1. The unspoken issue that haunts the UK general election • EDITORIAL
   Page 1515
   The Lancet

2. A US-led Peace Corps for Health • EDITORIAL
   Page 1516
   The Lancet

3. America's new “food pyramid” • EDITORIAL
   Page 1516
   The Lancet

4. Hormones and endometrial cancer—new data from the Million Women Study • DISCUSSION
   Pages 1517-1518
   Louise A Brinton, James V Lacey, Jr and Edward L Trimble

5. Still looking for answers in COPD • DISCUSSION
   Pages 1518-1520
   James F Donohue

6. Extending the benefits of deworming for development • DISCUSSION
   Pages 1520-1521
   Lorenzo Savioli, Dirk Engels and Hiroyoshi Endo

7. Non-HLA humoral immunity and chronic kidney-graft loss • DISCUSSION
   Pages 1522-1523
   Simone A Joosten and Cees van Kooten

8. Shanghai plans responses to AIDS • DISCUSSION
   Pages 1524-1525
   Scott Burris, Joanne Csete, Xia Goumei, Zhou Dan and Bebe Loff
<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Section</th>
<th>Pages</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Dissolving the dilemma over forced treatment</td>
<td>DISCUSSION</td>
<td>1525-1526</td>
<td>Vardit Ravitsky and David Wendler</td>
</tr>
<tr>
<td>10</td>
<td>World Diabetes Day: footing the bill</td>
<td>DISCUSSION</td>
<td>1527</td>
<td>William Jeffcoate and Karel Bakker</td>
</tr>
<tr>
<td>11</td>
<td>thelancet.com—new horizons</td>
<td>DISCUSSION</td>
<td>1528</td>
<td>Richard Lane and Richard Horton</td>
</tr>
<tr>
<td>12</td>
<td>Germany's reforms have yet to deliver insurance savings</td>
<td>NEWS</td>
<td>1529-1530</td>
<td>Ben Aris</td>
</tr>
<tr>
<td>13</td>
<td>Minority health care remains a problem for Canada's leaders</td>
<td>NEWS</td>
<td>1531-1532</td>
<td>Paul Webster</td>
</tr>
<tr>
<td>14</td>
<td>A dark epidemic: responding to the Black Death</td>
<td>BOOK REVIEW</td>
<td>1533-1534</td>
<td>Bill Bynum</td>
</tr>
<tr>
<td>15</td>
<td>End-of-life care</td>
<td>BOOK REVIEW</td>
<td>1534</td>
<td>Agnes van der Heide</td>
</tr>
<tr>
<td>16</td>
<td>Jörg-Dietrich Hoppe: forthright President of the German Medical Association (Bundesärztekammer)</td>
<td>DISCUSSION</td>
<td>1535</td>
<td>Ned Stafford</td>
</tr>
<tr>
<td>17</td>
<td>Katherine Austin Lathrop</td>
<td>PERSONAL REPORT</td>
<td>1536</td>
<td>Anne Harding</td>
</tr>
<tr>
<td>18</td>
<td>Increased risk of cardiovascular events with coxibs and NSAIDs</td>
<td>CORRESPONDENCE</td>
<td>1537</td>
<td>David GI Scott and Richard A Watts</td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>Authors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>Increased risk of cardiovascular events with coxibs and NSAIDs • CORRESPONDENCE</td>
<td>Rebecca Dachman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td>Increased risk of cardiovascular events with coxibs and NSAIDs • CORRESPONDENCE</td>
<td>Paul Emery, Andrew Moore and Chris Hawkey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td>Increased risk of cardiovascular events with coxibs and NSAIDs • CORRESPONDENCE</td>
<td>LG Thomas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>Screening for chlamydia • CORRESPONDENCE</td>
<td>Nicola Low, Roger M Harbord, Matthias Egger, Jonathan AC Sterne and Björn Herrmann</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>Screening for chlamydia • CORRESPONDENCE</td>
<td>John Macleod, Chris Salisbury and Nicola Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td>HCV-indeterminate blood donors or occult HCV infection? • CORRESPONDENCE</td>
<td>Juan A Quiroga and Vicente Carreño</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td>HCV-indeterminate blood donors or occult HCV infection? • CORRESPONDENCE</td>
<td>Nasser Semmo, Gillian Harcourt, Paul Klenerman, Craig Taylor and Neil Smith</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td>Appropriate technology in tuberculosis diagnostics • CORRESPONDENCE</td>
<td>David Moore, Carlton Evans, Robert Gilman, Daniel Vargas and Rod Escombe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td>Strangulation by intravenous tubes • CORRESPONDENCE</td>
<td>Philippe Lunetta and Maria Laari</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
28. Endometrial cancer and hormone-replacement therapy in the Million Women Study • ARTICLE
Pages 1543-1551
Million Women Study Collaborators

29. Effects of N-acetylcysteine on outcomes in chronic obstructive pulmonary disease (Bronchitis Randomized on NAC Cost-Utility Study, BRONCUS): a randomised placebo-controlled trial • ARTICLE
Pages 1552-1560
Marc Decramer, Maureen Rutten-van Mølken, PN Richard Dekhuijzen, Thierry Troosters, Cees van Herwaarden, Riccardo Pellegrino, CP Onno van Schayck, Dario Olivieri, Mario Del Donno, Wilfried De Backer et al.

30. Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis • ARTICLE
Pages 1561-1569
Charles H King, Katherine Dickman and Daniel J Tisch

31. Non-HLA transplantation immunity revealed by lymphocytotoxic antibodies • ARTICLE
Pages 1570-1576
Gerhard Opelz and for the Collaborative Transplant Study

32. Abdominal aortic aneurysm • REVIEW ARTICLE
Pages 1577-1589
N Sakalihasan, R Limet and OD Defawe

33. Aortic aneurysm: the physician as patient • DISCUSSION
Page 1590
Juan Martinez L de Letona

34. Multiplicity in randomised trials I: endpoints and treatments • REVIEW ARTICLE
Pages 1591-1595
Kenneth F Schulz and David A Grimes

35. A blood donor with bacteremia • SHORT COMMUNICATION
Page 1596
Marcia D Haimowitz, Louis A Hernandez and Ross M Herron, Jr
The unspoken issue that haunts the UK general election

Superbugs. Waiting times. Patient choice. More doctors. New investment. Efficiency savings. Foundation hospitals. National Health Service (NHS) reorganisation. Free personal care. These are some of the issues that have been identified by the main political parties as battle-grounds for the UK election on May 5. The Economist, more likely, perhaps, to identify financial and business health as key issues, opted instead for the NHS as the “most important” subject in the campaign. But, sadly, Labour, Conservative, and Liberal Democrat politicians have all failed to address the single most important factor hindering the improvement of health services—the catastrophic collapse in morale among doctors.

Doctors at all levels within the NHS are utterly demoralised. They believe that targets have compromised patient care, undermining clinical decision making in favour of mostly irrelevant managerial gratification. They see the focus on shift work as fatally compromising ideals of professionalism, which are based on getting the job done properly irrespective of the time it takes to do so. They see a lack of continuity of care for large numbers of patients, increasing the likelihood of needless error. They are cynical about a new cadre of managers who have little clinical understanding but create a massively over-managed health service. And they feel let down by their own leaders, who have consistently failed to articulate a positive and assertive vision about the contribution modern medicine makes to society.

The strategic mistake made by politicians is to put the patient before everything else in the quest to achieve effective health care. To quote Labour: “One principle underpins our reforms—putting patients centre stage.” Liberal Democrats pose a question: “Putting patients first. It is not asking much, surely?” And the Conservatives? While they pay lip service to professionalism, Conservative politicians offer few concrete proposals to restore medical morale. The mantra of patients having “the right” to select the hospital or doctor of their choice is repeated. The common error here is to make health an issue of power. Politicians believe that health-seeking behaviours can be equated with almost any other form of consumerism. They cannot. A successful market is the game consumers choose to play based on rules of competition and equal and maximum access to information. Yet none of these assumptions can ever be true in medicine. Health is not a game. It is a public good. Health services should not be about competition. They should be about equity. And access to information is appallingly unequal. Patients are not consumers. They are patients, pure and simple.

Too often, today, the debate about the NHS is based on a set of entirely false notions about what matters to patients. Nigel Crisp, the Chief Executive of the NHS, reminds the public that the NHS will prosper thanks to increased funding and more staff. He blithely assumes the continued commitment of health workers, without considering their attitudes to decades of political aggression aimed at the very values he is so proud to use in his own specious arguments. John Reid, the present Health Secretary, has spoken of a “personalised health service”, one in which the patient takes control of decisions about the prescription of medicines and the selection of surgical procedures. Yet this attitude is manifest nonsense. If doctors have any role at all—and perhaps they do not in Dr Reid’s world view—it is to establish a partnership with the patient. A partnership based not on power or untrammelled patient choice, but rather one based on mutual respect. Respect by the doctor for a patient’s anxiety and care preferences. And respect by the patient for a doctor’s skill and professional expertise.

Where are doctors’ leaders in this debate? Mostly silent or out of touch. To be fair, the Royal Colleges, which could and should play a much fuller part in politics, are heavily constrained by their charitable status. They cannot appear to be partisan in their pronouncements. The British Medical Association, meanwhile, has only one unifying political message across general practitioners, consultants, staff and associate specialists, and juniors—more pay. It is no wonder that the public and mass media sometimes, although wrongly, see doctors as self-serving.

What UK medicine needs is a new and stronger political voice, one that is more concerned with augmenting professional standards than with protecting professional status. That voice does not exist at present. Doctors want to strengthen their professional morale because they know that a more robust and motivated profession will mean better outcomes for patients. Currently, no politician recognises this truth. May 5 will therefore not be about democracy at all. The public will be voting in a vacuum of fact. And our collective health will surely suffer the consequences. ■ The Lancet
A US-led Peace Corps for Health

Oct 14, 1960, was a long day for US presidential candidate John F Kennedy. That evening he had sparred with Richard Nixon—during a debate in which Nixon had called Democrats the “war party”—and afterwards he flew to a speaking engagement at the University of Michigan. Although it was 0200 h in the morning by the time he arrived, more than 10 000 students were waiting for him. Kennedy asked these students whether they would be willing to spend a few years working in Asia, Africa, and Latin America. Thus the idea for the Peace Corps was born.

Nearly 45 years later, but at a rather more decent hour, a “Peace Corps for Health” was proposed. On April 19, the US Institute of Medicine published Healers abroad: Americans responding to the human resource crisis in HIV/AIDS, which outlines a proposal to send US health-care professionals to 15 countries hard-hit by HIV/AIDS. These 15 countries, in Africa, the Caribbean, and southeast Asia, are the focus of the President’s Emergency Plan for AIDS Relief (PEPFAR).

The core of the proposed Global Health Service (GHS) is a small group of full-time US health professionals (150 in the first year) who would work alongside local providers, aiming to bolster those countries’ abilities to fight AIDS, tuberculosis, and malaria. GHS participants would provide clinical services, training, and help in technical and management areas identified as priorities by each country’s health ministry and local PEPFAR teams. In addition to the senior-leadership core, a fellowship programme would provide opportunities for professionals in their early or mid-careers. As an incentive, participants could receive education-loan repayments of up to US$25 000 for every year of service they complete.

The proposal raises important questions. How will the GHS integrate with existing programmes? How will it encourage retention of local providers? Will it have the capacity to strengthen health systems? Despite these complexities, the GHS is, like its late-night off-the-cuff predecessor, a bold and compelling idea that deserves urgent implementation.

America’s new “food pyramid”

On April 19, US Agriculture Secretary Mike Johanns unveiled the new updated “food pyramid”—a contentious graphic symbol and interactive website designed to help Americans choose foods that will help them “live longer, better and healthier lives”. The pyramid draws on the government’s hefty volume of Dietary Guidelines for Americans for content, and was dreamt up as a way of communicating these recommendations to consumers in an easily digestible format.

The free advice enshrined in the pyramid seems well-meaning enough. But in its previous incarnation, first published in 1992, the pyramid sparked a decade-long row about exactly what, and how much, Americans should eat. Many of the issues with the old graphic were a problem of oversimplification. The coloured triangle was simply not detailed enough to denote the subtle differences in nutritional value between, say, a grain of white long-grain rice and its brown or wild counterparts. The new guidelines address many of these criticisms. “Good” fats, such as olive oil, are now separated from “bad” saturated fats and trans-fatty acids. Whereas potatoes were previously lumped together with all other vegetables in a single section, there are now five different groups representing these foods. Exercise has been added to diet as a key element of healthy living. And the pyramid’s architects have even tackled the portion problem; the new guidelines dispense with the ambiguous “serving” used in the 1992 version, opting instead for easily quantifiable cups and ounces.

But critics of the new campaign point out that the guidelines still do not take a hard enough line on such things as the amount of refined starches or red meat in the American diet, and that they inappropriately promote milk consumption. The problem, of course, is that the US Department of Agriculture has an inherent conflict of interest: its primary mission is to ensure the success of the American agriculture industry, so its nutritional advice is an inevitable compromise of vested interests. This conflict means the pyramid cannot communicate what American consumers really need to hear—and makes a powerful argument for leaving health advice to health professionals.
In this week’s *Lancet*, investigators from the Million Women Study (MWS) report important results on the extensively studied but incompletely understood relation between menopausal hormones and risk of endometrial cancer. Most epidemiological studies focus on oestrogens alone, which are used mainly by hysterectomised women. For women with an intact uterus, present formulations involve combined oestrogen-progestagen therapy, on the basis of findings that progestagens counteract the proliferative effects of unopposed oestrogens on endometrial tissue. However, there are several unresolved questions, including how progestagens should be prescribed and the effects of extended use.

