US National Institutes of Health or the European Union for the use of pharmacotherapy include a BMI of at least 27·0 kg/m² with a persistent comorbidity or a BMI of at least 30·0 kg/m². Asian medical groups propose lower BMI criteria reflecting their concern about higher rates of comorbidities at lower BMI in Asian populations. Phentermine, an analogue of dexamfetamine, on the market for decades, is permitted in the USA and elsewhere but was allowed back on the market in the European Union only after a legal challenge to the ban by the European Agency for the Evaluation of Medicinal Products (EMAE). The proposed restriction was based on the absence of long-term data on the efficacy and safety of phentermine, the latter being mainly based on post-marketing data. None of these drugs is a magic bullet to induce involuntary and substantial weight loss; they are most effective when used as ancillary therapy in a well-organised weight-management programme.

Surgical treatment is increasingly used,116 particularly in the USA, on patients with BMI of more than 40·0 kg/m² and those with severe comorbidity at BMI more than 35·0 kg/m². Laparoscopic adjustable banding of the stomach along with Roux-en-Y and other forms of gastric bypass are now favoured. In experienced surgical centres, the operative mortality is well below 1%, with average weight losses of 25–30% and rapid normalisation of glucose handling and blood pressure in patients with diabetes and hypertension.117 Long-term monitoring is needed, and patients can eat a nutritionally poor diet without fruit and vegetables. As yet there is only slight evidence of reduced mortality in long-term analyses of surgical treatment,118,119 but most patients feel transformed by the degree of weight loss. Schizophrenia, personality disorders, and uncontrolled depression are absolute contraindications for surgery and great care is needed in assessing the use of surgery in patients with eating disorders.

How are health-care systems going to cope with the obesity epidemic?

No health-service system has yet developed a useful strategy for managing the huge numbers of overweight and obese people in the community. Nursing, dietetic, and physical-activity expertise and collaboration with public and private community slimming groups are needed. The challenge of prevention as well as managing the millions already affected is overwhelming.120,121 The challenge to think in novel ways was also emphasised by the new WHO global agreement122 to develop strategies to deal with the burden of cardiovascular disease, cancer, and diabetes now being fuelled by the obesity epidemic.

The medical profession is only now waking up to the political and industrial challenges as well as the medical challenge. The industrial interests, with powers exceeding even those of the tobacco industry, are on the alert and often acting to slow the drive for change, by intense political lobbying at the highest level and by engaging in tactics well rehearsed by the tobacco companies. Our new scientific understanding of obesity is helping to validate a new approach to tackling the problem but the response of the medical profession to both its management and prevention is still at an early stage.

Conflict of interest statement

The National Obesity Forum receives or has received support for its activities from GlaxoSmithKline, Sanofi-Aventis, and Roche. DWH has received honoraria and expenses from conference organisers for individual lectures and advisory groups on obesity from the above organisations; he is an investigator for the SCOUT trial. The International Association for the Study of Obesity receives or has received support for its activities from Abbott, GlaxoSmithKline, Roche, and Sanofi-Aventis. WPTJ has received personal consultancies or honoraria and travel support from conference organisers for individual lectures and for chairing sessions on obesity from Abbott, GlaxoSmithKline, Johnson & Johnson, Pharmacia, Roche, and Sanofi-Aventis. Since August, 2002, he has chaired the Executive Steering Committee for the SCOUT trial on the effects of weight management on morbidity and mortality in patients with diabetes at high risk of cardiovascular disease.

References


Seminar

1208 www.thelancet.com Vol 366 October 1, 2005


Personal digital assistants in health care: experienced clinicians in the palm of your hand?

Daniel C Baumgart

Physicians and other health-care professionals are rapidly adopting personal digital assistants (PDA). Palm pilots and other hand-held computers are also increasingly popular among medical students. PDAs can be used for medical student education and physician training, daily clinical practice, and research. PDAs and their increasing integration with information technology in hospitals could change the way health care is delivered in the future. But despite the increasing use of PDAs, evidence from well-designed research studies is still needed to show how much these devices can improve the quality of care, save patients’ lives, and ultimately reduce health-care expenses. In this Review of PDA use in health care, the operating systems, basic functionality, security and safety, limitations, and future implications of PDAs are examined. A personal perspective and an introduction to medical PDA applications, software, guidelines, and programmes for health-care professionals is also provided.

Health care in the developed world is characterised by a rapidly increasing use of information technology in patient care, increasing documentation, coding and billing requirements, the desire and need to access a wealth of clinical and basic science data on the internet, a fast growing number of electronic medical publications and online supplements to paper journals, instant communication needs of mobile medical professionals, application of multimedia technology in academic teaching, ad-hoc scientific data collection, and intelligent scheduling.

Traditionally, these needs were addressed independently with separate devices, reference systems, and networks. Personal digital assistants (PDAs) are capable of changing how health care is delivered in the future, since they aim to merge and integrate this functionality in one device that is versatile, customisable, and portable. According to polls, the worldwide PDA market had 10·5 million devices in 2003.

Clinicians are rapidly adopting PDAs into their daily practice. In one study, more than half of all doctors younger than 35 years in developed countries used a PDA in 2003. In a survey from the University of California (San Francisco, CA, USA) 40–50% of all US physicians and junior doctors (also referred to as residents in the USA) use or can use a PDA. In 2005, the proportion of US doctors using PDAs is expected to be well above 50% and rising. This Review provides an overview of current PDA technologies, applications relevant to medical education and clinical practice, a guide to medical software, safety and security, a personal perspective, current limitations, and a future outlook.

Platforms and operating systems

In the past, hand-held computing was restricted to sophisticated programmable calculators with or without a data storage option. By comparison, most PDAs

Search strategy and selection criteria

I first searched in MEDLINE with the MeSH term “computers, handheld”. The National Library of Medicine (NLM) summarises the following entry terms under this MeSH: “computer, handheld”, “handheld computer”, “handheld computers”, “computers, palmtop”, “computer, palmtop”, “palmtop computer”, “palmtop computers”, “computers, palm-top”, “computer, palm-top”, “palm-top computer”, “palm-top computers”, “personal digital assistant”, “digital assistant, personal”, “PDA computer”, “computer, PDA”, “computers, PDA”, “PDA computers”, “palm pilot”, “palm pilots”, “pilot, palm”, “pilots, palm”, “pocket PC”, “PC, pocket”, “PCs, pocket”, and “pocket PCs”. For some subsearches, this MeSH term was combined with “evidence-based medicine” or “decision making, computer-assisted”, covering the terms “decision making, computer assisted”, “computer-assisted decision making”, “computer assisted decision making”, “medical decision making, computer-assisted”, and “medical decision making, computer assisted”. However, “computers, handheld” was only introduced in 2003 and now represents a subgroup of the microcomputer category (used to describe a broad spectrum of computers from 1992 to 2002). Therefore a second search was done of all non-indexed fields of MEDLINE for articles including the terms: “PDA”, “personal digital assistant”, “palm pilot”, “palm top”, “pocket computer”, “pocket PC”, “palm OS”, “Windows mobile”, “Windows CE”, “handheld computer”, “Smartphone”, “evidence based medicine”, “EBM”, “decision tool”, and “education”. Google was used with similar search strategies as MEDLINE. Medical journals, hand-held computer magazines, online user forums, technical data sheets, software manuals, websites of device and software manufacturers, user groups, and medical libraries of leading medical institutions, and PDA devices were also reviewed. Since current research in MEDLINE on hand-held computers is mainly comprised of surveys, uncontrolled experiments, and individual reports, no formal evidence grade selection criteria to base a systematic review on could be developed and applied. However, priority was given to articles that at least attempted a controlled, masked, multicentric design, had meaningful sample sizes, included a statistical analysis or other objective outcome measures, and were published within the past 5 years.
currently run on the mobile operating systems of either Palm OS (PalmSource Inc, Sunnyvale, CA, USA) or Microsoft Windows (Microsoft Corp, Redmond, WA, USA) that, in addition to their intrinsic functionality, allow customisation by the installation of third-party software applications. Furthermore, some Palm OS or Windows mobile-based PDAs have a Java (Sun Microsystems, Santa Carla, CA, USA) runtime that allows the use of platform-independent, Java-based applications. Other platforms such as Newton (Apple Computer, Cupertino, CA, USA), Psion (Psion Teklogix, Mississauga, ON, Canada), BeOS (PalmSource Inc), Symbian OS (Symbian, London, UK), and Blackberry (Research in Motion, Waterloo, ON, Canada) currently have no major role in the health-care market.

In 1996, Palm Inc introduced the Pilot 1000 and Pilot 5000 products running the Palm OS operating system (PalmSource Inc) that led the resurgence of hand-held computing. In 1999, the company added advanced wireless communications capabilities to the Palm OS platform to address the demand for mobile information appliances. Their company policy to provide registered developers with access to the source code of the Palm operating system led to the development of more than 40 000 software applications, to run on more than 36 million Palm OS devices sold, unmatched by any other hand-held operating system so far.4,5 Microsoft Windows mobile is Microsoft’s most recent operating system for hand-held devices. Its source code is proprietary and only available to professional-device and software manufacturers.6 Although Palm Inc still markets its own line of devices directly, both Palm OS and Windows mobile-based PDAs and smartphones (devices with a mobile phone and PDA combined) are also designed, manufactured, and distributed by several major computer manufacturers.

Basic functionality

PDAs are shirt-pocket-sized devices with a touch-sensitive screen, a dedicated input area or keyboard, customisable application buttons, and a multiway (button or mini joystick) navigator to browse information on the screen. Depending on the brand and model, some devices feature an expansion slot for memory cards or accessories, a built-in camera, headphone jacks, speaker, microphone, ports for infrared, Bluetooth, or Wi-Fi (Wireless Fidelity), and even built-in GPS (global positioning system) receivers.

PDAs are now generally equipped with a comprehensive suite of personal information management software or the option to integrate with common brands of such software, note-taking applications, and contact databases. PDAs can connect to desktop computers and wireless local area networks (W-LAN) using infrared, Bluetooth (first developed by Telefonaktiebolaget L M Ericsson, Stockholm, Sweden, now Bluetooth Special Interest Group [SIG], Delaware, DE, USA), or Wi-Fi communication technology. The desktop synchronisation software or additional add-on applications provide compatibility with popular office file formats. Most devices feature an e-mail application to integrate with current office suites, which allows users not only to carry critical files when travelling, but also to synchronise important files quickly and easily between desktop and hand-held devices.

Smartphones enhance the basic PDA functionality with wireless communication properties, including instant messaging, e-mail, web browsing, data synchronisation with remote servers and networks, and even video conferencing, if used in the coverage of commercial cellular telephone networks (figure 1).

Applications for health-care professionals

Physicians, nurses, dieticians, medical students and trainees, and other health-care professionals must review an ever-increasing amount of constantly changing information about their patients several times a day and correlate the data with the most recent diagnostic and therapeutic recommendations and management options to make sound decisions. Traditionally, health-care professionals consulted meticulously collected personal notebooks and article

Figure 1: Basic PDA functionality of a sample main application screen on Palm OS 6.1
cut-outs, white-coat-pocket manuals, subscription journals, medical reference books, or electronic references on desktop computers.

The wealth of information and its constant changes due to the accelerated pace in translational research in biomedical science mean that these traditional resources are very difficult to keep up to date. Fast approval and propagation of newly discovered therapies by regulatory agencies such as the FDA (US Food and Drug Administration) or EMEA (European Medicines Agency) can also lead to more frequent recalls of drugs, medical products, and devices (as well as newly issued warnings); labelling changes; and novel interactions with existing compounds. Additionally, with the advent of overzealous documentation, coding, and billing requirements in managed care, constantly overworked health-care professionals cause an increasing number of treatment and management errors, because the time available to spend with patients is sadly diminishing. PDAs can help to overcome some of these problems.

PDA use in medical student education

The education of medical students now relies heavily on computer technology, beginning with the replacement of animal experiments by computer simulations in basic science laboratories, multimedia study programmes and exercises, and the abolition of paper-and-pencil board examinations for fully computerised systems in the USA and other countries. PDAs fit very well with these concepts, and the fact that medical students were among the earliest adopters of PDA use is unsurprising.7

Many medical schools require students to acquire basic clinical skills in clerkships. Faculty staff and students generally complete lengthy assessment forms at the end of the respective rotation, which do not always allow for a timely feedback and balanced learning experience. Electronic records of patient encounter and procedure logs maintained by the students on their PDAs, which are synchronised with either a central database or the mentor’s desktop system, provide an interesting new approach. This concept has been assessed by several academic medical centres for rotations in internal medicine, family medicine, and emergency medicine in surveys. Medical students thought the logs were convenient to use. This system generally increased the number of patient encounters and recorded diagnoses, helped improve history-taking skills by alerting students to under-addressed issues such as women’s health, improved overall computer literacy, allowed to immediately identify large gaps in basic clinical skills, and provided an easy mutual feedback with faculty staff during clinical clerkships.8–14

The early use of a clinical management approach to evidence-based medicine is a worthy goal in undergraduate medical education. Two studies were undertaken to investigate whether PDAs could assist this approach at the point of care. In both studies, medical students were given PDAs preloaded with either university-developed clinical-decision support software (CDSS) or a bundle of commercial-decision support applications commonly used by clinicians. Multivariable regression analysis showed that improved perceived usefulness of PDAs with CDSS was associated with supportive faculty attitudes, good knowledge of evidence-based medicine, enhanced computer literacy skills. Greater satisfaction with the CDSS than with commercial-decision support devices was associated with increased use in a clinical setting and improved success in search rates.15 In the second study, pre-orientation and post-orientation questionnaires and a post-rotation assessment measured students’ comfort levels, and the perceived usefulness of PDAs with CDSS and ratings of programmes on their PDAs were analysed. PDAs almost always enhanced the clerkship experience, although the outcome measures were not as clearly defined as those in the first study.16

The education effectiveness of evidence-based-medicine learning was investigated objectively in a randomised controlled trial, in which students’ use of a PDA with CDSS was compared with the use of a pocket card containing guidelines and controls. Main outcome measures were factored and individual item scores from a validated questionnaire on personal, current, and future use of evidence-based medicine; use of evidence during and after the clerking of patients; frequency of discussions on the role of evidence during teaching rounds; and self-perceived confidence in clinical decision-making. The PDA showed significant improvements in all outcome scores, with the largest change in students’ educational experience with evidence-based medicine. No substantial deterioration was seen in the improvements even after the withdrawal of PDAs during an 8-week washout period, which suggested at least short-term sustainability of PDA effects.17

PDAs can also assist in telementoring and multimedia learning. Two studies have shown the feasibility of live wireless transmissions of laparoscopic surgical procedures to PDAs. One of these studies also compared the recognition of anatomical landmarks on PDA screens with that of standard computer monitors during the procedure and showed significant improvements.18–19 PDAs could also help enhance the classroom learning experience. In a pilot study, a histology class teacher polled the students about effectiveness, student interest, and comprehension with Bluetooth-equipped PDAs. End-of-class survey results indicated that students were enthusiastic about the polling device.20

Overall, current data lend support to the potential usefulness of PDAs in medical education. However, large randomised controlled trials with comparisons of PDA with non-PDA groups and with objective outcome measures, such as performance in in-house or board examinations, are needed to substantiate these early observations. Another important aspect of hand-held
PDA use in junior physician education

Several programmes for junior doctors at leading US academic institutions (such as Harvard Medical School, Boston, MA; Columbia College of Physicians and Surgeons, New York, NY; or Georgetown University Medical School, Washington, DC), have been early adopters of hand-held computers and provide their junior doctors with PDAs and software bundles.

Training programme accreditation authorities and medical specialty boards demand an ever-increasing documentation of patient exposure and procedural performance, to maintain and improve training standards. Apart from log cards, no simple and reliable mechanisms currently exist for directors of junior doctor programmes to assess how well their trainees are being exposed to teaching in their specialties and what curriculum weaknesses need to be addressed. Several studies in specialties such as anaesthesia, emergency medicine, family practice, general surgery, internal medicine, neurology, obstetrics and gynaecology, radiology, and urology, have demonstrated the usefulness of PDAs to simplify data collection and assess doctor and programme performance.23–28

A larger survey in junior doctors of six training programmes in family practice, internal medicine, neurology, paediatrics, radiology, and surgery concluded that, as advantages, many junior doctors readily adapted their personal organisers to help keep track of their clinical tasks and keep in touch with patients, and that commercial medical references were used most by the surveyed residents to answer immediate medical questions. The perceived drawbacks included: calculators and patients’ trackers were not clearly able to be tailored to residents’ needs (eg, to restrict and modify types of calculations to just those actually used), the physical size (both too small for display and too bulky overall), and several junior doctors mentioned a concern of becoming too dependent on one source of information, which was viewed as being too easy to lose or break. PDAs were widely used across the spectrum of specialties, irrespective of encouragement by the training programme.27 PDAs can also assist in assessing the performance of clinical educators and students in objective structured clinical examinations (OSCE).30–31

The available data suggest the potential usefulness of PDAs in junior physician education. However, as concluded for medical student education, larger randomised controlled trials and surveys are needed to compare PDA-assisted training with traditional training in institutions and specialties by use of objective outcome measures, such as performance in in-house or board examinations, to define the role of PDAs in postgraduate medical education.

PDA use in daily clinical practice

PDAs are widely used among health-care professionals across all major specialties. A study of 2130 paediatricians selected randomly from the American Medical Associations’ Physician Masterfile (American Medical Association, Chicago, IL, USA) aimed to calculate the percentage of paediatricians using PDAs, deduce the perceived strengths and weaknesses of PDAs, and explore characteristics associated with beliefs and use. The most commonly used applications were for drug reference (80%), followed by scheduling (67%), medical calculations (61%), prescription writing (8%), and billing (4%). PDA users were significantly more likely to be male, come from an urban community, have recently graduated from medical school, and work in non-private practice. Users were also more likely to believe that PDAs could reduce medical error, but often complained about memory capacity, although small screen size and system speed were not problems.22 With 35–40% of respondents using a PDA, this study is a good example for mainstream hand-held computer use by physicians in many clinical specialties.

Drug reference and treatment safety

PDAs can replace bulky drug reference books and help with the selection and comparison of drugs, identification of dosing schedules, and dose adjustment when drug excretion is impaired. A major advantage of PDA use over paper-based drug references are drug interaction checks and—if updated (synchronised) with an institutional or commercial server regularly—the most up-to-date drug information and immediate access to alerts or recalls from regulatory or government agencies, such as the FDA or CDC (Centers for Disease Control, Atlanta, GA, USA). The usefulness of PDA-based drug references, including parenteral nutrition, blood products, and chemotherapy, and drug interaction checks has been established in several different studies.33–39

The effect of PDA use on medication safety can be even greater if use is extended to nursing staff and combined with patient identification systems. To improve patient safety in hospitals by reducing drug treatment and treatment errors, the FDA has published a final rule about bar code label requirements for human drugs and biological products, in February, 2004.40 Bar codes are now required on most prescription drugs, blood, blood products, and specific over-the-counter drugs. This system begins when a patient is admitted to the hospital. The hospital gives the patient a bar-coded identification bracelet to link to his or her computerised medical record. As required by the FDA rule, most prescription drugs and specific over-the-counter drugs would have a bar code on their labels. The health-care team uses PDA-based bar-code scanners that are linked to the hospital’s computer system of electronic medical records. Before a health-care worker gives a drug to the
patient, the health-care worker scans the patient’s bar code, which allows the computer to access the patient’s computerised medical record. The health-care worker then scans the drug that the hospital pharmacy has provided for treatment. This scan informs the computer which drug is being given. The computer then compares the patient’s medical record with the drug being given to ensure that they match. Therefore some of the following problems (unfortunately not uncommon) could be easily avoided: wrong patient, wrong dose of drug, wrong drug, and wrong time to administer the drug. The technology is available and has already been implemented in some multisite facilities in the USA with some success.45–46

Patient scheduling, tracking, charting, and coding

Daily writing of progress notes with patients’ data interpretation, management plans, and coding of medical treatments and procedures are crucial clinician responsibilities. However, the quality and legibility of notes are often inadequate. The following studies illustrate how the quality of medical records can be enhanced with PDA use. In a paediatric critical-care unit, researchers recorded documentation discrepancies in 60% of daily-progress notes. Therefore, they undertook a before-and-after trial to determine whether a point-of-care, PDA-based patient record and charting system could reduce discrepancies in progress note documentation by junior doctors in a neonatal intensive-care unit. They recorded significantly fewer documentation discrepancies.47

Another randomised study investigated whether hand-held computer-based documentation could improve both the quantitative and qualitative aspects of medical records in orthopaedic surgery. The electronic documentation consisted of a specially designed software package on a hand-held computer for bedside use with structured decision trees for examination, access to a history, and coding. In the control group, chart notes were compiled on standard paper forms and were subsequently entered into the hospital’s information system. The number of documented ICD (International Classification of Diseases) diagnoses was the primary endpoint for sample size calculations. All patients’ charts were reread by an expert panel, which assigned quality ratings to the different documentation systems by scrutinising the extent and accuracy of patients’ histories and physical findings assessed by daily chart notes. Documentation with the hand-held computer significantly increased the median number of diagnoses per patients from four to nine, but it produced some over-coding for false or redundant items. Documentation quality ratings improved significantly with the introduction of the hand-held device with respect to the correct assessment of a patient’s progress and translation into ICD diagnoses. Various learning curve effects were recorded with different operators.48 These findings were confirmed by another orthopaedic surgery study in outpatients.49

A study among anaesthesiologists investigated their experience of using acute pain assessment software on a PDA for patient management. PDA assessments were more likely to contain documentation regarding pain and side-effects than paper assessments. The median time of the assessment period during the patient encounter was longer with the PDA than with paper; however, the median period for the total encounter time (chart review, assessment, documentation) was significantly shorter with the PDA than with paper.50

The battle between health insurers and physicians about claims is not over.51 Claims are frequently denied or delayed on technicalities such as over-coding or under-coding, which PDA use could help in the future. Many clinicians have difficulty determining the appropriate code for current procedural terminology (CPT) or evaluation and management (E&M) to assign to the type and intensity of patient care they provide. Several surveys reported PDA-based charge capture and billing programmes were more accurate than paper. The reimbursement advantage was estimated to be 20%.52–54