The MWS provides the most extensive data yet to address these questions. By comparison with never using hormones, continuous combined therapy was associated with reduced risk, cyclical combined therapy with no alteration in risk, and oestrogens alone with increased risk. The study also provides unique data on tibolone, a drug that is available in certain areas of Europe, but not in the USA. Tibolone was associated with significant increases in the risk of endometrial cancer, in a manner similar to unopposed oestrogen therapy, leading to a need for cautious clinical use. Although observational in nature and not a clinical trial, the study’s extensive data enabled adjustment for a wide variety of other lifestyle factors.

In addition to allowing an evaluation of the effects of different formulations and regimens, the design of the MWS enabled an assessment of hormone effects according to characteristics of the users. Thin women are more likely to use hormones than heavier women, who face higher risks of developing endometrial cancer. Thus, of particular interest is whether hormonal risks differ according to the users’ anthropometric characteristics. In the MWS, relative risks for all formulations were highest in thin women. Continuous or cyclical combined therapy was associated with substantially reduced risks in obese women, whereas tibolone led to substantially increased risks in normal and overweight women.

A full understanding of clinical implications requires consideration of relative risks, which compare cancer rates in exposed individuals (eg, exogenous hormone users) with rates in unexposed women (eg, never users), and absolute risks of disease incidence. The MWS responsibly provided both. The absolute risks emphasise the importance of obesity as a major predictor of the risk of endometrial cancer, but this pertained mainly to individuals who were not exposed to exogenous hormones. There was no gradient in risk with obesity in users of continuous combined therapy. The MWS shows that any benefits for endometrial cancer associated with continuous combined therapy are far outweighed by risks for breast cancer, which is adversely affected by this therapy. The absolute risks also remind us of the need to consider multiple disease outcomes when balancing risks and benefits.

A limitation of the MWS was that women were followed up, on average, for only 3.4 years. Of concern for endometrial cancer are recent data suggesting that long-term use of continuous combined therapy might carry increased risks compared with non-use. It is therefore noteworthy that the MWS showed that the risk for users of 5 years or more approximated that of never users. Additional follow-up of women in this study will be important for clarifying effects of long-term use. However, given the complex and dynamically changing nature of hormonal use, other studies will be needed to fully understand the spectrum of effects.
Although for years it was believed that hormones would reduce risk for various diseases, well-designed randomised trials have failed to substantiate these claims.1–3 Hormones effectively reduce the risk of fractures, but do not reduce the risk of most coronary, cerebrovascular, and cognitive events. There is also accumulating evidence that hormones might increase the risk of ovarian cancer, although reductions in the risk of colorectal cancer seem probable.4,5

Hormones clearly remain the most effective therapy for menopausal symptoms, their original indication when first marketed.6 Thus the important clinical question is how hormones can be prescribed in a fashion that will allow women to receive the greatest benefits without commensurate risks. To minimise cancer and other risks, clinicians should prescribe the lowest possible dose of oestrogen for short periods of time.10–11 Fortunately, recent evaluations support the idea that oestrogens prescribed at low doses (eg, 0.3 mg a day conjugated equine oestrogen) are generally as effective in controlling menopausal symptoms as the traditional higher doses.12

The million dollar—or pound or euro—question is: how long can women stay on hormones without adverse effects? Although the benefits of short-term use of hormones during the perimenopausal period appear to outweigh risks,12–13 how should patients be counselled as the time from menopause progresses? Given that women face various physiological changes as they age, other approaches for symptom relief and disease prevention must be found to replace the role once filled by hormone pills.9 Oestrogens given locally can alleviate urogenital symptoms, whereas vaginal lubricants might improve sexual activity. Some antidepressants, including serotonin and norepinephrine reuptake inhibitors, provide moderate relief from hot flashes. Regular physical activity, both cardiovascular and weight-bearing, can maintain cardiovascular health, preserve strong bones, and ward off obesity’s many adverse health outcomes. A balanced and moderate diet, supplemented by calcium (1500 mg a day) and vitamin D (400–600 IU daily), will prevent osteopenia and osteoporosis (which may also be reduced by bisphosphonate therapy). Regular use of aspirin might be warranted for cerebrovascular health. Continued mental vigour appears to decrease the risk of Alzheimer’s disease.

But for women whose menopausal symptoms persist, chronic oestrogen (with or without a progestagen) might be required. For these women and their clinicians, continued research on the long-term risks and benefits of hormone therapy, screening modalities, and effective risk communication remains an important priority.

*Louise A Brinton, James V Lacey Jr, Edward L Trimble
National Cancer Institute, Bethesda, MD 20892, USA
brinton@nih.gov

We declare that we have no conflict of interest.


Still looking for answers in COPD

See Articles page 1552

Despite rapid developments in our understanding of the pathophysiology of chronic obstructive pulmonary disease (COPD),1 clinical progress in outcomes that matter to patients, such as survival, shortness of breath, and exacerbations, has been slower. In today’s Lancet, the result of the Bronchitis Randomized NAC Cost-Utility Study (BRONCUS) is another example of this disappointing trend. Rennard wrote that two major problems facing clinicians treating patients with stable COPD are underdiagnosis and a nihilistic attitude.2 I might add a third—ineffective therapy.
Tobacco smoke contains a high concentration of oxidants, and oxidative stress is a significant contributor to the pathogenesis of COPD. The consequences of increased oxidative stress in the lung include increased transcription of inflammatory genes, protease activity, and mucus secretion. Systemically, skeletal muscle function is impaired, and oxidative stress might contribute to glucocorticoid resistance. These changes are associated with increased exacerbations, decline in forced expiratory volume in 1 s (FEV₁), and poor quality of life.

Consequently, there has been much interest in antioxidant therapy for COPD. There is an epidemiological relation between dietary intake of antioxidants in fruits and vegetables and improved lung function. Antioxidants, such as N-acetylcysteine, and antioxidant mimetics are also being tested. BRONCUS is the most comprehensive study of a pharmacological antioxidant.

Acute exacerbations are an important cause of morbidity and mortality in COPD. Earlier studies in small numbers of patients over short durations suggested that N-acetylcysteine reduced exacerbations, perhaps by upregulating the glutathione system. Given the broad effects of antioxidants on inflammation and bacterial adherence in COPD, the lack of effect of N-acetylcysteine on patients with baseline exacerbation rates of 2.5 a year is disappointing. Bronchodilators (e.g., tiotropium, salmeterol, formoterol, and inhaled steroids, and combinations of inhaled steroids and long-acting β₂ agonists, reduce exacerbations. Surprisingly, N-acetylcysteine reduced the exacerbation rate only in those patients not taking inhaled corticosteroids. Oxidative stress accounts for the relative steroid resistance of certain aspects of inflammation in COPD due to reduction in histone deacetylase activity in alveolar macrophages. Thus one might have assumed a complementary role for N-acetylcysteine and inhaled corticosteroids.

The annual rate in decline of FEV₁ has been used as a surrogate marker for the natural progression of COPD (Table). Smoking cessation improves this rate decline, and slows the accelerated decline from 60 mL a year in smokers to 30 mL a year in ex-smokers. Slowing the rate of decline in FEV₁ is difficult in COPD. A meta-analysis of six studies of inhaled corticosteroids in 3571 patients showed a reduction in the rate of decline in FEV₁ of −5 mL a year (95% CI −11.2 to 1.2 mL, p=0.011) compared with placebo. Therapy with ipratropium bromide failed to change the slope of the FEV₁ decline in the NIH Lung Health Study. Even augmentation therapy for those deficient in α₁-antitrypsin was inconclusive in slowing the decline in FEV₁. Ongoing long-term studies, such as TORCH, with an inhaled steroid plus a long-acting β₂ agonist and UPLIFT, a study of tiotropium, will provide additional data. The inability of N-acetylcysteine to reduce FEV₁ decline is not surprising given this history. Perhaps decline in FEV₁ is not the best outcome for COPD, and future studies should be powered on mortality or composite outcomes, such as BODE (a composite for body-mass index, airflow obstruction, dyspnoea, and exercise capacity).

One silver lining in BRONCUS is the effects of N-acetylcysteine on the secondary outcome of lung volumes. The reduction in functional residual capacity is intriguing. As an obstructed patient becomes more active and increases ventilation, progressively more air is trapped in the lung, causing dyspnoea. Bronchodilators, including tiotropium, salmeterol, and formoterol, reduce hyperinflation, sometimes in severely obstructed patients despite little change in FEV₁. The phosphodiesterase E4 inhibitors have modest effects on FEV₁, but seem to reduce hyperinflation to the same extent as N-acetylcysteine.

Is BRONCUS the final word on N-acetylcysteine? It is far and away the best of the many trials published on this theoretically promising agent dating back to 1976. N-acetylcysteine is safe and higher doses could be used. But N-acetylcysteine has no effect during acute exacerbations when added to bronchodilators and corticosteroids. Research has convincingly shown the need to supplement endogenous antioxidants and modulate oxidative stress. The fact that BRONCUS was negative on its primary outcomes should redirect investigators to pursue other approaches, such as antioxidant mimetics and newer antioxidants.
Comment

James F Donohue
Division of Pulmonary & Critical Care Medicine,
University of North Carolina School of Medicine, Chapel Hill,
NC 27599, USA
jdonohue@med.unc.edu

I declare that I have no conflict of interest.


Extending the benefits of deworming for development

In today’s Lancet, Charles King and colleagues, in a meta-analysis of functional disability due to schistosomiasis, show that schistosome infection is associated with significant anaemia, chronic pain, diarrhoea, reduced exercise tolerance, and malnutrition.

Previous assessments of the public-health relevance of schistosomiasis have focused mainly on symptomatic morbidity and late-stage disease. 1 Van der Werf et al estimated clinical morbidity associated with schistosome infection in sub-Saharan Africa—eg, 70 million cases of haematuria, 18 million cases of major bladder-wall abnormalities, and 10 million cases of major hydro-nephrosis associated with Schistosoma haematobium. 2 To this burden of disease, King and colleagues convincingly add much subtle morbidity. Disability-adjusted life years (DALYs) are increasingly used as a non-monetary measure of the impact of mortality and morbidity caused by a disease. Subtle functional disability is very relevant in soil-transmitted helminthiasis, another group of highly prevalent helminths. 3 Regular treatment is clearly linked with physical and cognitive development, educational outcome, and economic development. 4, 5 Consequently, the estimated DALYs lost due to these infections have been rated higher than those lost to schistosomiasis.

In 2001, a WHO Expert Committee concluded that the current figure for DALYs lost to schistosomiasis was considerably underestimated, and recommended that the figure should be revised to take into account the subtle morbidity induced by this disease. 6 King and colleagues provide this missing information. As a result, we can readjust the disability weight currently assigned to schistosomiasis—and the resulting DALYs lost—to a much higher level. King’s results should trigger a better quantification of the development impact of schistosomiasis. Beyond this, their analysis should encourage a...
comprehensive re-evaluation of the burden on human and economic development of a group of highly prevalent but still concealed communicable diseases of the poor, including soil-transmitted helminthiasis, lymphatic filariasis, onchocerciasis, cysticercosis, echinococcosis, foodborne trematode infections, and trachoma.

The past 20 years of schistosomiasis control have been characterised by two major advances. The first is the acknowledgment that even in areas where reinfection is intense, regular chemotherapy can effectively control morbidity. The second is the endorsement in 2001 by the World Health Assembly of a novel public-health strategy for the integrated control of soil-transmitted helminthiasis and schistosomiasis. The aim of this strategy, tailored specifically for areas with high transmission, is to remove the disease burden by regular treatment of high-risk groups within a broader context of preventive measures such as improvement of living conditions and hygienic behaviour.

When possible, regular treatment should be delivered through existing channels for the sake of sustainability. School health-programmes, also targeted at non-enrolled school-age children, are an excellent vehicle for the delivery of integrated interventions to a fundamentally high-risk group. Recently, we have also seen a multiplication of country experiences for the delivery of deworming to preschool children, packaged with vaccinations and/or vitamin A distribution. The community-directed treatment approach used in the onchocerciasis control-programme might be an option for delivery of combined treatment packages to remote communities.

We need to strengthen the links between deworming programmes and other chemotherapy-based programmes against endemic diseases affecting poor people. The delivery channels we mention above provide realistic opportunities for the health system to extend its capacity for the packaging and delivery of a series of simple health interventions to those most in need. The combined delivery of antiparasitic treatment is likely to be highly cost effective because most drugs are today cheap or donated.

However, in the current context, two concerns need to be raised. The first is the need to ensure sustainability of delivery, because regular treatment will have to be delivered for a long time before improvement of living conditions will eventually provide a permanent solution. The second concern is the potential limitation of a chemotherapy-based strategy should drug resistance arise. We therefore believe that appropriate tools need to be developed and mechanisms put in place to enable monitoring of any reduction in drug efficacy so that strategic changes can be made in a timely manner. We also believe that research for new drugs and new control tools, such as the possible development of a hookworm vaccine, should be pursued.

King and colleagues have added a further dimension to the effect that chemotherapy against schistosomiasis may have on disability. We believe that this novel information adds strength to the process of development of a comprehensive public-health strategy to control the burden of chronic endemic diseases in the world. We also hope that such a strategy will yield a high return on investment in terms of contribution towards reaching the Millennium Development Goals.

* Lorenzo Savioli, Dirk Engels, Hiroyoshi Endo
Parasitic Diseases and Vector Control (LS, DE), and Communicable Diseases Control, Prevention and Eradication (HE), WHO, 1211 Geneva 27, Switzerland
savioli@who.int

We thank David Brandling-Bennett, David W T Crompton, Alan Fenwick, Peter Hoetz, David Molyneux, and Martin G Taylor for their advice and suggestions. We declare that we have no conflict of interest.

50 years ago, on Dec 23, 1954, the first kidney was successfully transplanted, with an identical twin brother as donor. The graft functioned well without immunosuppressive drugs for 9 years until relapse of disease. Kidney transplantations between non-identical twins or non-twins were less successful, implicating the importance of close relationship. In the 1950s people were not aware of HLA matching or panel reactive antibodies as a cause of graft failure. However, in the past 50 years, renal transplantation has become increasingly successful with rates of graft survival now over 90% after the first year and about 50% after 10 years. The increased success of renal transplantation is due to increased knowledge about HLA matching, immune responses directed against graft antigens, and immunosuppressive drugs. Currently, acute rejection episodes can be treated successfully in most patients. On the other hand, about half the graft recipients encounter chronic rejection. The pathogenesis of such rejection remains uncertain and various factors might play a role, including immune-mediated damage. Chronic rejection is one of the most important challenges for successful renal transplantation.

In this issue of The Lancet, Gerhard Opelz describes, for the Collaborative Transplant Study, that panel reactive antibodies are strongly associated with long-term graft loss in HLA-identical sibling-kidney transplantation. The investigators analysed over 4000 HLA-identical (matched for A, B, and DR) sibling transplantations, compared with over 160 000 cadaveric transplantations. In the first year after transplantation, the presence of panel reactive antibodies was associated with significant reduction in graft survival in the cadaveric transplant group, but there was no effect in the HLA-identical sibling transplantations. However, long-term follow-up in the HLA-identical sibling group showed a major effect for panel reactive antibodies: in recipients with high levels of panel reactive antibodies (>50%), the number of functioning grafts was significantly lower at 10 years after transplantation. Thus the effect of panel reactive antibodies was highest in the first months after renal transplantation in cadaveric transplantations, started after the first year in HLA-identical sibling transplantations, and affected graft survival in the 10-year follow-up in the HLA-matched group.

What are the antibodies recognising? Before the transplantation, lymphocytotoxic panel reactive antibodies were assayed by incubation of serum with a panel of lymphocytes (from random donors of blood) in a dye-exclusion test. The level of presensitisation is determined by the number of donors that is recognised by the patient’s serum. The antibodies might bind both HLA and non-HLA antigens on the lymphocytes. These different reactivities cannot be discriminated with the assays used in Opelz and colleagues’ study. The effect of HLA antibodies might be more direct and result in the rapid effects found in the mismatched cadaver-transplant group. On the other hand, antibodies reactive with non-HLA antigens might be directed against minor histocompatibility antigens. The effects of these antibodies may be protracted and take years to induce damage. Alternatively, epitopes might be hidden or cryptic, and only exposed on damage to the graft. Hidden epitopes might be exposed by acute rejection episodes, as suggested by the finding that episodes of acute rejection are important risk factors for the development of chronic rejection. Unfortunately, data on the number of acute rejection episodes, which could partly explain the differences observed in Opelz's study, are not available in the registry. More specifically, recent molecular analysis of kidney biopsy specimens during acute rejection revealed strong heterogeneity, including a subset with major B-cell contribution. These different subtypes of acute rejection showed different prognostic implications and careful examinations may yield important information for the pathogenesis of chronic rejection.

The high level of panel reactive antibodies might represent high responsiveness to antigen encounter. Previous alloantigen sensitisation (eg, during pregnancies, previous transplantations, or blood transfusions) may then be responsible for the high levels of panel reactive antibodies. Indeed Opelz and colleagues describe an association with female sex and the number of blood transfusions. Extrapolation of this phenomenon also suggests that the recipient might develop antibodies against any other antigens that are
distinct between the new kidney graft and the recipient. If, hypothetically, antigens in the kidney itself (that are not present on lymphocytes) are variable or polymorphic between individuals, this might result in an antibody response by the high-responsive recipient.

In-situ evidence for humoral involvement is derived from the finding of C4d deposits in rejected kidney allografts that correlate with circulating anti-HLA antibodies.5 In addition, immunosuppression with mycophenolate mofetil, which inhibits not only T-cell function but also antibody production by B cells, seems promising for long-term function of the transplant. The risk of chronic rejection is decreased in patients treated with mycophenolate mofetil.6,7 The effect of this drug is additive to the effect of HLA matching,8 suggesting that mycophenolate mofetil not only inhibits anti-HLA antibodies but also other antibody responses.

Antibody responses receive increasing interest in chronic rejection and specificity, affinity, and pathogenicity need to be investigated to estimate their contribution. The antigens recognised by the antibodies might be antigens on lymphocytes but may also be antigens that are specifically expressed in the graft. An example of tissue-specific antigens recognised by antibodies of patients with a specific form of chronic rejection, transplant glomerulopathy, is the glomerular basement-membrane protein agrin.9 Agrin is a heparin sulphate proteoglycan involved in maintenance of glomerular filtration. Antibodies against the side chains of agrin can induce proteinuria and duplications of the glomerular basement membrane.10 Antibodies against agrin were not found in patients with chronic allograft nephropathy; however, these patients might have antibodies with other specificities.

The results described by Opelz and colleagues support a search for antibodies in renal transplant recipients with chronic rejection.

Simone A Joosten, *Cees van Kooten
Department of Nephrology, C3-P, Leiden University Medical Centre, 2333 ZA Leiden, Netherlands
kooten@lumc.nl
We declare that we have no conflict of interest.

Shanghai plans responses to AIDS

On Dec 15, 2004, a meeting was convened in Shanghai, China, to discuss options for laws dealing with HIV/AIDS at the provincial and municipal levels. The primary agencies responsible were the Shanghai Academy of Social Sciences, the Shanghai Law Society, and the Institute of Legislation and Temple University in the USA. With a population of some 13 million, Shanghai is China’s largest city and enjoys a legal status equivalent to a province, enabling it to create its own law. Participants in the meeting were mainly leaders of the municipal government (including members of the Municipal People’s Congress and the health department), and researchers from the Shanghai Academy of Social Sciences and other research institutions. The meeting showed Shanghai’s commitment to effective HIV/AIDS legislation and the challenge of steering HIV/AIDS policy through the shoals of fear and moral disapproval of drug use and prostitution.

Both the content and process of law reform were addressed. A law-reform exercise in Australia in the early 1980s was offered as an example of a consultative process in which affected populations were encouraged to share their experiences. Over 2 years, non-governmental organisations were encouraged to come forward to help shape the law, law that in turn supported an environment of partnership between government and community.

At the Shanghai meeting, three important areas of concern were identified, each tied to a different legal strategy for HIV/AIDS control. First, there remains a great deal of anxiety in some of those in the legislative process about low-risk exposures and attempts to intentionally transmit HIV. Speakers and participants offered anecdotes about vengeful people with HIV deliberately trying to infect others, and robbers using infected needles to intimidate victims. Similarly, some speakers alluded to the necessity of testing hospital patients to protect medical staff, and disclosing HIV test-results to employers to allow safety-related job reassignment.

Fears about deliberate exposure to HIV have elsewhere come to be regarded as exaggerated. Mandatory testing and disclosure as a means of protecting individuals from exposure to HIV is elsewhere seen as unnecessary and even counterproductive. The universal precautions approach, in which workers are trained and equipped to manage occupational exposure to blood or infectious body fluids, is now well established. The Chinese Centres for Disease Control plans to promote this approach, but it does not yet appear in Chinese HIV/AIDS legislation. Including a universal precautions approach in Shanghai’s legislation, and providing the support for the necessary training and equipment, could set a valuable example.

A second important theme was the need to enhance the involvement of non-governmental organisations in HIV/AIDS prevention and care in China. People in vulnerable populations, such as drug users and sex workers, have begun to organise in China to advocate for policy and law that respects their rights. The Shanghai meeting was notable for the participation of an openly gay lawyer, Zhou Dan. Although Chinese law formally contemplates a vital role for non-governmental organisations in HIV/AIDS prevention and response, there are many legal and social barriers to the development of a vigorous sector for non-governmental organisations in
China. The requirements for establishing non-governmental organisations are set for the most part at the national level, and include provisions requiring such organisations to have a government or official sponsor, and that allow only one non-governmental organisation on a topic for each jurisdiction. Partnerships between Chinese non-governmental organisations and non-Chinese funders and providers of technical assistance are also limited. Whilst eliminating these barriers would require national legislation, local governments can encourage the participation of non-governmental organisations by providing greater funding to these organisations and actively involving them in planning and evaluation of HIV policies and programmes, including the development of good law. Creative legal work might identify ways in which provinces and municipalities can reduce bureaucratic barriers to non-governmental organisations in the absence of reform at the national level. Representatives of the Shanghai administration reported that the problem of non-governmental organisations will be addressed in the municipal legislation.

Finally, the meeting also highlighted the importance of policy implementation. The approval of needle-exchange and drug-treatment programmes on paper, or the passage of privacy and discrimination protections, is not sufficient to assure that services are actually delivered in a way that meets policy goals. Legislative planning must include funding for training and monitoring of performance. The implementation problem is particularly great in the area of privacy and anti-discrimination law. Vitally, there was no disagreement that people with HIV should be protected from discrimination and from the inappropriate release of their medical information. At the same time, however, it was also widely accepted that China lacks an effective system for the enforcement of these basic human rights. An enforcement mechanism, such as a Human Rights Commission charged with investigating discrimination cases and promoting compliance, as well as legal services, will be necessary to provide real protection that people with HIV can rely on.

*Scott Burris, Joanne Csete, Xia Goumei, Zhou Dan, Bebe Loff*
Temple University Beasley School of Law, Philadelphia, Pennsylvania, PA19122, USA; Johns Hopkins School of Hygiene and Public Health, Baltimore, Maryland, USA; and Center for Law and the Public’s Health, Johns Hopkins School of Hygiene and Public Health, Baltimore, Maryland, USA (SB); Canadian HIV/AIDS Legal Network, Montreal, Quebec, Canada (JC); Centre for Public Policy Study of HIV/AIDS, Shanghai Academy of Social Science, Shanghai, China (XG); Richard Wang & Co, Shanghai, China (ZD); and Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria, Australia (BL)
scott.burris@temple.edu

We declare that we have no conflict of interest.


Dissolving the dilemma over forced treatment

How should doctors respond when patients who have the capacity to make their own medical decisions refuse treatment that is in their interests? Should societies support the right of patients to make their own decisions? Or should clinicians be allowed to force medically indicated treatment on patients who have decisional capacity?

Many western countries allow patients who have decisional capacity to refuse treatment that is in their interests. Recently, Michael Gross forcefully questioned this consensus, partly defending the Israeli Patients’ Rights Act which allows clinicians, under specified conditions, to force treatment on patients who have decisional capacity. We argue that this dilemma, widely thought to require a wrenching choice between respecting patients’ autonomy and promoting their interests, dissolves on analysis.

Decisional capacity refers to the set of abilities, including the abilities to understand and deliberate, that individuals must have to make their own decisions. Because the necessary abilities depend on the nature of the decision in question, individuals can have the...
capacity to make one decision, but not another. Decisional competence refers to a societal decision, often made by legislatures and courts, to respect the decisions of some individuals, but not others. Possession of decisional capacity is the most obvious, but not the only, reason to respect individuals’ decisions.

Clinicians should force treatment on patients who have decisional capacity only when there are compelling reasons to do so, ones that both outweigh the importance of respecting patients’ autonomy and justify forcibly interfering in their lives. Clinicians should not force treatments on patients, such as antibiotics for a sinus infection, that would promote their interests only to a minor extent. Neither should clinicians force treatments on patients, such as experimental therapies, that offer only a minor chance of promoting their interests.

When capacitated patients refuse treatments that would clearly and substantially promote their medical interests—eg, refusing antibiotics that would cure an otherwise fatal infection—clinicians first should attempt to remove any obstacles that might be interfering with the patient’s decision-making. Clinicians might, for example, provide painkillers to patients with pain that interferes with their ability to concentrate. If the patient continues to decline, the clinician should explain again why the treatment is in the patient’s medical interests.

Clinicians should not attempt to coerce or manipulate patients who refuse treatment that is in their medical interests. At the same time, respect for patients’ autonomy does not imply that clinicians should simply accept whatever decision capacitated patients happen to make. Instead, clinicians should engage with patients, treating them as rational beings, capable of understanding and responding to reasons. Sincere efforts at persuasion, providing reasons to which the patient may agree, engage patients’ autonomy; they do not constitute coercion or manipulation.

During this process, clinicians should remember that decisions that conflict with patients’ medical interests might nonetheless promote their overall interests. A patient might refuse to start chemotherapy today, to attend a daughter’s wedding tomorrow. Clinicians should also remember that living according to a set of fundamental values is part of a flourishing life; hence, is in patients’ overall interests. The decision to refuse a life-saving blood transfusion, for example, might conflict with the medical interests of life-long Jehovah Witnesses, yet promote their overall interests.

A dilemma is thought to arise on the failure of this interactive process, when patients continue to refuse treatments that would clearly and significantly promote their medical interests without a reason to believe that this refusal is in their overall interests. In our view, this scenario poses no dilemma at all.

While it is widely agreed that decisional capacity does not depend on the content of particular decisions, the content of patients’ decisions can provide evidence about their decisional capacity. Several factors might explain why a patient refuses treatment that would significantly promote their medical interests, in the absence of a reason to believe that this decision promotes their overall interests. The patient may fail to understand the severity of their condition, be depressed, or have a morbid fear of needles. The presumed dilemma dissolves once we recognise that all the potential reasons for refusal point to the conclusion that the patient lacks one or more of the abilities necessary for decisional capacity.

In these cases, clinicians need not appeal to ethicists or courts to determine whether they can ethically and legally over-ride the decisions of patients who possess decisional capacity. Instead, clinicians should appeal to an appropriate surrogate to determine how to treat a patient who has been found to lack the capacity to make this decision at this time.

Vardit Ravitsky, *David Wendler*
Department of Clinical Bioethics, National Institutes of Health, Bethesda, MD 20892, USA
dwendler@nih.gov

Our views do not represent any position or policy of the National Institutes of Health, Public Health Service, or Department of Health and Human Services. We declare that we have no conflict of interest.

World Diabetes Day: footing the bill

The world-wide prevalence of diabetes exceeds 200 million, which is predicted to rise to over 300 million in the next 20 years.1 There is no cure and the condition can be devastating in its consequences. The aim of World Diabetes Day next Nov 14 is not only to draw attention to the impact of the disease on health and the economy, but also to emphasise how this burden can be reduced. The spotlight of World Diabetes Day in 2005 falls on the foot.