PDA-based outcomes research to improve quality of care

Quality assessment and outcomes research in large medical associations require the acquisition, analysis of, and response to point-of-care data. Although most hospitals now process much of their clinical and administrative data electronically, data acquisition from the actual care providers and patients during encounters are still accomplished with an intermediate paper process. PDAs have the potential to simplify and accelerate this. Several studies, particularly in procedure-oriented specialties, have shown feasibility and measurable benefits of PDA-based data collection, because they allowed the quick modification of the study design, rapid data acquisition, and processing, to enable immediate effect of the results on clinical and administrative daily practice. This type of data collection increased performance almost instantly. Data were obtained with PDAs from either providers or patients to assess patient-perceived outcomes.55–57

PDA-based decision support, reference systems, and information retrieval

Quality of care can be improved with the implementation of CDSS (figure 2),58 evidence-based medicine,59 or other critically appraised publications and with alerting systems in hand-held computers. In a survey of 1538 health-sciences faculty staff and junior doctors, most responders indicated that they would like to learn more about clinical resources for PDAs.60 Although many health-care professionals already rely on various sources of medical reference applications (figure 3),61 their effect on the quality of care is currently under-explored. Pilot studies in which users either assessed an interface to
access institution-provided, critically appraised topics or headlines delivered to their PDAs alerting them to new books, National Guideline Clearinghouse guidelines, Cochrane reviews, and National Institute of Health (NIH) Clinical Alerts, as well as updated content in UpToDate (UptoDate, Waltham, MA, USA), Harrison’s Online (McGraw Hill, Princeton, NJ, USA), Scientific American Medicine (now renamed ACP Medicine; American College of Physicians, Philadelphia, PA, USA), and Clinical Evidence. Participants could request additional information for any of the headlines, and the information that was delivered via e-mail during their next synchronisation was perceived as helpful.65–68

The Lister Hill National Center for Biomedical Communications (Bethesda, MD, USA), a research and development division of the National Library of Medicine (NLM) of the NIH, has undertaken a project to discover and implement design principles for point-of-care delivery of clinical support information. PubMed on Tap is an application for PDAs that retrieves MEDLINE citations directly from the PDA through a wireless connection to the internet. PubMed on Tap features include several PubMed search options, a history of previous queries, the ability to save citations to an electronic memo pad, two clustered results options, and links to journal websites.69

The National Cancer Institute (NCI; Bethesda, MD, USA), another NIH branch, has also recognised the need for new information delivery methods and is currently undertaking a research study that investigates how health-care professionals use cancer information on hand-held wireless devices. The AvantGo Enterprise 4.2 Solution (iAnywhere Solutions, Dublin, CA, USA) provided the platform to deliver the website content of NCI’s cancer information service (CIS) onto hand-held devices. Several obstacles still need to be overcome before this service will be available to the general public.70

Other clinical settings where PDA-based decision support devices have been reported to be useful or advantageous include: emergency and mass casualty triage, data management of transplantation patients, management of patients with stroke, infection control, and enforcement of institution-specific, rational antibiotic use.71–79 Although these concepts undoubtedly have potential, no study so far has compared this approach with existing methods of information delivery or performance of users in board examinations or re-certifications.

PDA use to educate and interact with patients

Most patients feel comfortable with their physicians using a PDA in daily clinical practice.80 However, their use is not restricted to health-care providers. PDAs can serve as electronic patient diaries and prediction devices in diseases that are intermittently flaring, such as asthma or urticaria. The successful use of PDAs in diabetes care to improve glycaemia in patients with insulin pumps has been reported. PDAs can also help migraine patients to predict attacks.81–88 The new use of PDAs in patients has also been recognised by government agencies such as the US Public Health Service (USPHS), which released an interactive programme for Palm PDAs to help patients quit smoking. The programme is distributed through the Agency for Healthcare Research and Quality (AHRQ) and is available on their website.89 In addition to these professional applications, the internet is replete with software of the fitness, wellness, and personal health-care categories, such as menstrual calendars, diet, weight, calorie, and workout management applications, among others.

Other evolving applications and uses in clinical practice

PDAs could help patients with brain dysfunction or injury as cognitive-behavioural orthoses.90,91 A frequent outcome in these patients is memory impairment. One group of researchers designed and tested a mobile-distributed care system in a cognitive neurology day-care clinic of an academic medical centre.92 A PDA-based speech synthesiser for speech-impaired patients has also been reported.93 With an extended bandwidth of cellular
telephone networks (eg, universal mobile telecommunications system or UMTS) and high-speed institutional wireless networks, teleradiology on hand-held computers may become a reality. Pilot studies have shown promising data, such as CT scans that have been transmitted in the industry standard format of DICOM (digital image and communications in medicine) and that have been assessed remotely by radiologists. Echocardiograms have also been successfully read on PDAs.94–96

PDA use for data collection and processing in research
International, randomised, multicentre clinical trials usually need the collection, storage, and processing of large amounts of data. Data collection by investigators and study coordinators is traditionally done with specifically designed paper forms in clinical research files or complex telephone interview systems. Most trials also need the repeated completion of patient questionnaires to calculate standardised disease activity or quality-of-life scores. Unfortunately, paper-based, self-administered instruments remain inefficient for data collection because of missing information, respondent error, and slow data analysis due to processing delay from paper-to-computer file conversion. The advantages of PDAs to improve trial efficacy, quicken data analysis, and even improve patient safety due to earlier availability of results of interim analyses, among others, are obvious. Text and photo data capture, transmission feasibility, and visual analogue scales have been validated. 57,97–100

PDA appliances can record, store, and transmit virtual electrocardiograms and electrochemical data.101,102 There are comprehensive PDA-based data recorders that, in combination with a sensor vest, continuously encrypt and store patients’ physiological data (ie, blood pressure, blood oxygen saturation, electroencephalograms, electro-oculograms, periodic leg movement, core body temperature, skin temperature, end tidal CO₂, and cough) on a memory card. Patients could also record time-stamped symptoms, moods, activities, and other endpoint-specific information in the recorder’s digital diary. These features allow researchers to correlate multiple physiological indices that can be objectively measured with subjective input.103

Clinical research organisations have already discovered the advantages of PDA-based data collection in clinical trials. One such organisation and a major PDA manufacturer reported record sales of customised electronic diaries in 2004. This clinical research organisation has deployed 40 000 electronic diaries in 46 languages to 48 countries for use in clinical trials since 2000.104

How and where to find medical software for hand-held computers
Several thousands of medical software applications and documents are available for health-care professionals to use. Medical software can be grouped into major categories: standard medical textbooks and manuals adapted for PDAs, PDA-designed medical references, medical dictionaries, drug reference and interaction check programmes, medical calculators, medical prediction rule applets (a Java software component), document readers, medical image viewers, software for medical evidence retrieval, subscription platforms to electronic newsletters or journal digests, educational programmes for medical students, and medical alerting messaging; comprehensive medical enterprise solutions integrating with electronic medical records, patient management and scheduling systems, and electronic order, prescribing and pharmacy-dispensing systems, coding, billing and file-sharing. Software and content are available from commercial suppliers, shareware and freeware distributors, health-care organisations, and PDA enthusiasts.

The quality of medical software applications varies greatly and depends heavily on accessibility of the information. Initially, most suppliers offered static translations of traditional textbooks that were difficult to navigate. The market has now become more
sophisticated, demanding more dynamic content with frequent updates taking advantage of the implementation of wireless networking protocols in PDAs. Internet websites are available to link users to sites dedicated to medical PDA use (table 1), major medical software suppliers (webtable 1), guidelines from professional societies or health-care agencies (table 2), and helpful programmes for general PDA use (webtable 2). Additionally, some medical journals such as the Journal of the American Medical Association (JAMA) regularly announce and discuss novel hand-held computer software titles (figure 4).

**A personal perspective**

As an internist and gastroenterologist, I face the same challenges that all academic physicians do: attending on the wards, clinics, critical care units, and emergency rooms; doing consultations for other specialties; dealing with numerous conferences, administrative work, lecturing, and bedside teaching; being an investigator in clinical trials; mentoring doctoral students; and running a basic science research laboratory, which often hardly fit into those 24 hours, unless one is very organised. I perceived the arrival of the Palm Pilot (Palm Inc, Sunnyvale, CA, USA) in 1996 as a blessing; it quickly changed how I organised my day, kept abreast of the ever-changing specialties of medicine and biomedical science, obtained and accessed medical information, and taught students. My old spiral notebook is retired now.

On a typical day, my PDA wakes up 30 minutes before I do, logs on to my notebook as well as the internet, and synchronises and updates all PDA applications. Not only are contacts, appointments, and medical references kept up to date in this way, but my PDA e-mail application is also programmed to retrieve exclusive e-mails such as electronic tables of contents from medical journals, alerts from the FDA Medwatch system, and other resources in my e-mail inbox. On the way to work I can review, mark, and erase these e-mails. Once at the hospital, my PDA reminds me of conferences, meetings, and displays a to-do list for the day. When I see patients, I rely on drug reference and interaction applications, institutional microbial spectra databases, medical calculators, prediction rules, and specific topics in PDA editions of popular medical reference manuals. Additionally, I have many guidelines from our institution, professional organisations, and agencies, as well as pdf excerpts from journal articles stored in my memory card. I do not believe PDA versions of large medical textbooks are helpful, because they are often difficult to navigate. I am currently investigating the usefulness of the new PubMed on Tap programme, whenever I have access to the institutional W-LAN.

Our department receives a fair amount of patients with gastrointestinal cancers. Staging of uncommon cancers is easy with a TNM (tumour, node, metastasis) staging programme. I can customise and print actual chemotherapy protocols with a shareware application. An add-on to this shareware application allows me to programme and print protocols for rare cancers. At ward rounds with students, I take full advantage of the multimedia capabilities of my PDA: I can display images from my personal medical image library or other PDA reference materials, play heart murmur or lung sound recordings, and use the screen to quickly sketch something to make a teaching point. I can carry and share with students (via infrared) a self-created collection of text notes, customised to the patients we see together.

In clinics where I see many patients enrolled in clinical trials, I quickly enter, access, and sort essential data on spreadsheets. The spreadsheets were created with my notebook spreadsheet application, transferred to and updated on my PDA with a commercial programme.
This software also helps me to review and store my presentations for lectures and talks. New versions of PDAs can also act as USB memory sticks. At the end of the day, my PDA synchronizes and backs up the day’s data with my notebook before it charges for the night (webtable 1).

**PDA safety and security**

Information recording and interchange always raise the question of security and privacy. Overall, PDA security hazards are probably similar to other computers used in hospitals and elsewhere. Catastrophic data loss can only be prevented with regular backups. PDA viruses have been reported for the mobile operating system from Microsoft Windows and also, to a lesser degree, for the Palm operating system. Major security firms are addressing this problem with the development of commercial antivirus products for hand-held devices.