Disease of the foot ranks among the most feared complications of diabetes, and yet is one of the most neglected. It is neglected partly because the pathogenesis is multifactorial and the presentation complex,2 and partly because doctors and nurses rarely receive formal training in the subject. Optimum management requires the effective integration of many health-care professionals with different skills.3,4 The result is that diabetic foot-disease continues to exact a ghastly toll in patients, especially in those who are least able to cope: the elderly, the impoverished, and those with limited access to health-care services. Probably more than a million people lose a leg because of diabetes each year, which means that one major amputation is done somewhere in the world about every 30 s. For those who lose a leg, life never returns to normal. The patient often becomes permanently dependent on the support of others and both the patient and their family may be deprived of their livelihood. Amputation is also associated with a high mortality, even in developed countries.2 The social and psychological consequences can be immense,3 as can the costs to health-care services and the wider community.2 The costs of major amputation far outweigh those of preventing the operation. Because there is good evidence that the incidence of major amputation can be reduced by the provision of accessible expert integrated care,4 the problem of the diabetic foot is one which should receive far greater attention—from the public, from professionals, and from health-care planners.

The Lancet is keen to help draw attention to the enormity of the global burden of diabetic foot disease, and to support the efforts of those working to lessen the suffering and cost which result from this neglected problem. For this reason, the theme of The Lancet on Nov 12—to coincide with World Diabetes Day—will be on the foot in diabetes and related aspects of wound healing. The Lancet therefore invites submission of original research papers—full articles or research letters—on this theme. Details of the format can be found on the journal’s internet site. Manucripts should be addressed to William Jeffcoate at The Lancet, and should arrive no later than Aug 1, 2005.

*William Jeffcoate, Karel Bakker
C/o The Lancet, 32 Jamestown Road, London NW1 7BY, UK (WJ); and Consultative Section on the Diabetic Foot, International Diabetes Federation, Heemstede, Netherlands (KB)
w'lliam.jeffcoate@lancet.com

Manuscripts can also be sent to WJ, c/o The Lancet, 360 Park Avenue South, New York, NY 10010, USA. WJ is a Contributing Editor at The Lancet. KB declares that he has no conflict of interest.

Comment

thelancet.com—new horizons

The rapid emergence of the internet in the early 1990s created a new world for scientific journals and readers: an electronic medium for accessing information hitherto dominated by print and paper and archived on metres of shelving in the academic library. July, 1996, was an important milestone for The Lancet; 173 years after the first print issue, subscribers could access around a quarter of the weekly journal’s content via their computer. By March, 1998, full-text access was a reality, and 120,000 non-paying users of thelancet.com had visited the site to view article summaries and other limited free journal content. A new online Lancet community was being created.

Today there are more than a million users of thelancet.com, 30 times the number of print subscribers. This number does not marginalise the value of the print journal—it merely serves as a powerful illustration of how medical journals, by their very nature, command a towering online presence among their core subscribers and non-subscribing web users, including physicians, other health professionals, the mass media, and the wider general public.

This week thelancet.com relaunches again. Over the coming months and years the website will evolve from its origins as a relatively static electronic version of the print journal into a dynamic interactive forum of information exchange and debate. Importantly, the relaunched website has been designed to ensure that the journal and the three monthly specialty titles (The Lancet Infectious Diseases, The Lancet Neurology, and The Lancet Oncology) are brought closer together. Users of thelancet.com will be able to submit electronic comments about journal content. Comments accepted for online publication will appear within 24 h of submission. Enhanced journal searching will be a key feature of the new website: by specific journal, across all Lancet journals, or by MEDLINE searching. Article Collections will enable users to browse content by specialty, with links to relevant content from all Lancet journals. Global Medicine has been developed as the first such collection, reflecting The Lancet’s sustained commitment to publishing research and analysis about global public-health issues free of charge.

Fast dissemination of important advances in medicine remains a priority for The Lancet. Enhanced website functionality provides the forum for the publication of key journal content before it reaches the pages of the print issue. Early online publication is now routine for many scientific journals; we will aim to publish half of our original research articles in 2005 online ahead of print publication. Such rapid publication will enhance The Lancet’s fast-track process, and enable “normal-track” of original research within weeks of article acceptance. Early online publication of original research for the monthly specialty journals—recently started with the first original research articles in The Lancet Oncology—will also be an important aspect of our early online publication output.

Happy though we might be with the new features and enhanced functionality of the relaunched website, the real excitement lies ahead. Over the coming months we will commission regular web-only content to give further insight into the key issues in medicine, with a strong international focus. We look forward to collaborating with our Editorial Consultants, listed at the end of this week’s issue, to assist us in the delivery of new online content. This week’s relaunch of thelancet.com will be the start of a new era in electronic publishing for The Lancet journals. Please join us in shaping the future of the lancet.com by e-mailing comments to webeditor@lancet.com.

Richard Lane, Richard Horton
The Lancet, London NW1 7BY, UK
webeditor@lancet.com
Germany’s reforms have yet to deliver insurance savings

Tempers have flared in Germany over health reforms designed to tackle huge deficits in public health funds. Health minister Ulla Schmidt says these changes will ultimately make health insurance cheaper. But, says Ben Aris, this promise has not placated the plan’s dissenters.

“Its outrageous”, says Kirsten Schneider, a civil servant at Germany’s Foreign Ministry. “I didn’t have any of these tests so what are they doing on my bill?” Schneider, who is 6 months pregnant, took the unusual step of actually checking her gynaecologist’s bill and found that several tests listed were ones she had never had. In addition, there was a bill of €10 for “consultation” that actually consisted of a phone call to cancel and reschedule an appointment.

Germans have come to expect a generous universal health service since the Prussian leader Otto von Bismarck set up the system at the end of the 19th century. Bismarck founded the public health funds, the so-called Krankenkassen, as a way of enticing workers away from the growing Social Democratic Party (SDP).

A hundred years on, that same SDP has been battling to cut some of the fat off the now-bloated system, but is facing an uphill battle against popular dissent. The public health funds have been running at a loss for nearly a decade and have become a major drain on the beleaguered federal budget.

From Jan 1, 2004, as part of the ruling SDP’s Agenda 2010 reform plan, new rules came into force that have cut costs and put public health funds back into profit for the first time in 10 years. But these reforms have failed to deliver on their main promise: to make mandatory medical insurance cheaper.

The headline change was the introduction of a €10 charge for the first visit to the doctor each quarter. The annual extra €40 citizens will have to pay is a drop in the ocean compared with the €3000 a year in contributions, but the charge strikes to the heart of Germans’ expectation of Bismarckian state-sponsored health care, and tempers have flared.

The nominal charge is the first attempt to introduce some accountability and force patients to think a little before going to the doctor. For example, as Germans pay only the monthly premiums, patients often ask for third or fourth opinions before accepting a diagnosis, leaving the Krankenkassen to pick up the bill.

Implementation of the new rules has not gone smoothly. The health ministry accused German doctors of deliberately sabotaging the government’s health-care reform last spring. A survey found some doctors took advantage of confusion over the introduction of limited payments to levy charges they are not entitled to.

A report presented in Berlin in April last year lists cases of abuse in the first months of 2004 including an eye specialist asking for €25 for eye tests, doctors that asked for the quarterly €10 fee every visit, and another doctor refusing treatment unless the patient accepted a much more expensive private bill. All justified these charges by citing the reforms.

Chancellor Gerhard Schröder is struggling against a wave of popular dissent, culminating in mass protests across the country at the end of last year. While the majority of Germans are paying the €10 fee, the estimated 300,000 people who refuse have put a substantial strain on doctors’ associations.

Under the new rules, doctors’ associations are responsible for collecting the fee even though the money ends up in public health funds. In March, the increasingly frustrated doctors’ association in the North Rhine region tried to sue a patient who refused to pay the quarterly fee, but lost the case on April 19. “If we pursued every case, we would be stuck with €3.5 million in court costs and €500,000 in administrative costs”, says the exasperated North Rhine association spokesman Klaus Enderer. “This madness has to end.”

The ruling has made associations even more worried because it sets a precedent that could lead to a mass boycott of the €10 fee.

Despite opposition, the reform does seem to have achieved one of its main goals: reducing Germans’ casual attitude toward visits to the doctor’s surgery. In the first two quarters of 2004,
Health minister Ulla Schmidt claims that the reforms have saved billions.

Visits were down 10% and 7%, respectively, compared with a year earlier. However, critics worry the new fee hits the poor disproportionately; in the run-down district of Neukölln, in Berlin, doctors’ visits plunged 16.8% during the first 3 months of 2004 after the reform was introduced, compared with a Berlin-wide average drop of 8.6%.

Health funds spent a total of €130 billion on care in 2003, but Health Minister Ulla Schmidt was crowing by the middle of last year that the reforms had shaved €10 billion off the bill and the public health funds made a profit. Health economist Karl Lauterbach, who advises the government, estimated that public health funds saved €9.5 billion in 2004 after a decade of deficit spending.

The €10 charge has become a cause célèbre but the meat of the reform—and the aspect that will hit patients hardest in the pocket—is the wide ranging re-evaluation of treatments that the Krankenkassen will pay for.

Patients now have to pay for all their non-prescription drugs and contribute towards the cost of prescription drugs. The rules covering claims for travel expenses have been tightened, and the way hospitals are funded has been changed. Since January, 2004, the costs for drugs, sick pay, and patients’ travel expenses all fell by double-digit percentage rates during the first 6 months of 2004, while spending on hospital services, which are usually the single largest outgoing, rose by only 0.1%. All other costs fell by 4.4%, compared to 2003.

The €10 for visits to the doctor and increased patient contributions for prescription drugs accounted for €3 billion of the savings, and the abolition of subsidised non-prescription drugs accounted for the bulk of the rest. The health funds turned a loss of €805 million in 2003, financed by the budget, into a profit of €960 million during the first 6 months of 2004.

Not everyone is happy, however. Critics worry that these reforms will again hit the worst-off citizens. All the new health-care legislation limits deductibles for doctors’ visits and drugs to 2% of a patient’s gross annual income. But this does not include the cost of non-prescription medication.

"Already in January, some patients with very little income had reached that threshold with the fees they paid. And on top of that they had to pay for their non-prescription drugs", says Stefan Egeton, health expert at the Federation of German Consumer Organisations.

The next phase of reforms were launched in January. Hospitals are the largest consumers of the public health funds’ money and improved cost control will be bolstered by a switch to the “Diagnosis Related Groups (DRG)” system of billing. Hospitals used to get a flat €150–400 per day for each patient depending on the hospital, but they will now be funded on the basis of the illnesses they treat. The reform is designed to introduce some competition among hospitals and will be phased in over the next 2 years.

"With the new system, the money will be spent where it is medically necessary", Klaus Theo Schröder, the permanent secretary to the German Ministry of Health, said in an interview last month.

However, the doctors’ associations are less happy with the change saying the DRG rates are too low and will not cover costs. They recently warned that 10–15% of Germany’s roughly 2200 hospitals would have to close or merge over the next 10 years because of lack of funds.

Health minister Schmidt is feeling pleased with the progress, but she has failed to deliver on her key promise to reduce the monthly premiums that people have to pay.

The immediate goal of the health reforms was to put health care back on its feet, but it is also part of the Agenda 2010’s wider challenge to make Germany’s economy more competitive. And as employers pay for half of the monthly contributions, health-care premiums are a major addition to labour costs.

Schmidt launched the health-care reforms with the promise of reducing premiums from 14.3% of gross wages to 13.6% by 2005 and 12.6% by 2006. However, by the start of this year workers had seen virtually no change in their monthly contributions.

The Krankenkassen complain that, while they are finally making profits, they are still collectively carrying debts of €8.3 billion built up over the past 10 years, which still need to be paid off. Last year, 28 million patients from the 70 million insured by public health funds saw their premiums fall slightly while 3 million saw them rise. Schmidt has backtracked and says she hopes to see some reduction in premiums this year.

Ben Aris
Minority health care remains a problem for Canada’s leaders

Doubling federal funding for health is a key pledge of the Canadian government’s ambitious plan to strengthen health care. The strategy should appease critics who blame underfunding of health services for recent outbreaks. But will the reforms help those most in need? Paul Webster reports.

It is long past midnight and the temperature is well below zero when a patrol van operated by the Native Men’s Residence, a homeless men’s shelter, finds Jeff, a 35-year-old Ojibway man, at a bus stop near Toronto harbour. He is asleep in his wheelchair. Numbed nearly to the point of speechlessness by cold and by alcohol, Jeff declines the hot food on offer in the van. While arrangements are made to find him a bed for the night, he accepts a bottle of water. Homeless since 1992, he says life got much harder after he lost his feet to frostbite last winter. “I woke with both my feet frozen”, he explains in a slow voice. “They amputated my feet and fitted me with prostheses”. But according to Jeff, his weight gain since the operation now makes it too painful to walk.

Thanks to Canadian government programmes that provide free medical care to aboriginal people, Jeff can get his prosthetic feet changed. But that is not likely to happen in a hurry. “The problem we have with the medical system is getting access to people who understand our culture, and who actually want to help us”, says Brian, a homeless Mohawk man, who approached the patrol van for a cup of soup and a pair of socks later in the evening.

Despite Canada’s universal access to free health care, every significant health indicator for the country’s 1·1 million aboriginal population trails those for the rest of the country. On Canadian native reserves, life expectancy for men is 9 years shorter than for non-aboriginal men. The suicide rate is six times that of the general Canadian population. At 8 births per 1000, aboriginal infant mortality is half again as high as it is for the general population. Diabetes has increased 8-fold among aboriginals since 1980. And for several specific aboriginal groups—most notably the Arctic Inuit—the health indicators are even worse. Canada ranks eighth among the 174 nations graded in the most recent United Nations Human Development Report. If aboriginal Canadians were ranked separately, however, they’d rank 48th.

“This is a really damaging black eye”, says Roy Romanow, a former head of the Saskatchewan government, who recently completed an official review of the national health system. “A country of our wealth, with our reputation for compassion, showing these outcomes for aboriginal health is a really urgent problem.”

But aboriginals aren’t the only people with complaints about Canada’s 40-year-old government-run health-care system. A study of mistakes in Canadian hospitals last year concluded that between nine and 2300 deaths each year are avoidable. Another study in British Columbia, Canada’s westernmost province, found mistakes in a quarter of all prescriptions. And the steadily breaking news of major medical crises—such as the Clostridium difficile hospital outbreak which killed 100 in Quebec last year; the SARS epidemic, which killed 44 in Ontario in 2003; and the tainted blood disasters, which infected hundreds with HIV and hepatitis in the 1990s—all reinforce the sense that the health system is in crisis after a decade of budget cuts. And with Canadian demographers warning that demand for medical services will increase sharply as unprecedented numbers of individuals from the post-war baby-boom generation edge closer to infirmity, the pressure on Canada’s government-run health system—and the elected leaders accountable for it—isn’t lifting.

Last September, Canadian leaders set out to rescue the system with a sweeping 10 Year Plan to Strengthen Health Care. Centring on reforms to paediatric care and increases in staffing aimed at reducing waiting times, the new plan commits the federal government to doubling its financial contributions to what is already the world’s fourth most expensive health system, per capita.

But not everybody is convinced that the new plan will restore a healthy health system. Under the terms of the new plan, Canada’s 10 provinces are no longer required to agree to national standards for care in return for the federal contributions, which currently finance almost 40% of public health spending nationwide. That marks a major, and possibly ominous, departure for the public system as a whole, and, according to Romanow, for Canadian unity. National standards have historically offered a bulwark against local officials’ zeal for service restrictions of the sort that compounded several notable public-health disasters. Now, provinces such as Quebec in the east, and Alberta in the west, are free to limit...
access to publicly-financed services, while encouraging patients to rely more on private services, which already account for 30% of Canadian health expenditures.

Even Michael Decter, a former Ontario health administrator who heads the Health Council of Canada, the organisation charged with monitoring implementation of the 10-year plan, says there is no guarantee it will restore credibility—let alone serviceability—to publicly-financed health care.

"The boomers are starting to hit the system now, looking for minor repair," Decter says. "This is a generation used to getting its pizzas delivered in 30 min, or they're free. They're starting to use the system and they're finding they don't like what they paid for. They want a faster, more modern system."

What's needed, says Decter, are swift remedies for the primary-care system, and upgraded information technologies aimed at saving costs. "During the 1990s, we consolidated acute care and shifted 70–80% of surgical cases to a day basis", he notes. "Now, we must build a more robust primary-care system. If you're dealing with asthma and diabetes in the emergency room, you need better primary care."

Of the US$41 billion in new money devoted to the 10 year plan, almost 15% is earmarked for reducing waiting lists through training, hiring new staff, clearing case backlogs, and expansion of ambulatory and community care programmes. A national physicians' alliance recently recommended waiting-time benchmarks in an effort to pressure the federal government to develop an evidence-based approach.

The physicians suggest that routine hip and knee replacements be done in under 9 months, non-urgent bypass surgery within 6 months, cataract surgery within 4 months, radiation therapy for cancer within 10 days, and diagnostic imaging, including MRI and CT, within 7 days for all patients.

But whether the plan will deliver is anyone's guess, says Decter. "Canadians are saying we squeezed money from the system pretty hard in the 1990s, so let's reinvest. But if there's a recession, and public reinvestment dries up, I think you are going to see a lot more money flowing into private care."

With waiting-time reductions widely viewed as the make or break issue for middle-class Canadians, aboriginal leaders say what is mostly a convenience issue for the affluent is siphoning money away from the severe health challenges they face.

The Canadian government has committed CAN$700 million for aboriginal health programmes, along with a further $745 million for child care, family programmes, and housing on native reserves. Spending strategies include bolstering aboriginal medical education (there are desperately few aboriginal doctors), and new programmes targeting suicide prevention, maternal and child health, and diabetes.

These issues mostly echo Romanow's recommendations. But his main proposal—a call for establishment of Aboriginal Health Partnerships designed to draw aboriginals into health-care administration and delivery—was rejected. "There is a lot of money floating around for aboriginal health. The problem is that it is not pooled, or directed", he says.

Joe Hester, executive director of Anishnawbe Health Toronto, which serves both as a clinic for hundreds of homeless aboriginal people in Toronto's poorest neighbourhood, and as a national advocacy group, says far more money is needed for aboriginal health care, along with deep structural reforms. He worries that the 10-year plan amounts to little more than tokenism. He had hoped that at very least it would acknowledge that large numbers of Canadian aboriginal people now live in big cities. "Most government health programmes are geared toward the rural reserves, where our people traditionally lived", Hester, a Cree, notes. "But large-scale health services, tailored for our cultural needs, are needed in the cities, where we mostly live now."

Speaking in an office pleasantly scented with sage and sweetgrass after an early morning spiritual ceremony, Hester thumbs through literature on Canada's health plan before summarising his views with a provocative question: "If a million non-aboriginal Canadians faced the same health crises our people do, don't you think this plan would focus on radically different priorities?"

Paul Webster
What would it take to reduce our civilisation to the anarchic state grimly encapsulated by Thomas Hobbes, when human life was “solitary, poore, nasty, brutish, and short”? A couple of decades ago, the answer would have been the nuclear winter so powerfully described by Carl Sagan and others. There is still some smart money there, but apocalyptic thinking tends these days to focus on microorganisms. HIV/AIDS led the way, but new variants of influenza, SARS, and a host of emerging diseases now occupy centre-stage. In Angola, health workers are currently attempting to stem the outbreak of the deadly Marburg virus. Even smallpox, celebrated in the 1970s as an eradicated disease, causes some anxiety, with new generations rendered vulnerable.

These modern concerns bestow currency to books on ancient epidemics. A generation ago, a volume on the Black Death (as it came to be called) would have been simply that: a historical exercise analysing a cataclysmic occurrence that affected our ancestors almost seven centuries ago. Since history is always about the present as well as the past, John Kelly’s fine study of plague in medieval Europe takes on contemporary reverberations. “Since history is always about the present as well as the past, John Kelly’s fine study of plague in medieval Europe takes on contemporary reverberations.”

Samuel Pepys’ diary and several other graphic accounts; nothing similar survives for London in 1348. There are, however, some authentic voices to give substance to Kelly’s aim to examine plague as it was experienced. Consider, for instance, the upwardly mobile shoemaker Agnolo di Tura, of Sienna, Italy. He had worked his way up, collecting taxes, engaging in property speculation, and currying favour with local bureaucrats. He was intensely proud of his city, writing with pride of its burgeoning cathedral, and vast piazza. He wrote more about his city than himself, but we know that his wife was above him in the social pecking order, and that he had five sons. We also know, because he told us, that in 1348 he buried his wife and all of his sons, with his own hands.

Agnolo di Tura lived to tell his tale. Other would-be chroniclers of the disaster were not so lucky. In Florence, the former banker and historian of his city, Giovanni Villani, recounted the history of the plague then enveloping his beloved Florence. He traced its slow march westwards, from the Levant and through the Italian cities. His history ends, “And the plague lasted until . . .”; Giovanni Villani never finished this sentence. Elsewhere, trials were stopped because all the witnesses were dead, and even kings and the children of kings died. The people of medieval Europe only had two things going for them. The first was the Age of Faith, and the widespread feeling that things were out of control for a good reason. This permitted stoicism, even fatalism in the face of personal and social tragedy. The second was limited and poor communication, which kept panic to a minimum until the disease actually struck.

When the plague finally did come to each locality—it took some 2 years to spread in a slow-moving wave throughout Europe—the results were everywhere the same, and everywhere different: different local social structures, but everywhere a similar mix of heroism and cowardice, benevolence and greed, decency and malevolence. As Kelly wryly comments, the only thing that hasn’t changed in seven centuries is human nature.

We know a good deal about this episode, even if much of our knowledge is inferential. There are still fundamental issues that are open to debate. One of these is the overall mortality and morbidity. The population of Europe before the plague can only be estimated, and contemporary statements about the numbers of people who died in any locality need to be treated with caution. Kelly offers a sober reading of these, and still leaves us with the global range the continent of between a quarter and a half of the population dead. It is easy to write this, but almost impossible to imagine what this level of devastation would mean, in terms of personal relations and disruption of the ordinary functions of life.
gathering the crops, making bread, and maintaining some sort of social order.

A second perennial puzzle is the disease itself. The clinical picture is not like modern bubonic plague. Victims developed widespread petechiae, had buboes in the wrong places, spat blood, and seemed to pass the disease directly to other people. Many places had higher mortality in the winter than in the summer and early autumn. Lots of domestic animals seemed to have died, but rats went unnoticed. These and other clinical and epidemiological features of the Black Death have recently led several historians and microbiologists to argue that *Yersinia pestis* was not actually the causative organism. Anthrax and various more or less hypothetical viruses have been put forward. It is easier to write a noticeable book on the subject when a completely new microorganism is advanced. To Kelly’s credit, he examines the alternatives carefully, and concludes that conventional wisdom is almost certainly correct. For one thing, strands of *Yersinia* DNA have been found in plague pits. Each of the alternatives poses as many historical problems as the old favourite. Besides, if one looks at the longer European experience with plague, we can see the modern disease evolving. For 300 years, the recurrent plague epidemics were perceived as the same disease. By the plague of London in 1665, the disease’s clinical and epidemiological features can be recognised as “our” plague.

Are there crumbs of comfort in this book, or lessons to be learned? Kelly provides a careful and convincing account of why the Great Mortality happened. Malthusian forces operated in early 14th century Europe, as a series of bad harvests, cold weather, and natural disasters left the European population badly nourished and deprived. But history explains things only after they have happened. There is comfort only in the knowledge that society did survive, and that some of the eventual effects of the plague were salutary. The value of labour rose and some of the worst excesses of the old feudalism were weakened. The plague was not, as Agnola di Tura feared, the “end of the world”. It may even have hastened the coming of the modern world, and the forces that may come back to haunt us.

**In brief**

**Book**  
**End-of-life care**  
One of the most important questions my colleagues and I had when we, in the late 1990s, planned an international comparative study of end-of-life decision making, was whether doing such a study would be feasible. In the Netherlands at that time, we had established a research tradition that was based on the willingness of many physicians to share their experiences. We considered the Dutch tradition of tolerance and transparency to have been an important prerequisite for the success of studies in this area. However, we soon realised that we were less unique than we thought. Working together with researchers from various parts of Europe turned out to be rewarding and far less complex than envisaged. We all shared a common goal of providing valid knowledge to support an important public and professional debate.

The international sharing of values and concordance in developments is also what struck me most about the book.* End-of-Life Decision Making: A Cross-National Study* Robert Blank and Janna Merrick collected information about many aspects of end-of-life care in 12 countries worldwide. Of course, they found much diversity. Euthanasia, for example, has been extensively discussed in some countries, but is virtually ignored in others. Some countries have rather detailed legal regulations on patients’ rights, and in others doctors often just paternalistically do what they believe to be right. In developed countries, death most often occurs in hospitals, whereas in many developing countries people typically die at home under the care of their family.

But these and many other differences are less absolute than they may seem at first sight. First of all, the book proves that end-of-life care is an emerging field in all the countries studied, irrespective of their wealth, culture, or religious background. The medicalisation of the last phase in life and the institutionalisation of dying are unanimously considered undesirable. In all countries, paternalism is to a greater or lesser extent on the way out in favour of an approach that focuses on the self-expressed needs of patients and their families. Timely discussion of wishes and preferences is difficult for all people in all countries, and advance directives are not particularly helpful anywhere.

Many of the contributors report their personal experiences or views. Although data on end-of-life issues are not available in many countries, this lack of information does not warrant some contributors’ rash assumptions about how things are, or should be. Nevertheless, this is a largely successful effort to provide international comparative information about death and dying, an especially important issue that may be indicative of the quality of how other sensitive public issues are dealt with in society.

**Agnes van der Heide**  
a.vanderheide@erasmusmc.nl
Profile

Jörg-Dietrich Hoppe: forthright President of the German Medical Association (Bundesärztekammer)

Regret still gnaws at Jörg-Dietrich Hoppe now and then—usually striking in concert halls. What if more than 40 years ago he had decided to pursue a career in music instead of switching to medicine? As a teenager, he was an accomplished violinist and dreamed of being a great conductor. But partly to ease the financial burden on his family, the oldest child chose the sure thing—he studied medicine, meeting a young nurse who later became his wife. Hoppe, now 64 years old, is probably Germany’s best known medical doctor. He is midway through his second 4-year term as the outspoken President of the German Medical Association (Bundesärztekammer). When asked if he regrets the path he chose, words fail him for the first and only time. When pressed a bit harder, he concedes, “Sometimes, yes, I feel regret. Sometimes at concerts I listen and I think that I would do it differently, give it a different interpretation.”

Despite his regret, Hoppe has not been unhappy as a physician. Far from it. “It’s not a job, not a burden. I’m always happy when I go back from vacation. I like my work a lot.” So much so, that he recently accepted an extension in his contract as chief of pathology at Krankenhaus Düren (Hospital Düren), meaning he will not retire in October when he turns 65. Indeed, Hoppe shows no signs of slowing down: he is president of the Düsseldorf-based North Rhine regional medical association, a member of the council of the World Medical Association (WMA), and helps former communist eastern European nations set up medical associations. He says he has always been politically inclined and contemplated running for the Bundestag in 1976, but this would have meant the end of his medical profession. Instead, he became active in doctors’ organisations. Otmar Kloiber worked closely with Hoppe at the German Medical Association until earlier this year when he became Secretary General of the WMA. He says Hoppe is a unifier, able to convince diverse, sometimes competing, groups of doctors to speak with one voice. “He understands people, relates to people. He’s calm. It is not possible to make him angry or aggressive.” And Kloiber says Hoppe can still play the violin: “I can’t tell the difference between him and a professional player.”