In the USA, PDA-based patient data processing and storage must comply with the Health Information Portability and Accountability Act (HIPAA) of 1996. The Centers of Medicare and Medicaid Services (CMS) of the

<table>
<thead>
<tr>
<th>Website</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.ahrq.gov/clinical/">http://www.ahrq.gov/clinical/</a></td>
<td>The ACC has contracted Skycape to convert its guidelines into a PDA accessible format. The application can be downloaded for free from the ACC website.</td>
</tr>
<tr>
<td><a href="http://www.chestnet.org/education/guidelines/currentGuidelines.php">http://www.chestnet.org/education/guidelines/currentGuidelines.php</a></td>
<td>The seventh ACCP conference on anthrscopic and thrombotic therapy. ACYP/ACCP management of acute exacerbations of COPD algorithm, guidelines for diagnosis and management of lung cancer, assessment of diagnostic tests for ventilator-associated pneumonia, guidelines for weaning and discontinuing ventilatory support, and a guideline on cough management as a defence mechanism and as a symptom, as well as pulmonary rehabilitation are available for download.</td>
</tr>
<tr>
<td><a href="http://www.qualitymeasures.ahrq.gov/">http://www.qualitymeasures.ahrq.gov/</a></td>
<td>The AHA distributes its guidelines through Apprisor.</td>
</tr>
<tr>
<td><a href="http://aidsinfo.nih.gov/mobile/">http://aidsinfo.nih.gov/mobile/</a></td>
<td>The US Department of Health and Human Services offers a programme for possible sexual, injecting-drug-use, or other non-occupational exposure to HIV including considerations related to antiretroviral therapy.</td>
</tr>
<tr>
<td><a href="http://www.hin.nhlbi.nih.gov/as_pal.htm">http://www.hin.nhlbi.nih.gov/as_pal.htm</a></td>
<td>The National Heart, Lung, and Blood Institute (NHLBI) offers evidence-based applications for asthma treatment its own website.</td>
</tr>
<tr>
<td><a href="http://www.medpalm.com/index.cfm?fuseaction=software.showsoftware&amp;PartnerREF=+&amp;prodID=7862">http://www.medpalm.com/index.cfm?fuseaction=software.showsoftware&amp;PartnerREF=+&amp;prodID=7862</a></td>
<td>The NCI Cancer Information Service offers one of the most comprehensive cancer databases. It contains peer-reviewed summaries on cancer treatment, screening, prevention, genetics, and supportive care, and complementary and alternative medicine; a registry of about 2000 open and 13 000 closed cancer clinical trials worldwide; and directories of physicians, professionals who provide genetics services, and organisations that provide cancer care.</td>
</tr>
<tr>
<td><a href="http://www.guideline.gov">http://www.guideline.gov</a></td>
<td>Public resource for evidence-based clinical practice guidelines. NGC is a collaborative initiative of the US Agency for Healthcare Research and Quality (AHRQ) and US Department of Health and Human Services. NGC was originally created by AHRQ in partnership with the American Medical Association and the American Association of Health Plans. All NGC summaries are available in a text format and downloadable to PDAs.</td>
</tr>
<tr>
<td><a href="http://hin.nhlbi.nih.gov/obgdpalm.htm">http://hin.nhlbi.nih.gov/obgdpalm.htm</a></td>
<td>The NHLBI offers evidence-based medicine applications for asthma treatment, cholesterol management, and obesity education in its own website. Through Apprisor the NHLBI distributes the joint national committee on management of hypertension (JNC 7) report reference card and the NCEP ATP III quick reference.</td>
</tr>
<tr>
<td><a href="http://www.qualitymeasures.ahrq.gov/about/pdadmindownload.aspx">http://www.qualitymeasures.ahrq.gov/about/pdadmindownload.aspx</a></td>
<td>The National Quality Measures Clearinghouse (NQMC), sponsored by the AHRQ, US Department of Health and Human Services, is a database and website for information on specific evidence-based health care quality measures. NQMC is sponsored by AHRQ to promote widespread access to quality measures by the health-care community and other interested individuals. Brief summaries of all measures can be viewed and downloaded in various formats, including PDA-compatible formats.</td>
</tr>
<tr>
<td><a href="http://www.cdc.gov/nchstp/tb/pubs/PDA_TBGuidelines/PDA_treatment_guidelines.htm">http://www.cdc.gov/nchstp/tb/pubs/PDA_TBGuidelines/PDA_treatment_guidelines.htm</a></td>
<td>The CDC publishes a PDA guide on the management of tuberculosis and yearly updated guide on the management of sexually transmitted disease. All resources can be downloaded for free of charge.</td>
</tr>
</tbody>
</table>

**Table 2: Professional society and health-care agency guidelines**
US Department of Health and Human Services provide information technology professionals and the general public with extensive resources to address the issues on their information security programme website. In addition to a general policy manual, the CMS has outlined fundamental regulations as well as system architecture and security requirements for the acquisition, storage, management, and transmission of patient data. As a general rule, these policies are developed to provide a defence-in-depth security structure, along with a least-privilege approach and a need-to-know basis for all information access. Developers can download security threat identification resources based on their occurrence and importance in the current CMS environment. Before approval, applications have to pass the contractor assessment security tool (CAST) test to record their compliance with the CMS.

Several specific new risks and vulnerabilities arise with wireless networks. Bluejacking (ie, unauthorised accessing of Bluetooth-enabled devices) in airports and other public places is advancing to a new hacker sport. These problems need to be addressed by hardware and software makers with improved encryption and authentication technology. Currently, no evidence shows that wireless-enabled PDAs interfere with the functioning of implanted cardiac pacemakers or defibrillators.29,106–108

**Challenges of current PDA technology and future outlook**

Evidence of PDA use and dominance in medical education, clinical practice, and research is still evolving. Most studies available so far have not been randomised, controlled, or are multicentric in design. The fact that physicians can carry an entire shelf of medical reference textbooks on a hand-held computer’s memory card does not automatically mean that physicians know their contents or can apply their knowledge appropriately in clinical practice. The increasing incidence of the so-called palmomental reflex by residents and medical students should remind clinical educators that PDAs are not peripheral brains and are a poor substitute for ad-hoc clinical knowledge.109

At a time when governments, health-care organisations, and insurers worldwide cannot stop entertaining the themes of necessity assessment, cost saving, and down-sizing, we need convincing arguments that the extra expenses of investment into PDA technology can actually improve quality of care, save lives, and ultimately save health-care costs.110 The IT industry has recognised health care as the next big market.111,112 It will be up to health-care professionals who depend on PDAs to inform PDA manufacturers of users’ true needs, do the necessary research, and actively direct the development of new hardware and software.

The future of information exchange in medicine is digital and wireless.113 What will a medical PDA look like in 2015? It will probably be housed in a ceramic or lightweight alloy case, and hopefully be no larger but substantially lighter than current shirt-pocket-sized devices. New semiconductor technology will allow hand-held computers to be equipped with processors that can handle much more work than the best desktop systems that are currently available, while consuming less power to extend battery life. Memory will no longer be an issue, because data will be mainly kept in network storage systems. Manual data entry is still a problem in current versions of PDAs. In the future, authorised, secure logons to the PDA and data entry will be done with combined speech and fingerprint recognition by sophisticated audio hardware and a new high-resolution generation of touch-sensitive screens. Graffiti 2 (PalmSource Inc) characters will be further developed into true handwriting recognition. Speech processing will also be a reality, replacing many dictation methods currently used.

Very high network speeds will provide immediate access to clinical and administrative data, including imaging information such as procedural movies; three-dimensional ultrasonography; CT, MRI, or PET scans; histological slides; microbial cultures; and institutional and remote reference systems at any place and time. Medical applications will go beyond organisation and
storage of information. PDAs could evolve into expert systems that access information from many sources (ie, classic textbook style references, data from basic and clinical research and genome scans, environmental and public-health information, and results from ongoing clinical trials, match the information with the patient’s medical records from current or past admissions or visits, apply prediction rules, calculate clinical equations, and integrate all the data into an overall information package for clinicians. Users will be able to obtain and share opinions on patients with colleagues and international experts with ad-hoc medical multimedia conferencing.

PDA-based medical information management could even have an environmental effect that goes beyond paper-saving. The environmental effects of two applications of wireless technologies were compared with those of conventional technologies. Compared with the use of a newspaper, users receiving the news on a PDA resulted in the release of 32–140 times less CO₂, several orders of magnitude less NOₓ and SO₂, and the use of 26–67 times less water than the use of newspapers. Wireless teleconferencing resulted in one to three orders of magnitude less CO₂, NOₓ, and SO₂ emissions than those from business travel.

Is this future scenario widely off the mark? Perhaps so, but critics should remember that many theoretical predictions of the future have inspired the design of devices used today. However, it is still certain that no computer system can ever replace dedicated, experienced clinicians and their empathic interaction with patients and families.

Conflict of interest statement
I declare that I have no conflict of interest.

Acknowledgments
I would like to thank Palm, Inc., and Epoocrates, Inc. for providing information about their products and technologies, and Epoocrates, Inc. for providing market research data.

The Palm OS Cobalt Simulator was used with permission to create the application screen shots. The Palm OS Cobalt Applications Menu Screenshots are copyrighted materials of PalmSource, Inc. PalmSource, HotSync Graffiti and Palm OS are trademarks of PalmSource, Inc. or its licensors or affiliates in the United States, Germany, the United Kingdom, Japan, France and other countries. The trademarks and copyrighted images may not be used in connection with any product or service that does not belong to PalmSource, Inc., except as expressly authorized by PalmSource by express written license. All rights reserved. All other brand and product names are or may be trademarks of, and are used to identify products or services of, their respective owners.

No funding was given for writing of this review.

References
1 Medicine on the move. PDAs and tablet PCs make the rounds with doctors and nurses. Postgrad Med 2004; 115 (suppl: i–vi.
2 Martin S. More than half of MDs under age 35 now using PDAs. CMAJ 2003; 169: 952.
13 Denton GD, Williams RW, Panagaro L. Core problems reported by students in a palm OS and Internet-based problem entry system predicts performance on the third-year internal medicine clerkship. AMIA Anna Symp Proc 2003; 827.
45 Turner CL, Cashard AC, Murphy MF. Barcode technology: its role in increasing the safety of blood transfusion. Transfusion 2003; 43:1200–09.
74 Barschak B. How PDAs can speed up the organ transplant process. Nursing 2004; 34:68.


Genetic Epidemiology 4

Shaking the tree: mapping complex disease genes with linkage disequilibrium

Lyle J Palmer, Lon R Cardon

Much effort and expense are being spent internationally to detect genetic polymorphisms contributing to susceptibility to complex human disease. Concomitantly, the technology for detecting and genotyping single nucleotide polymorphisms (SNPs) has undergone rapid development, yielding extensive catalogues of these polymorphisms across the genome. Population-based maps of the correlations amongst SNPs (linkage disequilibrium) are now being developed to accelerate the discovery of genes for complex human diseases. These genomic advances coincide with an increasing recognition of the importance of very large sample sizes for studying genetic effects. Together, these new genetic and epidemiological data hold renewed promise for the identification of susceptibility genes for complex traits. We review the state of knowledge about the structure of the human genome as related to SNPs and linkage disequilibrium, discuss the potential applications of this knowledge to mapping complex disease genes, and consider the issues facing whole genome association scanning using SNPs.