As head of the German Medical Association, Hoppe frequently issues statements on medical issues and posts many press releases on the association’s website (www.baek.de). In late February, Hoppe issued a statement praising a European Commission initiative to ban alcoholic drink sales to young people aged under 18 years, adding that parents need to set good examples. A few weeks later, he praised a handful of states that have banned cigarette smoking in and around public schools, calling for a national school smoking ban. “My children don’t smoke and I am very happy with that”, he says, adding that one daughter is a nurse, the other daughter a science adviser, while his son will complete medical studies this summer. Hoppe has also recently spoken out against living people donating organs and proposals for legislation to allow doctors to assist critically ill patients who want to die. “If doctors are allowed to do this, it would totally destroy doctor-patient trust”, he says. He even issued a statement indirectly criticising the decision by US courts to allow life-giving feeding tubes to be pulled from Terri Schiavo. “In Germany, this would not be allowed to happen”, he says.

But some of Hoppe’s most forceful statements are directed at Chancellor Gerhard Schröder and his Health Minister, Ulla Schmidt, who have passed new laws to cut health-care costs in an attempt to reduce premiums for Germany’s massive public-health insurance system. Hoppe has called for revision of the new laws, complaining that some low-income people with health problems are not visiting doctors. When asked what he thinks of Schröder and Schmidt, Hoppe says he would prefer not to say. When asked again, he said they have “no vision”. A Health Ministry spokesperson asked to comment on Hoppe first chuckled, and then says: “I am not sure we want to do that”. Hoppe would support measures to save money, but not if those measure mean inadequate medical care. “Doctors love people, have affection for people”, he says. “Politicians and lawyers think in the abstract, not about the effect their decisions will have on sick individuals or on the interaction between doctor and patient.” The quality of German medical care is among the best in the world, Hoppe maintains, and asserts “we are fighting to keep it that way”. When asked about medical care in other major nations, he admits: “I don’t believe I would want to be treated in the UK. I think I would come back to Germany.”

As a child Hoppe became acquainted with physicians because his brother had haemophilia, and he was impressed by the good they did. “That was a major influence on me”, he says, noting that his brother contracted HIV from a tainted blood transfusion in the late 1970s and died of AIDS in 1991. Kloiber says there was another influence in Hoppe’s life: “He lost an eye, you know, in the last days of the war. I think this is what makes him see the problems of patients, of people, their sorrows, their needs.” Huh? He lost an eye? “He was just a boy, 4 years old”, Kloiber explains. “Bomb shrapnel. He didn’t tell you that?”

Ned Stafford
ns@europefn.de
Katherine Austin Lathrop

Key member of team that introduced technetium-99m into clinical use as radiotracer. Born in Lawton, OK, USA, on June 16, 1915, she died of complications from dementia, in Las Cruces, NM, USA, on March 10, 2005.

When Katherine Lathrop learned the first atomic bomb had been dropped, she was horrified, her son David recalls. As a member of the Manhattan Project, she studied the effects of radiation on living tissue and knew as well as anyone what the bomb would do to the people of Hiroshima. Although she believed that shortening the war was worth the bomb’s human cost, her son said, she was grateful for being able to make a contribution to human health through her subsequent research in nuclear medicine. “To find peaceful, utilitarian uses for nuclear energy was very, very important to her and certainly to the group that she was part of”, David Lathrop, now the leader of the arrhythmias group at the US National Heart, Lung and Blood Institute, said.

Katherine Austin’s first encounter with science was a textile chemistry course she took as a home economics major. She impressed the professor, who asked her to teach the course the next year. She went on to get a bachelor’s degree in biology in 1936, and 3 years later a bachelor’s in physics and a master’s in chemistry, all from Oklahoma State University, Stillwater, OK, USA. She married Clarence Lathrop and moved with him and their two children to Chicago so he could attend medical school. She quickly saw she would have to work to keep the family afloat. A friend told her the University of Chicago was hiring people with scientific training for a secret project. She applied, and was hired in the Metallurgical Laboratory of the Manhattan Project as a junior biochemist, studying the metabolism, tissue distribution, and excretion of radium and fission products in animals. “In high school . . . I wanted nothing to do with the cats being dissected in Biology”, she recalled in a 1995 interview. “Faced with earning enough money to care for my two children and get my husband through medical school, I quickly decided I could learn to work with animals.”

In 1947, Lathrop became associate biochemist at the Argonne National Laboratory. 7 years later, tired of the long commute and the toll it took on her as a working mother—she now had five children—she joined Paul Harper’s laboratory at the Argonne Cancer Research Hospital at University of Chicago, Chicago, IL, USA, gaining her first title, biochemist in the department of surgery, but taking a 50% pay cut. “She was really Dr Harper’s right hand”, recalls Robert Beck, who worked with the team to develop an imaging system and scanner to follow the pharmacokinetic distribution of radioactive materials in small animals. With her background in physics, chemistry, and biology, as well as her diligence and organisation, Lathrop helped ground and guide the more volatile Harper, Beck said. “I provided the ingenuity and Katherine provided the scholarship”, Harper himself told an interviewer in 1995. She rejoined: “Well, it just worked.” Their partnership lasted more than 40 years.

Initially the team focused on implanting radioactive materials for treating malignant disease, but their efforts gradually shifted toward imaging. Working with Beck and Don Charleston, Harper and Lathrop came to the realisation that technetium-99m, with its 6 h half-life, might be a good agent. The team built their own brain scanner and produced “the first really spectacular scan” in 1963, Beck said, publishing the results in 1964. “That paper really set off a kind of growth spurt in nuclear medicine”, he added. Lathrop developed a series of molecules tagged with technetium 99-m for use in imaging the thyroid, liver, and bone, and with Harper introduced the commercial method for producing iodine-135. She also helped devise a technique for imaging stable iodine in the thyroid gland by using fluorescence that produced a tiny fraction of the radiation exposure of the standard iodine-131 scan.

Lathrop became a full professor at Chicago University, winning emerita status in 1985. Greg Karczmar, now an associate professor of radiology and medical physics, joined the university as an assistant professor at about this time, and remembered Lathrop as “friendly and welcoming” and still excited about her work. She had accumulated a unique database on the effects of radiation on tissue, Karczmar said, and was hoping to compile it and make it available to others. Her last paper, published in 1999, dealt with this material. Lathrop retired in 2000, after a stroke.

“I’m sure she’s probably involved in some grand experiment some place trying to figure out how it all works”, says her son. She is survived by her son, three daughters, a sister, a brother, ten grandchildren, and five great-grandchildren.

Anne Harding
anne_harding@yahoo.com
Increased risk of cardiovascular events with coxibs and NSAIDs

The withdrawal of rofecoxib and David Graham and colleagues’ findings (Feb 5, p 475)1 on the cardiovascular side-effects of selective cyclo-oxygenase-2 inhibitors (coxibs) and non-steroidal anti-inflammatory drugs (NSAIDs) pose major concerns for the management of patients with inflammatory conditions. For many years, NSAIDs have been known to be potentially dangerous, mainly because of their gastrointestinal toxicity. Guidelines from both the American College of Rheumatology2 and European League against Rheumatism3 on the management of osteoarthritis suggest that NSAIDs should not be first-line treatment for these patients. However, in a nested control study from the USA, only a maximum of 8% of patients and 5% of controls seemed to have an inflammatory disorder.4 This finding suggests that coxibs—especially rofecoxib—are very widely and possibly inappropriately prescribed for non-inflammatory disorders.

Whether the cardiovascular toxicity is specific to rofecoxib or is a class effect is unclear,1 and possibly also relevant to other NSAIDs. Unfortunately, most data compare a coxib with a conventional NSAID, usually naproxen. No long-term placebo-controlled trials of older NSAIDs have looked specifically for cardiovascular side-effects, and it is therefore impossible to assess the true cardiovascular risks from these older NSAIDs. The mixing of comparator studies (showing differences between different coxibs and suggesting cardiovascular risks of naproxen greater than celecoxib) and placebo-controlled studies only of coxibs makes interpretation of the data impossible.

Over the past few years, there has been increasing evidence that chronic inflammatory disorders such as rheumatoid arthritis are associated with an increased cardiovascular mortality.5 This increase has been ascribed to chronic inflammation, but most patients with rheumatoid arthritis take NSAIDs for many years. These studies will need to be reassessed.

In the UK, substantial harm and anxiety has been caused to our patients with inflammatory arthritis. Some general practitioners have stopped all coxib prescriptions, either leaving patients with no anti-inflammatory treatment and more pain, or substituting older NSAIDs with potentially increased risk of gastrointestinal toxicity (and maybe even increased cardiovascular toxicity). In the current state of knowledge, a less hysterical approach is called for, especially since the data presented by David Graham and colleagues suggest that celecoxib might actually be safer than some of the older NSAIDs.

We have received support from many pharmaceutical companies, including those making conventional NSAIDs and coxibs.

*David G I Scott, Richard A Watts
david.scott@nnuh.nhs.uk
Department of Rheumatology, Norfolk and Norwich University Hospital NHS Trust, Colney Lane, Norwich NR4 7UY, UK (DGIS); Department of Rheumatology, Ipswich Hospital NHS Trust, Ipswich, UK (RAW)


The current crisis in assessing risk precipitated by the finding of increased cardiovascular events with selective cyclo-oxygenase-2 inhibitors could be mitigated by a cohesive protocol for effectively assessing adverse reactions, which should include the following.

(1) Acknowledgment that safety is relative. Contraceptives substantially increase the risk of thrombosis in healthy individuals, yet the risks are accepted.6
(2) Acknowledgment that design constraints limit assessment of every type of risk before licensure. Rare idiosyncratic events and increases in common events are examples.
(3) Industry maintenance and reporting of surveillance registries and literature for 4 years after approval.
(4) A distinctive mark and the Medwatch number or other adverse drug reaction hotline number on the label for 4 years after approval.1
(5) Clinical and biochemical subgroup analysis by industry and regulatory agencies of patients who sustain adverse events. In the case of contraceptives, we have learned that smoking increases the risk of thrombosis significantly,1 and recent data indicate that individuals with particular coagulation profiles have the greatest risk of thrombosis.4
(6) Institution of statistical methods to detect sentinel events. The Poisson equation is suited to surveillance assessment because it is not denominator-dependent.3 If needed, standard event incidence rates and marketing data could be used to determine a basic estimate of use. Establishing statistical criteria that trigger labelling changes, particularly when relative risk data are available, could ameliorate the current confusion that results in the face of increased incidence of events. For example, 1–2-fold increases in common conditions could result in an addition in the warning section, whereas 2–4-fold increases could result in black-box labelling or banning, depending on drug need. 5-fold increases could lead to restricted use in limited or previously unresponsive populations with written patient consent, or to banning of the drug.

(7) The Agency for Healthcare Research and Quality, or other agencies with experience in best practice analysis, should be tasked with promulgating best practice guidelines for use of drugs with complex safety profiles.
Institution of strategic protocols will avert future crises in safety monitoring and enable seamless and effective drug development.

I declare that I have no conflict of interest.

Rebecca Dachman
beede2@mindspring.com
PO Box 3370, Silver Spring, MD 20918, USA


5 DuMouchel W. Bayesian data mining frequency tables with an application to the FDA spontaneous reporting system. Am Stat 1999; 53: 177.

The paper by David Graham and colleagues' prompts commentators Simon Maxwell and David Webb2 to criticise our earlier contention that British prescribing was conservative.1 In fact both publications reinforce our view. In Graham and colleagues' study, the risk of serious coronary heart disease on rofecoxib and non-selective non-steroidal anti-inflammatory drugs (NSAIDs) was compared with remote NSAID use and with celecoxib. The effect of rofecoxib (all doses) did not reach significance in either analysis, the effect of 25 mg rofecoxib or less per day was significant in one but not the other analysis, and significant results with more than 25 mg rofecoxib per day were based on ten events in patients and eight in controls.

By contrast, there was a clear increase in risk of acute myocardial infarction with both naproxen and other NSAIDs that was consistent in both analyses and similar in magnitude to that seen with therapeutic doses of rofecoxib. However, Graham and colleagues' interpretation of the data differs nicely for rofecoxib ('increases the risk of serious coronary heart disease') and naproxen ('does not protect against serious coronary heart disease').

Moreover, neither Graham and colleagues nor Maxwell and Webb see fit to discuss the implications of the data with all other non-selective NSAIDs at all—a truly extraordinary omission that could not more tellingly illustrate the bias we originally wrote about.1 By ignoring the fact that risk was increased with current NSAID as well as rofecoxib use, Graham and colleagues draw the spurious conclusion that "an estimated 88 000–140 000 excess cases of serious heart disease probably occurred over the market-life of rofecoxib".

Focusing on the effects and hazards of a new (and in this case now withdrawn) drug might support catchy headlines, but dying from the effects of older drugs that are still available could be rather more important to a patient.

CH has received research funding or honoraria from AstraZeneca, GlaxoSmithKline, Merck, Nitromed, Novartis, Pfizer, Takeda, Serono International, and Wyeth.

Paul Emery, Andrew Moore, *Chris Hawkey
cj.hawkey@nottingham.ac.uk
Academic Unit of Musculoskeletal Disease, Leeds University, Leeds, UK (PE); University of Oxford, Wellington Square, Oxford, UK (AM); Institute of Clinical Research, Wolfson Digestive Diseases Centre, University Hospital, Nottingham NG7 2UH, UK (CH).


I have been following with great interest the Lancet articles on rofecoxib, especially David Graham’s report1 and the accompanying Comment.2 Graham's article makes the very valid point that, since the public pays for its regulatory agencies, it has the right to expect those bodies to work in the public interest. As far as my experience with rofecoxib goes, it seems that a certain sector of the regulatory bodies— the ethics committees—are not performing at all well and that their role in the setting-up of clinical trials is perfunctory, to say the least. Please allow me to enlarge on the above.

In 2003, a clinical trial was set up to examine the possible use of rofecoxib in the prevention of prostate cancer. Despite the availability of studies that showed a link between rofecoxib and increased heart attacks or strokes, the ethics committee allowed the trial to go ahead. At the request of my general practitioner, I volunteered to take part in the trial because I felt that it might prove beneficial in combating that scourge of my sex, prostate cancer. I accepted the potential side-effects mentioned in the patient information sheet, although, in retrospect, certain sentences can now be seen as rather evasive. The sheet states: “Isolated cases of heart attack and stroke have been reported in patients treated with rofecoxib. However, it is not known whether these were caused by treatment with rofecoxib.”

The trial ended for me when I was rushed into hospital almost dead. I entered the trial a fit, active man. I left, after 3 weeks, with ulcerative colitis and all the psychological problems concerned with this recurring illness, and with no desire to ever volunteer for any further clinical trials.

It is my view that those involved in setting up this trial should be severely censured. I do not take kindly to being used as a human guinea-pig, especially when the dangers were known by people who really should have acted more responsibly. It seems to me that the profitability of the drug company involved was more important than trialists’ lives.

When you are lying in hospital for 3 weeks with intravenous fluids and steroids being pumped into your body; when you are being fed a diet of mesalazine, prednisolone, and antibiotics; when you are having blood samples taken two or three times a day; when you are being X-rayed frequently;
when you are having to examine and weigh your excreta; when you are squirting Predfoam up your anal passage (need I go on?). It is quite easy to feel justified in taking a rather jaundiced view of the actions of drug companies, regulatory bodies, and ethics committees.

It would have been quite pleasant for someone to have said, “Sorry, but we made an awful mistake”, instead of someone to have said, “Sorry, but we feel justified in taking a rather jaundiced view of the actions of drug companies, regulatory bodies, and ethics committees.”

I declare that I have no conflict of interest.

L G Thomas

23 Chelmer Way, Ely CB6 2WS, UK


Screening for chlamydia

William Miller (Feb 5, p 456) asks whether we are doing enough to screen those at risk of genital chlamydial infection. We believe that the answer is no. Opportunistic screening such as that provided by the English National Chlamydia Screening Programme and screening recommendations in the USA are unlikely to be sufficient to control the spread of chlamydia. Sweden’s experience demonstrates this point.

In Sweden, opportunistic screening and mandatory partner notification for chlamydia have been undertaken since the 1980s. Screening is organised at county level within a range of health-care settings, and 75–80% of those screened are women.1 However, although rates of diagnosed chlamydia, pelvic inflammatory disease, and ectopic pregnancy initially declined, comprehensive surveillance shows that chlamydia rates in Sweden doubled between 1997 and 2003 to pre-screening levels.1

The Uppsala Women’s Cohort Study provides some insight into the reasons underlying the failure of long-term chlamydia control activities in Sweden.2 This study linked a population register including all women of reproductive age who were resident in Uppsala County between 1985 and 1989 with records of chlamydia tests from 1985 to 1996 and hospital admissions from 1985 to 1999. We used survival analysis to estimate chlamydia screening uptake, and examined hospital-diagnosed disorders of the female reproductive tract resulting from complications of chlamydia in women who were screened and had positive results, in those screened with negative results, and in those never screened for chlamydia.

We included 52 580 women aged 15–24 years, with 719 717 woman-years of follow-up until Dec 31, 1999. By the age of 35 years, an estimated 70.7% had been screened at least once, but nearly half were only ever tested once. The number of women screened each year decreased from 1991 onwards. During the study period, 3169 women had pelvic inflammatory disease, ectopic pregnancy, or infertility. Of these women, 1007 (31.8%) had never been screened, 1802 (56.9%) had only negative tests, and 360 (11.4%) had had a positive chlamydia test.

Low annual coverage and infrequent screening for this communicable disease mean that only a minority of women in the target population had the opportunity of benefiting from early treatment and partner notification each year. Most complications therefore occurred in women who had never been screened or had negative tests, and transmission in the population was able to continue.

Lessons about the inherent difficulties of attaining and sustaining adequate coverage through opportunistic screening can also be learnt from the UK. Opportunistic cervical cancer screening in family planning clinics and general practices started in the 1960s, but coverage only increased sufficiently to prevent deaths after a systematic register-based system was introduced.1 If chlamydia screening in England is not to suffer the same fate, empirical data about the limitations of opportunistic screening in both Sweden and the UK should now be used to influence health policy about chlamydia control.

We thank Fowzia Ibrahim for help with data analysis. We declare that we have no conflict of interest.

Nicola Low, Roger M Harbord, Matthias Egger, Jonathan A C Sterne, Bjorn Herrmann

low@ispm.unibe.ch

Department of Social and Preventive Medicine, University of Berne, Finkenhofweg 11, CH-3012, Switzerland (NL, ME); MRC Health Services Research Collaboration, Department of Social Medicine, University of Bristol, Bristol, UK (NL, RMH, ME, JACS); Department of Women’s and Children’s Health, University of Uppsala, Uppsala, Sweden (BH); Department of Clinical Microbiology, University Hospital of Uppsala, Uppsala, Sweden (BH).


4 Low N. Will a women-only chlamydia screening programme reduce the incidence of PID? Presentation at Royal College of Obstetricians and Gynaecologists/British Association for Sexual Health and HIV joint meeting; London, UK; Dec 3, 2004.


The importance of rigorous assessment of screening programmes has been emphasised by the UK National Screening Committee. Unfortunately, these strictures do not seem to have been applied to chlamydia screening, for which an opportunistic model of unproven effectiveness is being introduced in England.1 Free pharmacy-based testing1 would be an extension to this model.

Pharmacy-based testing, however innovative, does not seem promising as...
an approach to substantially increasing screening opportunities among young people. Free pharmacy-based testing was recently assessed in Amsterdam, the Netherlands. Women aged 15–29 years who were collecting contraceptives were actively offered a chlamydia test by a pharmacist. Only 13% of the eligible population were tested over a 2-year period, distribution declined over time, and the intervention was not cost-effective. It seems reasonable to assume that, without active encouragement, uptake would be even lower. Further, pharmacy testing is unlikely to increase screening coverage among men, who probably have less reason to attend pharmacies.

By contrast, postal chlamydia screening of home-collected specimens is an innovative approach that can reduce the incidence of pelvic inflammatory disease. Those testing positive can receive treatment and contact tracing in primary care. Uptake among eligible 16–24-year-olds in England is about 30%, which is greater than pharmacy testing and reaches a substantial proportion of people who do not regularly use primary health-care services.

To reach young men, opportunistic screening in primary care probably offers the greatest potential. We have shown in England that 60% of men and 75% of women) aged 16–24 years attended their practice at least once during a 3-year period. Thus, a combination of opportunistic screening in primary care with home-based screening of infrequent attendees might be the most rational approach.

We are not saying that pharmacy-based testing has no place, rather that there is currently no evidence of its value. Introducing eye-catching policies that have not been properly assessed seems unlikely to benefit public health. The bottom line is that any approach to screening needs to be rigorously assessed and properly resourced. The phased introduction of the English National Chlamydia Screening Programme provides the ideal opportunity for the former, through randomised controlled trials of competing screening models. With regard to the latter, sound science is needed to ensure that money is well spent.

We declare that we have no conflict of interest.

*John Macleod, Chris Salisbury, Nicola Low
j.a.macleod@bham.ac.uk
Department of Primary Care and General Practice, University of Birmingham, Birmingham B15 2TT, UK (JM); Department of Community Based Medicine, University of Bristol, UK (CS); Department of Social and Preventive Medicine, University of Berne, Switzerland (NL.)


HCV-indeterminate blood donors or occult HCV infection?

Nasser Semmo and colleagues (Jan 22, p 327) report on the detection of hepatitis C virus (HCV)-specific T-cell responses in 15 of 30 antibody-indeterminate blood donors, suggesting previous exposure to the virus. Whether such virus-specific responses develop during transient exposure in self-limited acute HCV infection as suggested by Semmo and colleagues, or instead reflect HCV persistence in the liver or within immune privileged sites is still unknown. In fact, Semmo and colleagues propose that T-cell responses could arise in situations with limited antigen exposure or potentially when the virus replicates poorly.

Indeed we have described the presence of occult HCV infection characterised by the detection of HCV RNA in the liver (as well as in peripheral blood mononuclear cells in most cases) in the absence of serum HCV RNA and circulating anti-HCV antibodies (as tested by enzyme immunoassay) in individuals with abnormal serum aminotransferase concentrations of unknown cause. Additionally, HCV persistence has been reported in the liver in sustained virological responders many years after successful treatment, as deduced from the fact that HCV RNA was undetectable in serum but present in mononuclear cells. We have found virus-specific T-cell responses in the peripheral blood in patients with occult HCV infection.

Screening blood donations for anti-HCV antibodies and HCV RNA, and even measuring aminotransferase concentrations, will not exclude HCV infection because these enzyme concentrations can normalise transiently in some patients with occult HCV infection. However, an indeterminate confirmatory antibody test might be recorded in few cases.

We agree with Semmo and colleagues in that future studies of the natural history of HCV should include donors with indeterminate HCV antibody tests, even if HCV persistence rather than transient virus exposure is confirmed in any of them. Although such confirmation needs to be done on liver tissue, a surrogate measure can be carried out in peripheral blood mononuclear cells: HCV RNA is detectable in at least two-thirds of patients with occult HCV infection or in sustained responders to antiviral treatment.

Whatever the case, the report by Semmo and colleagues highlights the importance of introducing more sensitive assays to test for HCV exposure in blood donations. Such analyses could provide important information about the status of HCV with potential implications in the clinical setting. Other pop-
ulations at risk of HCV infection such as haemodialysis patients await this type of screening.

We declare that we have no conflict of interest.

Juan A Quiroga,* Vicente Carreño fehvp@fehv.org

Fundación para el Estudio de las Hepatitis Virales, 28015 Madrid, Spain


Authors' reply

Juan A Quiroga and Vicente Carreño raise the issue of to what extent hepatitis C virus (HCV) is truly cleared from the body, even for individuals in whom virus is no longer detectable by conventional analyses of plasma or serum. This issue applies not only to patients whose recombinant immunoblot assay (RIBA) results are indeterminate, but more commonly to those in whom RIBA results are positive, indicating an acute resolving infection. Additionally, the question has been applied to those who have had persistent infection, which has been apparently successfully treated by interferon-based regimens. The detection of viral RNA, including the negative strand (indicative of replicating virus), has been possible with very sensitive assays of both peripheral blood mononuclear cells and liver tissue in some such patients.

The replicative capacity of these viral forms is currently not understood. However, virally derived antigen might be produced in such settings, which could contribute to the maintenance of immunological memory, and thus the T-cell responses we saw. The issue of the role of antigen in the maintenance of long-term immunological memory after viral infection has been extensively investigated in mouse models, such as lymphocytic choriomeningitis virus. Memory CD8+ T cells seem to persist in situations where antigen is no longer detectable, although if low-level antigen is present it can lead to restimulation of T cells and increases in functional capacity, including protection against rechallenge.

In HCV, recent studies suggest that virus-specific T cells found in blood after infection exist in a quiescent memory state. However, we have seen quite some variation in both CD4+ and CD8+ T-cell responses, and some of this variability could result from the presence of viral antigens in some individuals, in whom it is associated with low-level viral persistence, but not in others. Additionally, the observation that such immune responses can change in immunodominance over time in selected individuals also suggests that antigen presentation could continue in some form long after apparent resolution of infection.

Since we saw maintenance of cellular immune responses in individuals with RIBA-indeterminate antibody status, a proportion of such patients could also possess very low-level viral RNA in particular sites, as could RIBA-positive individuals after spontaneous resolution of acute infection. However, the key point of the study was that T-cell responses were detectable in the RIBA indeterminate group, suggestive of exposure to HCV antigens at some point. Further detailed studies are clearly required to define the origins of this state and its natural history, as well as the potential role of persistence of viral RNA.

We declare that we have no conflict of interest.

Nasser Semmo, Gillian Harcourt,* Paul Kleneman, Craig Taylor, Neil Smith

Nuffield Department of Medicine, Peter Medawar Building for Pathogen Research, South Parks Road, Oxford OX1 3SY, UK (NS, GH, PK), National Blood Service, Birmingham, UK (CT, NS)

Appropriate technology in tuberculosis diagnostics

We would like to make three points in response to Alwyn Mwinga’s Comment (Jan 8, p 97) on our comparison of diagnostic techniques for tuberculosis in HIV patients with inadequate sputum production.

First, the statement that the use of the string test “will be hampered by the need for sputum induction” seems to be a misunderstanding: sputum induction added nothing to the results of the string test and was used only as a comparative test for this research. In fact, our findings clearly show that the string test alone offers better diagnostic sensitivity than sputum induction. By obviating the need for sputum induction, it could remove an important risk factor for nosocomial transmission of tuberculosis, particularly in resource-poor settings with a high tuberculosis burden, no isolation facilities, and wards crowded with highly susceptible HIV-infected patients. We would therefore suggest that the string test should supersede sputum induction if these data are borne out in other settings.

Second, we believe that the string test might indeed have a potential role in the difficult diagnosis of paediatric pulmonary tuberculosis, and we have shown that the string test procedure is well tolerated by children with sus-
Strangulation by intravenous tubes

The risk of accidental childhood asphyxia in various recreational and domestic settings has been repeatedly stressed, and represents an important target for preventive action. In infants, most unintentional deaths associated with asphyxia are sleep related—eg, suffocation by an overlying adult, smothering by bedding or a pillow, or wedging between a mattress and bed frame.1 Strangulation is less common in infants than in toddlers, and is very seldom reported in the hospital setting.