Genomic approaches to disease association mapping

Genomics is transforming epidemiology, medicine, and drug discovery,1-3 and attention is being directed towards population-based genetic association studies for complex phenotypes.4-12 For many complex conditions, the genetic basis of susceptibility to disease, disease progression and severity, and response to therapy has been increasingly emphasised in medical research, with the ultimate goal of improving prevention, diagnosis, and treatment.13,14

Completion of the human genome sequencing project has been followed by three advances that provide novel opportunities for understanding the pathogenesis of common diseases:15-17 (1) compilation of extensive catalogues of DNA sequence variants across the human genome (polymorphic loci);18-20 (2) more rapid and cheaper molecular genetic techniques for investigating polymorphic sites; and (3) increasing availability of large, population-based samples such as the European Prospective Investigation into Cancer and Nutrition, the International Study of Infarct Survival,19 and the Million Women Study.20 Large, national cohorts (eg, the UK Medical Research Council and Wellcome Trust Biobank) are attracting funding bodies in many countries. Although the genomics revolution and the generation of high-density single nucleotide polymorphism (SNP) maps has benefited the investigations of mendelian (single-gene) diseases, our discussion will be restricted to common complex conditions such as obesity and cardiovascular disease that are determined by multiple genetic and environmental factors. Such diseases constitute the main health burden in developed countries.15,21-23

Given the rapidly changing nature of the field of genetic epidemiology, the large amounts of genomic data being generated at considerable cost, as well as the apparent and unforeseen obstacles facing progress, it is important to consider these initiatives in the context of expediting the discovery of complex human disease genes. We review knowledge about the human genome as related to SNPs and linkage disequilibrium (LD), discuss the potential applications of this knowledge to mapping complex disease genes, and look at the feasibility of whole genome association using SNPs.

Genomic information in mapping complex disease genes

We are at the beginning of our ability to map complex disease genes. Sequencing of the human genome remains the key to this enterprise, but the focus of that project was the consensus human sequence, which by definition cannot contain information about individual differences of medical relevance.21 To make use of the consensus sequence, the SNP Consortium was formed in 1999, with other public and private projects, with the aim of discovering common polymorphism sites in the human genome.22 The increasingly complete catalogue of common genetic variants that is being applied to association studies of complex phenotypes is a direct extension of the consortium’s work. The natural next step to the SNP discovery phase was to genotype identified SNPs in individuals to begin to assess their potential usefulness for disease mapping. This work is ongoing in the International HapMap project. The next stage will involve applications to gene discovery. Some genes associated with complex diseases have been discovered by association-based genetic mapping.26 Genetic association studies are discussed in detail in other papers in this series.26-28

The association of an allele with a phenotype due to correlation (ie, LD) between the allele and a nearby causal variant—so-called indirect association—is the main thrust of whole-genome association studies and large-scale genomic projects like the International HapMap project (discussed below).
SNPs
Because the mutation rate is low (around $10^{-8}$ per site per generation) when set beside the most recent common ancestor of any two people (around $10^6$ generations), most SNPs are thought to arise from a single historical mutational event. Across the human genome, there are far more SNPs than any other types of polymorphism— at least 10 million SNPs with frequency greater than 1%, yielding an average spacing of one every 290 base-pairs. These common SNPs are thought to account for around 90% of human genetic variation.\(^{17,13-15}\)

There are four important advantages of using SNPs rather than other types of genetic polymorphism to investigate the genetic determinants of complex human diseases.\(^{8,14,35}\) First, SNPs are plentiful throughout the genome, being found in exons, introns, promoters, enhancers, and intergenic regions,\(^{8,37}\) and some of these polymorphisms might themselves be functional. Second, groups of adjacent SNPs might exhibit patterns of correlations that could be used to enhance gene mapping\(^{16}\) and which may highlight recombination hot-spots.\(^{39}\) Third, interpopulation differences in SNP frequencies can be used in population-based genetic studies.\(^{36,41}\) Fourth, SNPs are less mutable than other types of polymorphism,\(^{42,43}\) and this greater stability could allow more consistent estimates of gene-phenotype associations.

The common SNPs have been subject to large cataloguing projects funded by both government and industry.\(^{16,17,44}\) These efforts have involved targeted SNP discovery by mutation detection\(^{15}\) or primary resequencing in candidate genes or pathways have independently sought to identify sequence variants by primary resequencing in candidate genes or regions.\(^{30,46}\) Of more than 10 million SNPs so far identified, more than 5 million have been validated.

Many other SNPs present in major ethnic groups are likely to be discovered. SNP databases are constantly being updated (panel 1).\(^{17,11}\) However, the data are not infallible, as some putative polymorphisms turn out to be sequencing errors or rare or population-specific variants often not detected in subsequent studies.\(^{16,17}\) Limitations due to cost and the incomplete status of SNP databases mean that the association analysis of SNPs in complex disease genetics has been mostly limited to polymorphisms within biologically plausible candidate loci. Many investigators interested in specific genes or pathways have independently sought to identify sequence variants by primary resequencing in their own study populations.\(^{31,46}\)

SNPs are finding widespread use in fine mapping of genetic disorders, in the delineation of genetic influences in multifactorial diseases such as breast cancer, cardiovascular disease, type 2 diabetes and asthma, and as genetic markers to predict responses to drugs and adverse drug reactions.\(^{32}\) There are at least six primary areas of potential application for SNP technologies in improving our understanding of complex disease: (1) hypothesis-free gene discovery and mapping; (2) association-based candidate polymorphism testing; (3) pharmacogenetics; (4) diagnostics and risk profiling; (5) prediction of response to non-pharmacological environmental stimuli; and (6) homogeneity testing and epidemiological study design.\(^{3}\) There are thus dual imperatives to develop advanced technologies to detect and genotype SNPs, and for improved statistical approaches and study designs to enable SNP data to be incorporated into epidemiology and clinical medicine.

**Linkage disequilibrium**
Most SNPs lie outside genes and are not likely to alter gene structure or function, so they might not be directly associated with any change in phenotype.\(^4\) We need to know whether the DNA sequence variant under consideration is potentially directly functional (ie, could lead to the observed biology) or is indirectly correlated with another DNA sequence variant that is the actual cause of the phenotype of interest. Since candidate genes are usually difficult to select\(^4\) and since functional data are rarely available for a given SNP, testing for indirect association is the model which most attempts at gene

---

**Panel 1: Selected websites**

**SNP databases**
- Cancer Genome Anatomy project [http://cgap.nci.nih.gov/]
- Génome Québec [http://www.genomequebec.com/index_e.asp]
- The Golden Path [http://genome.ucsc.edu]
- Human Genome Variation Database [http://hgvsbase.cgb.ki.se/]
- The Human Genome Variation Society [http://www.genomic.unimelb.edu.au/mdi/]
- Human Gene Mutation Database [http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html]
- The International HapMap Project [http://www.hapmap.org/]
- LocusLink [http://www.ncbi.nih.gov/LocusLink/]
- NHLBI Programs for Genomic Applications Resources [http://pga.lbl.gov/PGA/PGA_inventory.html]
- SNP Consortium [http://snp.cshl.org/]
- SNP View [http://snp.gnf.org/]
- The Sanger Centre [http://www.sanger.ac.uk/]

**Software**
- An extensive list of genetic analysis software [http://linkage.rockefeller.edu/soft/list.html].
discovery use. LD is discussed in other papers in this series.27,28 Loci in LD are generally close together, but the relation varies (figure 1). When a variant is first introduced into a population by mutation, it will be perfectly correlated with nearby variants, but over successive generations meiotic recombinations will break up the correlations, and LD will decay (figure 1). Indirect association mapping relies on LD in the sense that the functional variant need not be studied at all, so long as one measures a variant that is in LD with it.

Many factors can influence LD, including genetic drift, population growth, admixture, population structure, natural selection, variable recombination and mutation rates, and gene conversion.49,50 The International HapMap project was started to describe disequilibrium patterns in some ethnic groups and it should help clarify the value of SNPs for the indirect association mapping of disease genes25 (see below).

Haplotypes and haplotype estimation
Indirect association mapping by LD relies on gene-phenotype associations at the level of population,51 and requires a dense map of markers.52 It may be enhanced by examining multiple markers simultaneously or using haplotypes, which are linear arrangements of closely linked alleles on the same chromosome inherited as a unit. Haplotype analysis in the context of disease association studies is difficult,53 but haplotypes do contain at least as much information as the genotypes at each component locus, so may prove essential for some disease gene studies.

For M biallelic markers there are $2^M$ possible haplotypes (though often many fewer are evident), and because we usually do not know in advance which haplotypes might be associated with disease, all are tested. Testing SNPs one at a time would require $M$ tests so the greater information in haplotypes is offset by the cost of testing more of them. The growing use of phylogenetic approaches derived from population genetics in human gene discovery investigations holds promise in this area,54 as it helps to form natural groupings of haplotypes.

When LD is high, the redundancy amongst markers means that haplotypes can be used in association studies to efficiently map common alleles that might influence the susceptibility to common diseases, as well as for reconstructing genomic evolution.55 When LD is low, haplotypes will generally be useful in refining SNP-phenotype associations only if they help delineate rare allele frequencies or if there are significant interactions among the SNPs in their effect on the trait. In complex diseases, where multiple variant loci contribute to disease susceptibility, haplotypes are therefore also potentially important since different combinations of particular alleles in the same gene may act as a meta-allele or meta-SNP and have different effects on the protein product and on transcriptional regulation.56

In population-based studies based on unrelated individuals, the parental origin of each allele of a genotype is not known (so-called phase unknown status); haplotypes for double heterozygotes are uncertain and must be estimated.57 Statistical methods and software are available to estimate haplotypes from phase unknown genotype data in large population-based samples of unrelated individuals or in family data,57–62 and new maximum-likelihood methods have been developed to allow the testing of statistical association between haplotypes and binary, ordinal, and quantitative traits.58 However, the use of haplotypes derived from phase-unknown genotype data is not always straightforward, and the value of these techniques for gene mapping is not yet clear.57,58,65

![Figure 1: Theoretical (upper) and observed (lower) patterns of LD decay](image-url)
Series

LD patterns across the genome

Large sets of SNPs and improved genotyping technology and statistical methods for haplotype estimation are necessary for improving gene discovery via indirect association analysis, but there is more information available. The extreme variability in the correlation between physical distance and LD in a given genomic region (figure 1) means that two genetic variants that are physically close will sometimes be completely independent, whereas loci that are very far apart will sometimes be highly correlated. Thus, when LD is low, screening nearly all of the SNPs in a given region could still miss the relevant locus. When LD is high, evidence for association can be found for most of the loci examined, which would reveal little about the precise localisation of the aetiological variant. These two extremes are depicted in figure 2, where a chromosome region in which many markers are associated with the outcome (top left) is contrasted with a region in which only a single marker reveals evidence for association. The different patterns of disease association are due to different LD patterns in the chromosome regions.

Until recently, little was known about LD patterns in the genome except for a few well-characterised genes and gene families. However, studies of large genomic regions or entire chromosomes are now adding to this knowledge base, highlighting the importance of dense marker panels and revealing extensive variability in LD patterns and recombination rates. Further information is needed to enable appropriate study design and more accurate interpretation of association studies. The International HapMap was initiated in recognition of this need (panel 2).

Figure 2: The role of LD in facilitating allelic association

A and B: disease association profile for hypothetical disease in which aetiological locus confers OR of 3.0. Markers in A show extensive background LD, so many are associated with trait. Markers in B show little LD, so only causal locus is associated. Distribution of LD for these two scenarios shown below to illustrate that knowing local patterns can help to delineate expected patterns of association and design efficient novel studies. Data from chromosome 22, in which arbitrary locus was designated disease gene in high and low regions of the chromosome. Decay in odds ratio computed as described.

LD (r^2)

Distance between genetic markers (Kbp)
The focus on SNP genotyping has made it clear that new statistical methodologies and study designs are needed for LD mapping of complex trait genes, and has led to re-examination of mapping methodologies and study designs. The fundamental issue of how to deal with the volume of data produced is only now being addressed; developments in biostatistics have been lagging behind the capacity to generate SNP genotypes. The best way to apply SNPs and LD mapping data to the genetic epidemiology of common diseases remains unclear. A number of statistical methods for selecting haplotype-tagging SNPs are available and more are in the pipeline. The differences between these diverse approaches will need to be understood to make efficient use of genome-wide LD data. Additionally, the applicability of a tagging approach developed in one population to other populations has not yet been fully examined, leading in part to the wide range of differences in the estimates of the potential gain in genotyping efficiency resulting from the use of HTSNPs.

One practical challenge facing haplotype tagging (panel 2) is the definition of the genomic region to be tagged. Tagging was initially described as a means of efficiently genotyping, but it was later wedded to the notion of haplotype blocks, which are regions of very high LD delineated by regions of low LD. As block boundaries are not always consistent within or between populations, or between statistical definitions it is not clear that block-tags defined in one sample will capture the same information in another. Ultimately, the region definition problem may be addressed empirically by examining multiple samples drawn from many populations, or theoretically by statistical methods that do not depend on physical boundaries.

Missing data, an issue for genetic analysis generally, are a particular problem for haplotype analysis. Sequencing or genotyping a given set of SNPs is rarely 100% complete and missing data with each additional SNP included in a haplotypic analysis. Other branches of statistical investigation have learned that ignoring missing data or restricting analysis to individuals with complete data can lead to biased or inefficient analyses, even when data are missing completely at random.

This problem worsens if data are not missing at random, as may be the case with systematic errors in genotyping assays. Methods for dealing with missing data have seldom been applied to genetic epidemiology but more needs to be known about the extent to which missing data are a problem in genetic association analyses of SNPs and haplotypes and about the application of methods for dealing with missing data in such studies.

### Statistical methods

The focus on SNP genotyping has made it clear that new statistical methodologies and study designs are needed for LD mapping of complex trait genes, and has led to re-examination of mapping methodologies and study designs. The fundamental issue of how to deal with the volume of data produced is only now being addressed; developments in biostatistics have been lagging behind the capacity to generate SNP genotypes. The best way to apply SNPs and LD mapping data to the genetic epidemiology of common diseases remains unclear. A number of statistical methods for selecting haplotype-tagging SNPs are available and more are in the pipeline. The differences between these diverse approaches will need to be understood to make efficient use of genome-wide LD data. Additionally, the applicability of a tagging approach developed in one population to other populations has not yet been fully examined, leading in part to the wide range of differences in the estimates of the potential gain in genotyping efficiency resulting from the use of HTSNPs.

One practical challenge facing haplotype tagging (panel 2) is the definition of the genomic region to be tagged. Tagging was initially described as a means of efficiently genotyping, but it was later wedded to the notion of haplotype blocks, which are regions of very high LD delineated by regions of low LD. As block boundaries are not always consistent within or between populations, or between statistical definitions it is not clear that block-tags defined in one sample will capture the same information in another. Ultimately, the region definition problem may be addressed empirically by examining multiple samples drawn from many populations, or theoretically by statistical methods that do not depend on physical boundaries.

Missing data, an issue for genetic analysis generally, are a particular problem for haplotype analysis. Sequencing or genotyping a given set of SNPs is rarely 100% complete and missing data with each additional SNP included in a haplotypic analysis. Other branches of statistical investigation have learned that ignoring missing data or restricting analysis to individuals with complete data can lead to biased or inefficient analyses, even when data are missing completely at random.

This problem worsens if data are not missing at random, as may be the case with systematic errors in genotyping assays. Methods for dealing with missing data have seldom been applied to genetic epidemiology but more needs to be known about the extent to which missing data are a problem in genetic association analyses of SNPs and haplotypes and about the application of methods for dealing with missing data in such studies.

### Power, p values, and multiple testing

In complex disease genetics, both type I and type II error needs to be reduced. Power for studies of allelic association will depend primarily upon sample size, the effect size of the susceptibility locus, the strength of LD with a marker, and the frequencies of susceptibility and marker alleles. Figure 3 illustrates sample sizes needed to detect a true odds ratio of 1.5 with 80% power and
Sample size is cases plus controls, with one control per case; detectable difference of OR 1·5; power 80%. α=type I error probability (α) of either 0·05 or 0·005. Even for the best-case scenario, a common SNP acting in a dominant fashion, more than 800 people are needed at the 0·05 level, which is still in widespread use by researchers and journal editors.

Multiple testing is an issue in many genetic association studies of candidate loci where multiple SNPs in one gene or multiple SNPs in several loci are tested, or both, and an α on the order of 0·005 could be more realistic, even for only a small set of genetic markers. Use of α=0·005 or with an uncommon SNP that acts in a recessive fashion leads to large sample sizes. This problem will be exacerbated in studies with more SNPs, such as whole genome association designs (and even larger samples have revealed that the larger the sample size, the greater the potential bias from stratification.128 Population-based studies of unrelated controls should begin to correct for it,111,114,121–127 coupled with careful population-based studies of unrelated controls should reduce confounding by population stratification.128 Research on the performance of genomic control with large samples has revealed that the larger the sample size, the greater the potential bias from stratification.111,114 We may need to type many hundreds or even thousands of markers to detect and control subtle stratification in large samples.111 Fortunately, genotyping costs are falling.111,114

Understanding how aetiological factors act at a population level will be a critical step for the clinical application of knowledge about the genome.4,129,130 Genetic knowledge will only become clinically useful when it is placed back in an epidemiological and public health context.71 Very large, longitudinal, well-characterised population-based studies drawn from multiple ethnic groups will have a vital role in the implementation of SNP-based gene discovery and in diagnostic tests for complex phenotypes in the outbred, highly admixed populations that increasingly characterise human societies today.71

Population heterogeneity
Population heterogeneity is a serious issue for gene discovery in any population-based study of complex diseases.108–114 Disease prevalence often changes with geography and ethnic origin, and allele frequencies can vary widely throughout the world.115 Additionally, there is likely to be a high degree of variation in LD between populations of different origins,112,116 and between different genomic regions,110,116 leading to differences in genetic-physical map correlations, estimates of LD and haplotypes, tagging SNP selections, and other outcomes. This heterogeneity can complicate or even prevent gene discovery and cloud apparent evidence for replication.

For association studies of many complex diseases, case-control designs have become the approach of choice. The biggest criticism of such studies has been the potential for undetected population stratification: spurious association may arise when allele frequencies vary across subpopulations (eg, people from different ethnic groups119). This is a potential issue for both direct candidate gene approaches and indirect association.120 Such stratification may result from recent admixture or from poorly matched cases and controls. Genetic control, genotyping of random panels of SNPs to assess population structure and begin to correct for it,111,114,121–127 coupled with careful population-based studies of unrelated controls should reduce confounding by population stratification.128 Research on the performance of genomic control with large samples has revealed that the larger the sample size, the greater the potential bias from stratification.111,114 We may need to type many hundreds or even thousands of markers to detect and control subtle stratification in large samples.111 Fortunately, genotyping costs are falling.111,114

Understanding how aetiological factors act at a population level will be a critical step for the clinical application of knowledge about the genome.4,129,130 Genetic knowledge will only become clinically useful when it is placed back in an epidemiological and public health context.71 Very large, longitudinal, well-characterised population-based studies drawn from multiple ethnic groups will have a vital role in the implementation of SNP-based gene discovery and in diagnostic tests for complex phenotypes in the outbred, highly admixed populations that increasingly characterise human societies today.71
Rare alleles

Current attention in population-based association studies is focused almost entirely on genetic markers and aetiological variants that are common (>1% frequency). This is true for SNP detection studies, public databases, the HapMap project, haplotype tagging, and most sample collections are powered to detect effects only arising from common variants. There are several reasons for this emphasis. The most cited one relates to the common-disease, common-variant hypothesis, which holds that genetic influences on diseases of high population prevalence are old, and are thus typically very common. There are arguments and evidence for and against this hypothesis, as well as empirical support and counterexamples.

Another reason for the emphasis on common alleles is purely practical. Common diseases are assumed to be influenced by many genetic and environmental factors, all with a modest effect on the trait. If the genetic influences are rare, the sample sizes required to detect the modest effects become impossibly large (figure 3). Thus, in the absence of so-called low-hanging fruit (genes with major effects on complex phenotypes) it is impractical to search for rare genetic effects using the allelic association design. This practical consideration explains the current focus on gene discovery strategies aimed at common alleles and implies that real effects associated with rare alleles will go undetected.

Allelic heterogeneity accentuates the problem of rare alleles. With the breast and ovarian cancer loci BRCA1 and BRCA2, the phenotype results from a very large number of different mutations in the same gene(s), so that many people have extremely rare or unique mutations. Such heterogeneity would possibly not be detected by population-based association, no matter how large the sample size or the number of common SNPs genotyped (the BRCA1 and BRCA2 loci were identified by family-based linkage). Thus there are genetic aetiologies that are not amenable to discovery by population association analysis. As these are not known a priori, it is important to emphasise that the vast SNP datasets being constructed, the HapMap project, enhanced genotyping capacity, and all the other resources being brought to bear on this problem will not always lead to gene discovery.

Replication

Several recent articles have addressed the features of a good genetic association study. This focus on study design stemmed from the realisation that genetic association studies of complex phenotypes often fail to discover susceptibility loci or fail to replicate studies that did. Despite the widespread use of genetic case-control studies, their inconsistency is a generally recognised limitation. This lack of reproducibility is often ascribed to small samples with inadequate statistical power, biological and phenotypic complexity, population-specific linkage disequilibrium, effect-size bias and population stratification.

Other reasons for the non-replication of true positive association results include inter-investigator and inter-population heterogeneity in study design, analytical method, phenotype definition, genetic structure, environmental exposures, and markers genotyped. It is now routinely argued that large sample sizes (generally, thousands rather than hundreds), rigorous p-value thresholds, and replication in multiple independent datasets are necessary for reliable results. For most complex human diseases, the reality of multiple disease-predisposing genes of modest individual effect, gene-gene interactions, gene-environment interactions, heterogeneity of both genetic and environmental determinants of disease and low statistical power mean that both initial detection and replication will likely remain difficult.

Ironically, the advances in SNP genotyping and LD mapping that offer promise for association studies also highlight some of the difficulties that large SNP studies face. Decreasing costs mean that more SNPs will be typed, and thus more spurious results will be obtained. This places a greater burden on establishing robustness via replication. However, different definitions of replication are emerging. Descriptions of so-called confirmatory replication are often attached to findings that appear non-confirmatory. For example, different genetic markers are significantly associated in the follow-up study differ from those in the original report; or the same genetic markers are reported in both studies, but with opposite alleles (ie, the disease allele is protective in another sample drawn from the same population) look biologically less plausible than other explanations (eg, a risk allele in one sample appearing as protective in another sample drawn from the same population) look biologically less plausible than other replication scenarios. Although there is no disputing the importance of heterogeneity within and between samples and genes, there is a risk that heterogeneity could be abused to rationalise negative follow-up studies in positive terms.

In general, studies showing similar results in terms of phenotypes tested and specific SNP associations found offer strong evidence for association. However, those lacking such clear overlap, even with positive association evidence, may require validation using other strategies.
or datasets. Future studies of large numbers of SNPs will need to approach these issues carefully lest replication lose its status as a gold standard for genetic association.

**Whole genome association**

High density SNP maps and the identification of genes by the Human Genome Project have made whole genome association analyses technically feasible for many conditions. However, despite costs heading down to US$0.01 per genotype (a target once regarded as highly ambitious), testing all of the 10 million common SNPs would cost at least US$100 000 per individual or US$200 million for a single study of 1000 cases and controls. Exhaustive genotyping for association is therefore currently impractical.

**What is a whole genome association study?**

Forms of whole genome association are now being explored. Whole genome implies complete coverage but not all such analyses are the same. For example, marker sets of 100 000 or more SNPs are now commercially available as whole genome panels. In constructing such panels, one could select SNPs in a variety of ways—eg, with a focus on genes only, via haplotype tagging or at random throughout the genome. None of these covers all variation in the genome, so by a strict definition, none offers a whole genome study.

Indeed, genotyping 100 000 SNPs in many populations would probably cover less than 50% of common genetic variants. Whole genome association studies will require qualifiers describing their aims, assumptions and presumed coverage. The concern is not so much that what they do find will be false but how many and of what composition are the genetic variants that they missed.

Complete resequencing of the entire genomes of case and control individuals would be ideal, but this technology is not yet available or affordable. The high-density panels being genotyped in the International HapMap project (panel 2) and in industry offer the most immediate form of whole genome coverage. Although rare variants are under-represented, 85–90% of the genetic variants that are common in the samples evaluated may soon be available for disease-genome research.

**Reducing the genotyping burden**

There are at least two strategies for reducing the number of SNPs that need to be genotyped, one based on indirect association and haplotype-tagging SNPs across the genome (map-based) and the other based on direct association and the genotyping of all potentially functional SNPs across the genome (sequence-based).

The map-based approach makes no assumptions about the genes involved or the type of the mutation, though it does assume that disease alleles or haplotypes are sufficiently frequent to have been captured by the original tagging study. Estimates for the number of tag SNPs needed to represent most common variants across the entire human genome range from 200 000 to more than a million. A single genome-wide study would still cost several million US$ for 1000 cases and controls. Moreover, the SNPs genotyped in such a study would be highly selected in order to reflect the underlying LD patterns in the relevant population. In this regard, the feasibility of whole genome association scans in the map-based model depend critically upon knowledge of genomic LD patterns in multiple populations. Random sets of uniformly spaced SNPs, though cheaper, easier to genotype and increasingly available commercially, do not yield the same efficiency or robustness. Further decreases in genotyping cost or savings in the number of markers to genotype are needed for well-powered association studies across the genome.

The sequence-based approach makes savings by assuming that specific variants are more likely to influence complex traits than others. Prioritised lists of such variants decrease the number of SNPs to 50 000–100 000 and study costs less than US$1–2 million for 1000 cases and 1000 controls. However, despite the availability of over 10 million SNPs in public databases, further work may be needed to identify all SNPs at the top of the priority list (ie, non-synonomous, non-conservative coding changes). In addition, many coding changes are rarer in their allele frequencies than non-coding changes, thus creating sample size challenges unless the genes have large effects.

One approach that can reduce genotyping requirements under both strategies is the use of generic or universal controls—or a large set of representative controls from which subsets are matched to individual disease samples. Genotyping a genome-wide set of markers on such a sample would allow re-use of the genotypes across the disease samples. Genomic control could facilitate matching and reduce potential confounding. One potential role for large cohort initiatives such as UK BioBank will be to provide such universal controls. Another labour-saving strategy is staged genotyping, so that not all markers are genotyped on all individuals. By genotyping all markers on a subset of the sample and liberally selecting the marker set to be genotyped on the remainder of the sample, it should be possible to retain most of the statistical power while reducing the genotyping load. Savings of up to 75% of potential genotyping reactions with minimal loss of power have been demonstrated with genetic analysis of type 1 diabetes samples.

The map-based and sequence-based approaches both hold promise for genome-wide studies. It is not clear which will prove more fruitful, and it is certain that no single approach will work for all situations.

**The future**

Explosive growth in technical capacity and genomic knowledge has been tempered by initial failures to find genes for complex phenotypes using any strategy and our
A comprehensive understanding of complex disease pathophysiology has been hampered by the overinterpretation of marginal results. Our current state of knowledge in genetics, many of which can be blamed on poor design and study execution, is justified. Another cause for hope is the assimilation of large numbers of successful gene localizations for complex diseases. The genomics revolution has been accompanied by an unfortunate tendency to overstate potential. This has led to unrealistic expectations among clinicians and to cynicism and pessimism within the genetics community. For genetics researchers, one of the most important tasks now is to not add to the hyperbole but to establish and communicate realistic expectations.

Where does LD-based association mapping stand today? For most complex human diseases, the reality of multiple disease-predisposing genes of modest individual effect, gene-gene interactions, gene-environment interactions, inter-population heterogeneity of both genetic and environmental determinants of disease, and low statistical power mean that both initial detection and replication are likely to remain difficult.12,13,15 However, our understanding of the complexity of the task is improving and new tools and a growing knowledge base (eg, rapid progress in SNP detection, complete catalogues of SNPs, and the attention being paid to methodological problems in LD mapping and haplotypic approaches) do offer prospects for success in gene discovery. These and other developments, taken together with a small but growing number of successful gene localisations for complex phenotypes, suggest that cautious optimism about discovery of genes underlying common human diseases is justified. Another cause for hope is the assimilation of genetic epidemiology into mainstream epidemiology and public health in many academic institutions. The involvement of epidemiologists should improve some of the difficulties that have plagued complex disease genetics, many of which can be blamed on poor design and overinterpretation of marginal results. Our understanding of complex disease pathophysiology has already begun to enter into the realm of clinical genetics,16 and we have every reason to anticipate that the impact of genomics upon clinical practice and upon our understanding of biology and epidemiology will continue to accelerate.

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

Acknowledgments
This work was supported in part by the Wind-over-Water Foundation (LJP) and by a Wellcome Trust Principal Research Fellowship and NIH grant EY-12562 (LRC).

References
16. Varmus H. Genomic empowerment: the importance of public health experiences, genomic region, and gene under investigation. There is no one paradigm for gene discovery and no single ideal study design or analytical approach. Despite ex cathedra statements on optimum study design and analytical, it is clear that flexible, mixed approaches and hypothesis-free designs are desirable. The genomics revolution has been accompanied by an unfortunate tendency to overstate potential. This has led to unrealistic expectations among clinicians and to cynicism and pessimism within the genetics community. For genetics researchers, one of the most important tasks now is to not add to the hyperbole but to establish and communicate realistic expectations.

Where does LD-based association mapping stand today? For most complex human diseases, the reality of multiple disease-predisposing genes of modest individual effect, gene-gene interactions, gene-environment interactions, inter-population heterogeneity of both genetic and environmental determinants of disease, and low statistical power mean that both initial detection and replication are likely to remain difficult.12,13,15 However, our understanding of the complexity of the task is improving and new tools and a growing knowledge base (eg, rapid progress in SNP detection, complete catalogues of SNPs, and the attention being paid to methodological problems in LD mapping and haplotypic approaches) do offer prospects for success in gene discovery. These and other developments, taken together with a small but growing number of successful gene localisations for complex phenotypes, suggest that cautious optimism about discovery of genes underlying common human diseases is justified. Another cause for hope is the assimilation of genetic epidemiology into mainstream epidemiology and public health in many academic institutions. The involvement of epidemiologists should improve some of the difficulties that have plagued complex disease genetics, many of which can be blamed on poor design and overinterpretation of marginal results. Our understanding of complex disease pathophysiology has already begun to enter into the realm of clinical genetics,16 and we have every reason to anticipate that the impact of genomics upon clinical practice and upon our understanding of biology and epidemiology will continue to accelerate.

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

Acknowledgments
This work was supported in part by the Wind-over-Water Foundation (LJP) and by a Wellcome Trust Principal Research Fellowship and NIH grant EY-12562 (LRC).


End of life: the humanist view

Julian Baggini, Madeleine Pym

A central tenet of humanist thought is that this world and this life are the only ones we have, and that, in the absence of an after-life and a soul, death brings a natural end to our existence. Broadly defined, a humanist is a morally concerned person who is not religious. However, individual humanist beliefs vary immensely—from atheists, who regard God as a human creation, to agnostics who might merely assert that although they can neither prove nor disprove God’s existence, it is of no practical concern to them. But all humanists will tend to share a scepticism toward, and a rejection of, traditional religion and religious ritual, and a positive commitment to living a morally responsible life.

In view of this practical emphasis on the world, the community, and the individual, rather than a transcendental emphasis on God and the after-life of the soul, how can those in the medical profession best meet the needs of humanists as they approach death? In answering this question, we will begin with a brief introduction to humanism and its main tenets, going on to consider what differentiates the humanist approach to death from the religious one. We will explain the main criteria for what might constitute a good death for humanists and how medical staff can help them achieve this end.

The key issues are: the fundamental requirement to accord the needs and beliefs of humanists the same respect given to religious believers; the need to recognise that humanists vary in their attitude towards their own death and might or might not want to be informed of the full facts of their case, despite a genuine respect for truth and honesty in all matters; and the need for autonomy and control over the means of death and the treatment of their body after death. We also will discuss some of the difficulties incurred when the dying person’s family neither respects nor shares his or her beliefs. We will include recommendations about how the needs of humanists can best be accommodated in the hospital.

Humanism

Some might argue that without a belief in God, and a set of God-given moral rules, there is no requirement for a person to embrace any moral code whatsoever. Although this view describes some atheists, it is not a definition of humanism, which consists not merely in the denial of the existence of God, but in a positive creed that asserts, first, that the moral life is desirable in itself, and second, that because there is no supernatural source of moral values, it is up to people to identify these principles and abide by them.

Historically, humanism is a continuation of the western European tradition of non-religious thought that can be traced back some 2500 years to the philosophy of the ancient Greeks. One of humanism’s moral roots can be found in the enlightenment tradition of liberal values that is also enshrined in the United Nation’s Universal declaration of human rights; the belief that all people are equal and should be free to hold and express their beliefs as long as they do not harm others.

Because humanists (panel 1) believe that morality is independent of religious faith, and that humans cannot rely on anyone or anything other than themselves to solve human problems, humanists have often been active social reformers. The early ethical societies of the Victorian era did much social and educational work in city slums, and 20th-century humanists were instrumental in opening up adoption services to non-religious people, at a time when these services were administered by the church, and unless you were a practising Christian you could not adopt a child. Today, the British Humanist Association meets the need to mark the major life transitions of birth, marriage, and death, by training and accrediting humanist officiants to conduct non-religious ceremonies. It is a typical feature of humanists that they prefer to take ownership of these occasions by scripting the ceremony themselves, selecting music, poetry, or prose that has a special significance for them and that focuses on the people concerned rather than on a transcendental world.

For professionals in medical and palliative care, humanists’ emphasis on individualism can make meeting the needs of humanists difficult. The rituals and beliefs of many organised religions are quite well prescribed: however, humanism, with its tradition in rationalism and free thought, leaves a great deal to the individual. Although for hospital staff to discern the needs of humanists is difficult, it can be just as difficult for humanists to decide their own needs in the absence of a set of prescriptive rules governing the way they should approach death, and the disposal of the body, or of rituals that should be performed.

Respecting the positively non-religious

The first step towards according humanists the same respect accorded to religious believers is a simple one—keep hospital and hospice environments free of religious symbols and texts, which can be made available on

Panel 1: Some 20th-century humanists

request to those who want them. This basic consideration not only avoids causing offence to humanists, but also to those of other faith groups as well.

We should remember that not all non-religious people are humanists—many people reject religion without embracing the positive moral position of humanism. They might, for example, ascribe to new age spiritualism, eastern religion, a belief in reincarnation, or just a vague belief that there is something more. However, these people are not humanists, because a prerequisite of being a humanist is that the supernatural has been rejected.

Identification of humanists can be difficult because, unlike with organised religions, it is possible to be a humanist without knowing that humanism is the best description of one’s beliefs. A humanist is any atheist or agnostic who believes in the possibility of a meaningful and moral life.

Humanists can be identified at the point of entry into hospital care. Pre-admission forms include a space for people to self-identify according to religious belief; however, these forms generally do not include an option for non-religious beliefs. Some members of the British Humanist Association have reported that when filling out these forms they have informed hospital staff that they are atheists or humanists, and the space has been filled with the word “none” or has been struck through, or they have been marked as “C of E”, an abbreviation for Church of England, a term that is often used as a default category for those who profess no particular religious affiliation. An adapted form would allow humanists to identify themselves on point of entry while also alerting staff.

Humanists are too often characterised as hardened atheists and rationalists, and it might be imagined that all humanists are hard-nosed stoics, determined to face-down death with steely determination. But this assumption should not be made. Certainly, some humanists do approach death in just this way, subscribing to the view that there is nothing to fear about being dead since once we are dead we will simply not be there—so nothing good or bad can happen to us. But as Woody Allen said, “It’s not that I’m afraid to die. I just don’t want to be there when it happens”. Dying can be as fearful for humanists as for anyone else, and although they accept the inevitability of death, that does not mean they are going to be happy when it happens. For example, if a priest gives last rites over a humanist they might, for example, ascribe to new age spiritualism, eastern religion, a belief in reincarnation, or just a vague belief that there is something more.

Similarly, although humanists reject traditional, religious, and confessional forms of counselling, they might want to talk about their coming death with a neutral third party. As death approaches, people of all beliefs often feel the need to unburden themselves of past secrets or achieve some form of closure in unresolved difficulties. Providing a secular counselling service can meet this need, and in some parts of the UK, humanist officiants, in addition to conducting weddings, funerals, and baby-naming ceremonies, have extended their work into areas such as palliative counselling and secular chaplaincy. If counsellors are used, their approach should be entirely secular.

Understandably, onlookers want to offer some sort of consolation in the face of death, and religion has for many years offered just that; however, doing nothing is sometimes preferable. Faced with a person who has perhaps suffered a terrible and painful lifelong illness, or who is dying at a young age and before there are any real achievements to look back on, what consolation can possibly be offered? Humanists do not expect easy answers here, and would be unlikely to provide any, other than the fact that their misery is soon to end. Humanists would generally prefer hospital staff to say and do nothing, rather than attempt to fill what appears to be a void with religion.

For those charged with the care of the dying, and for family and friends, the desire to fill this void can be hard to resist. But if this means providing something that is in conflict with the patient’s belief system, this desire must be resisted. Sometimes hospital staff can do nothing except provide the best medical care possible and an environment in which patients can come to terms with their death themselves.

Can religion really do any harm?
If humanists do not believe in the truth of religion, can they be harmed by religious rituals they are unaware of? For example, if a priest gives last rites over a humanist who has slipped into terminal unconsciousness, what harm can possibly be done? The claim here is that humanists cannot be harmed by something they don’t believe in the power of, and have no awareness of. But it does not follow that because humanists believe death is the end, they also have no interest in, or claim on, what happens after their death. To draw a comparison, I do you an injustice if I slander you behind your back, even if you never find out about it, and I can equally do you an injustice if I slander you after your death. And whether or not a humanist can be wronged after death, a person’s wishes still need to be respected after death. Humanists are committed to specific values: a resistance to the dominance of traditional religious forms of ceremony and to the assumption that meaning, purpose, and ethics can be supplied only by religion. Thus for a humanist to consent to any form of religious ceremony is to assert exactly that which humans oppose, and would constitute a gross disregard for their views. In the wider context of society as a whole, part of the humanist project is to gain recognition that people can live happy, full, and moral lives without recourse to religious dogma and rituals. To do anything to a dead or dying humanist that would undermine this project or fail to show the project proper respect would therefore be inappropriate.
These are difficult issues for medical staff to address, and staff can and do come under pressure from family—from a parent, for example, who has a strong desire (even a psychological need) for last rites to be given, even though the patient explicitly rejected these rituals. Faced with an anguished parent and an unconscious patient, to grant the parent’s wish might seem harmless. However, to do so would be to behave with gross inconsistency. After all, would it not be unimaginable to allow a non-Christian death ritual to be done with a patient who was a committed Christian but whose parents were of another faith, or vice-versa? So it is for the humanist. Similarly, what harm might there be in allowing a priest to drop by on his rounds for a chat—after all, he can always be sent away? But being visited by a priest will not be part of the script humanists are trying to write for themselves as they face their last days. Such a visit can cause both harm and offence to humanists, who might or might not be able to express how offended they would be to staff if, at the point of death, they are faced with someone offering the consolations of religion and a life after death. Moreover, it would be a signal failure to respect the humanist’s right to confront death without these props.

Again, this is not merely a thought experiment. Hospital staff have been known to allow priests to approach dying humanists, just to check they do not want to change their mind and embrace the faith before it is too late. For humanists who actively reject religion such actions cause much offence. Indeed, one humanist officiant categorically stated, when asked how he would want to be treated on his death-bed, “Make sure no religious twist stops by to talk”. Clearly, allowing a priest to visit this humanist’s bedside would not only cause offence but also anger. However, each humanist is different and some might be happy to talk to a chaplain in some circumstances.

Personal autonomy has an important role for humanists. Humanists typically value the way in which they can, and must, become the authors of their own lives. They will strongly resist being caught up in standard procedures or those that assume a theistic view of the world. The need to retain control or authorship of their lives right until the end is typical, and is shown by the large number of members of the British Humanist Association who are also members of the Voluntary Euthanasia Society. Humanists want to take responsibility for their deaths just as much as their lives, and their support for euthanasia shows that they are more likely than others to want to have the final say in what treatment they are given and the chance to refuse treatment that merely postpones death a little longer.

Humanism encourages open discussion of death and is opposed to pretending it does not happen. But individual humanists approaching death have different needs, and many might prefer not to know everything. Fear and sensitivity are not alien to the humanist psyche.

Conclusion
There can be a tendency to assume, perhaps because of the absence of dogma and rules, that humanists have fewer needs than religious believers. If people say they are Sikh, for example, their needs are well known and documented and care is taken to meet these needs. But if the person is a humanist, the assumption might be that there are no specific needs to be met, other than the minimum and negative need not to offer religious support. We have attempted to demonstrate that humanists’ needs are not fewer than, just different from, those of religious believers. We argue that ignoring these needs is a form of discrimination because it fails to take the belief system of the humanist as seriously as that of the religious believer.

The needs of the humanist must, necessarily, be varied, because every humanist is an individual forging his or her own meaning in the world. We cannot therefore offer specific prescriptions for how all humanists should be treated as they approach death. But what remains true is that we should respect the humanist belief—that this life is all there is and that death is the end, and that at the same time life is of value. To do this we have to keep religion out of the hospital environment and allow patients to approach death on their own terms. For further reading see panel 2.

Panel 2: Recommended Reading
Alcock’s canal syndrome revealing endometriosis

Hélène Nehme-Schuster, Cherif Youssef, Catherine Roy, Jean-Phillipe Brettes, Thierry Martin, Jean-Louis Pasquali, Anne-Sophie Korganow

In October, 2004, a 32-year-old woman presented to us with symptoms associated with her menstrual cycle. She had a history of dyspareunia, dysmenorrhoea, and chronic pelvic pain. Over the previous year, she complained of recurrent knife-like truncated sciatalgia occurring close to her menses. Over the previous 3 months she also had dysuria, right-sided perineal neuralgia, and vulvodynia, occurring at the onset of menstrual bleeding.

Physical examination on the 8th day of her menstrual cycle showed no neurological abnormalities, but bladder distension and a painful mass inside the right rectovaginal septum were present. Full blood count, C-reactive protein, liver function and kidney function tests were normal. Pelvic ultrasonography confirmed urinary retention. Pelvic MRI showed an endometriosis nodule (2·5 cm diameter) in the right upper-part of the rectovaginal septum. Inflammation adjacent to the nodule had entrapped the right sciatic nerve emergence and the right pudendal nerve origin, above the entrance to Alcock’s canal (figure). Perineal electrophysiological investigations, after intrarectal stimulation of the right pudendal nerve, showed an increase in distal motor latency, consistent with chronic compression. We prescribed our patient a non-steroidal anti-inflammatory drug for 1 month and initiated gonadotropin-releasing hormone agonist treatment. When seen in February, 2005, all symptoms had resolved, and in May, 2005, pelvic MRI showed a substantial regression of the endometriosis (figure). Her treatment still includes gonadotropin-releasing hormone agonist and oestrogen add-back.

Alcock’s canal syndrome results from prolonged compression of the pudendal nerve in its osteo-musculo-aponeurotic tunnel. Alcock’s canal is bound laterally by the ischium and medially by the obturator internus aponeurosis (ischiorectal fossa). The pudendal nerve comprises sacral roots S2, S3, and S4, and gives rise to the perineal nerve, which innervates anal and urethral external sphincters, and perineal and vulvar skin in women. Pudendal nerve entrapment causes pain in the lower central pelvic areas (anus, perineum, scrotum, penis, or vulva). Pain can affect one, several, or all areas, and is usually worse in the sitting position. Frequently, there is also urinary or anal incontinence and sexual dysfunction. It was first described in 1987 in male cyclists with transient genitopsinchteral dysfunction and perineal and genital paraesthesia/hypoaesthesia due to prolonged compression of the pudendal nerve in Alcock’s canal. Electrophysiologically exploration of the perineum is necessary to confirm the diagnosis. There is no data available about the general prevalence of Alcock’s canal syndrome, but it seems to be a rare event. Causes include: compression and microtrauma; perineal stretching (during delivery); spinal cord lesions; sacral meningoradiculitis; pudendal neuritis; direct trauma; and post-surgical lesions. Alcock’s canal syndrome is often misdiagnosed. In our patient, pudendal nerve entrapment was confirmed electrophysiologically and morphologically. Sympathetic stimulation of the bladder neck by afferent fibres of the pudendal nerve could explain the urinary retention. MRI of the pelvic endometriosis showed an abnormal nodular structure with spiculated borders associated with low signal intensity on T2-weighted images. Low intensity strandings, due to fibrous tissue, which obscure organ interfaces are typical of adhesions, and are a common in endometriosis; this supports a diagnosis of deep infiltrating pelvic endometriosis. The regression of clinical and radiological features after hormonal therapy also confirms the diagnosis. We suggest that clinicians check for endometriosis in nulliparous women suffering from Alcock’s canal syndrome associated with menstrual cycles.

References