In 2004 we investigated a fatal strangulation by intravenous (iv) tubing during hospital treatment. A 10-month-old girl with a history of acute lymphoblastic leukaemia was admitted to a local hospital for cytostatic therapy and received iv antibiotics and liquids via a central venous catheter for treatment of suspected sepsis. The patient’s temperature decreased over the next few days to about 36°C. 3 days after admission, the girl was restless in the evening but eventually fell asleep after being given an analgesic. The nurse on duty checked the infant early the next morning, at which time she was sleeping; an hour later, the nursing staff found her lying prone in her crib. She had no pulse, and was cyanotic and apnoeic. The iv tubing inserted into her right clavicular vein was tightly wrapped twice around her neck. The tubing was cut immediately, but despite prompt efforts at resuscitation the girl was declared dead shortly after.

The case was referred to the Department of Forensic Medicine, University of Helsinki, where an autopsy was performed. The cause of death was lymphoblastic leukaemia. An exsanguination with iv tubing was noted, and there were no other signs of non-traumatic injury. Strangulation posed by iv devices and monitoring equipment is a well-recognised risk, especially in resource-poor settings.

Reference


Endometrial cancer and hormone-replacement therapy in the Million Women Study

Million Women Study Collaborators

Summary

Background Postmenopausal women who use hormone-replacement therapy (HRT) containing oestrogen alone are at increased risk of endometrial cancer. To minimise this risk, many HRT users who have not had a hysterectomy use combined oestrogen-progestagen preparations or tibolone. Limited information is available on the incidence of endometrial cancer in users of these therapies.

Methods 716 738 postmenopausal women in the UK without previous cancer or previous hysterectomy were recruited into the Million Women Study in 1996–2001, provided information about their use of HRT and other personal details, and were followed up for an average of 3.4 years, during which time 1320 incident endometrial cancers were diagnosed.

Findings 320 953 women (45%) reported at recruitment that they had used HRT, among whom 69 577 (22%) last used continuous combined therapy (progestagen added daily to oestrogen), 145 486 (45%) last used cyclic combined therapy (progestagen added to oestrogen, usually for 10–14 days per month), 28 028 (9%) last used tibolone, and 14 204 (4%) last used oestrogen-only HRT. These HRT types had sharply contrasting effects on the overall risk of endometrial cancer (p<0.0001 for heterogeneity). Compared with never users of HRT, risk was: reduced with last use of continuous combined preparations (relative risk 0.71 [95% CI 0.56–0.90]; p=0.005); increased with last use of tibolone (1.79 [1.43–2.25]; p<0.0001) and oestrogen only (1.45 [1.02–2.06]; p=0.04); and not significantly altered with last use of cyclic combined preparations (1.05 [0.91–1.22]; p=0.5). A woman’s body-mass index significantly affected these associations, such that the adverse effects of tibolone and oestrogen-only HRT were greatest in non-obese women, and the beneficial effects of combined HRT were greatest in obese women.

Interpretation Oestrogens and tibolone increase the risk of endometrial cancer. Progestagens counteract the adverse effect of oestrogens on the endometrium, the effect being greater the more days every month that they are added to oestrogen and the more obese that women are. However, combined oestrogen-progestagen HRT causes a greater increase in breast cancer than the other therapies do. Thus, when endometrial and breast cancers are added together, there is a greater increase in total cancer incidence with use of combined HRT, both continuous and cyclic, than with use of the other therapies.

Introduction

Use of oestrogen-only hormone-replacement therapy (HRT) increases the risk of endometrial cancer. To counteract this effect, many postmenopausal women who have not had a hysterectomy use combined HRT—regimens containing progestagens and oestrogens. Epidemiological evidence suggests that use of such combined oestrogen-progestagen therapy attenuates, and perhaps even reverses, the oestrogen-associated increase in endometrial cancer. However, published findings are sparse and mostly come from the USA, where HRT preparations often differ from those available elsewhere. Furthermore, tibolone, a synthetic steroid with oestrogenic, progestagenic, and androgenic properties, has not been licensed in the USA, but is prescribed for HRT in many countries, and has been linked to an increased risk of endometrial cancer. We report here on the relation between use of different types of HRT and incidence of endometrial cancer in the Million Women Study, a cohort study including about one of every four women in the UK who were aged 50–64 years in 1996–2001.

Methods

Data collection, follow-up, and definitions

Between May, 1996, and March, 2001, more than a million women joined the study. They completed a recruitment questionnaire containing questions about sociodemographic and other personal factors, including use of HRT. A follow-up questionnaire was sent to study participants 2–3 years after recruitment, to update information on use of HRT and other factors. These questionnaires can be viewed on the study website. To date, 1.3 million study participants have been flagged on the NHS Central Registers, so that cancer registrations and deaths can be routinely notified to the investigators. The registers provide information on the date of each such event and code the cancer site using the 10th revision of the International Classification of Diseases (ICD). The main endpoint for this report is endometrial cancer (ICD C54) first diagnosed after recruitment into the study; some analyses use incident invasive breast cancer (ICD C50) as the endpoint. Women with any type of cancer registered before recruitment, except non-melanoma skin cancer (ICD C44), are excluded.
Women were classified according to their reported use of HRT and other relevant factors at recruitment. Information recorded about use of HRT included: ever use; age at first and last use; total duration of use; and the name of the proprietary preparation last used and the duration of its use. Information was not gathered on the name of proprietary preparations used before the last one. The specific constituents and formulation of each proprietary preparation of HRT were obtained from the British National Formulary.29 Women were classified as having last used: continuous combined preparations (progestagen added to oestrogen on a daily basis); cyclic combined preparations (progestagen added to oestrogen on some, but not every, day of the month); Tibolone; oestrogen-only HRT; other or more than one HRT type simultaneously; or an unspecified or unknown HRT type. They were defined as exclusive users of Tibolone if the reported duration of use of Tibolone was the same as the reported total duration of use of HRT.

Statistical analysis
Person-years were calculated from the date of recruitment up to the date of cancer registration, death, or last date of follow-up, whichever came first. The last date of follow-up was Dec 31, 2002, except for areas covered by the South West, Midlands, and Trent cancer registries, where the date was Dec 31, 2001, and for Scotland, where the date was Dec 31, 1999. Women who had had a hysterectomy before recruitment were excluded from analyses of endometrial cancer. (Follow-up of a sample of such women showed that 1.4% had a hysterectomy for reasons other than cancer in the 2–8 years after recruitment; however right-censoring could not be done since information on subsequent hysterectomy was not routinely available for all women included in the analyses; table 1). Women who joined the study after the date of last follow-up or who had not been flagged on the NHS Central Registers could not be included in analyses. More women were potentially eligible for these analyses than had been previously21 because of the extended period of follow-up and the additional flagging of study participants in the interim.

Cox-regression models were applied to estimate relative risks and their corresponding 95% CIs, using the computing package STATA version 8.1 (Stata, TX, USA). Data were stratified by age at entry (<50, 50–54, 55–59, 60–62, 63–65, >65 years) and adjusted by: region (ten areas covered by ten cancer registries); quintiles of socioeconomic group, based on deprivation index;22 time since menopause (<5, 5–9, 10–14, 15–19, >20 years); body-mass index (<18.5, 18.5–24.9, ≥25 kg/m²); parity (<2, ≥2); alcohol consumption (g/week); current smokers (yes, no); and use of oral contraceptives (yes, no).

Table 1: Characteristics of women with no previous cancer or hysterectomy, their patterns of use of HRT, and details of follow-up, according to use of HRT reported at recruitment

<table>
<thead>
<tr>
<th>Characteristics at recruitment</th>
<th>Continuous combined</th>
<th>Cyclic combined</th>
<th>Tibolone</th>
<th>Oestrogen only</th>
<th>Other or unknown</th>
<th>Never use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of women</td>
<td>69 577</td>
<td>145 866</td>
<td>28 028</td>
<td>14 204</td>
<td>63 658</td>
<td>395 785</td>
</tr>
<tr>
<td>Age (mean [SD], years)</td>
<td>57.0 (3.6)</td>
<td>56.2 (3.7)</td>
<td>58.0 (3.8)</td>
<td>57.1 (4.3)</td>
<td>57.4 (4.0)</td>
<td>58.0 (4.3)</td>
</tr>
<tr>
<td>Socioeconomic status (%) [n]</td>
<td>33% (22 872)</td>
<td>34% (49 541)</td>
<td>34% (9517)</td>
<td>34% (4798)</td>
<td>31% (39 782)</td>
<td>31% (122 220)</td>
</tr>
<tr>
<td>Parity (mean [SD])</td>
<td>2.1 (1.2)</td>
<td>2.1 (1.2)</td>
<td>2.1 (1.2)</td>
<td>2.1 (1.2)</td>
<td>2.2 (1.2)</td>
<td>2.1 (1.3)</td>
</tr>
<tr>
<td>Past use of oral contraceptives (%) [n]</td>
<td>64% (44 472)</td>
<td>64% (92 998)</td>
<td>62% (17 990)</td>
<td>62% (85 065)</td>
<td>64% (40 443)</td>
<td>47% (382 800)</td>
</tr>
<tr>
<td>Body-mass index (mean [SD], kg/m²)</td>
<td>25.5 (4.2)</td>
<td>25.5 (4.3)</td>
<td>26.0 (4.3)</td>
<td>25.6 (4.3)</td>
<td>26.4 (4.6)</td>
<td>26.3 (4.8)</td>
</tr>
<tr>
<td>Strenuous physical activity (%) [n] more than once a week</td>
<td>40% (26 698)</td>
<td>41% (57 998)</td>
<td>40% (10 196)</td>
<td>42% (58 14)</td>
<td>39% (23 598)</td>
<td>38% (142 222)</td>
</tr>
<tr>
<td>Alcohol consumption (mean [SD], g/week)</td>
<td>38.1 (3.3)</td>
<td>39.4 (4.4)</td>
<td>40.0 (4.5)</td>
<td>39.0 (4.4)</td>
<td>37.5 (4.5)</td>
<td>30.3 (4.0)</td>
</tr>
<tr>
<td>Current smoker (%) [n]</td>
<td>19% (12 811)</td>
<td>21% (29 591)</td>
<td>19% (5039)</td>
<td>19% (2539)</td>
<td>26% (23 508)</td>
<td>23% (74 109)</td>
</tr>
<tr>
<td>History of hypertension (%) [n]</td>
<td>25% (17 082)</td>
<td>25% (34 647)</td>
<td>27% (7348)</td>
<td>24% (34 787)</td>
<td>25% (15 856)</td>
<td>25% (99 603)</td>
</tr>
</tbody>
</table>

Women with missing values are not included in the percentages. "Based on information reported by 111 301 women with no previous cancer or hysterectomy who were recruited in 1996–98, who also completed a follow-up questionnaire in 1999–2001 (an average of 2.8 years after recruitment), and who did not have endometrial cancer diagnosed during follow-up, did not have a hysterectomy during follow-up, or both. * Questionnaire data from follow-up of the remainder of the cohort are not yet available for analysis. The proportion having a hysterectomy during the subsequent 2.8 years, for reasons other than cancer, was 1.4% overall and 1.8%, 1.5%, 1.4%, 2.0%, 1.6%, and 1.0%, respectively, for women who last used continuous combined, cyclic combined, Tibolone, oestrogen-only HRT, other or unknown types of HRT, and never users. Among women who specified the type of HRT last used, both at entry and 2.8 years later.
Standardised incidence rates for endometrial and breast cancer and corresponding 95% CIs were calculated per 1000 women over a 5-year period. Never users of HRT were taken as the standard, and incidence rates in users of HRT were standardised by age, region of residence, socioeconomic status, time since menopause, parity, use of oral contraceptives, body-mass index, and alcohol consumption.

Role of the funding source
The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The authors (V Beral, D Bull, G Reeves) had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
716738 postmenopausal women with no previous cancer or hysterectomy were recruited in 1996–2001, flagged on the NHS Central Registers, and followed up for incident cancer. Average age at entry into the study was 57.5 years. 320953 women (45%) reported that they had ever used HRT, and among them 22% last used continuous combined preparations, 45% last used cyclic combined preparations, 9% last used tibolone, 4% last used oestrogen-only HRT, 2% last used other or unspecified types are shown in table 1. When never users of HRT were compared, most women who last used the four main types of HRT (continuous combined, cyclic combined, tibolone, and oestrogen only), and of those who reported last using other or unspecified types are shown in table 1. When never users and ever users of HRT were compared, most sociodemographic, health and lifestyle characteristics did not differ substantially. The main difference was that never users were less likely than ever users to have used oral contraceptives (47% vs 64%) and tended to drink less alcohol (average consumption 30.1 vs 39.0 g/week). Among the users of specific types of HRT, there was little material variation in the characteristics of women according to the type last used (table 1).

Information on use of HRT after recruitment was available for 111201 women who completed a follow-up questionnaire an average of 2.8 years after recruitment (table 1). At follow-up, 95% of those who reported at entry that they had ever used HRT continued to report that they were ever users, and, as expected, some ceased use during that period, so that fewer were current users 2.8 years after entry. Also, some women who were initially never users of HRT began use subsequently, with 3% of the never users at entry reporting 2.8 years later that they were current users.

Most women who last used the four main types of HRT were current users when they were recruited and, except for oestrogen-only HRT, most reported last using the same type of HRT, both at entry and follow-up (table 1). Agreement over time was especially strong for

![Figure 1: Relative risk* of endometrial cancer according to type of HRT last used, reported at recruitment](image)

*In users of HRT compared with never users, stratified by age and adjusted by time since menopause, parity, oral contraceptive use, body-mass index, alcohol consumption, region of residence, and socioeconomic status. At recruitment.
The relative risk was 1.12 (0.58–2.17), based on nine cases. To look for possible confounding by other factors, results were further adjusted in turn by oral contraceptive use, body-mass index, and alcohol consumption. To test for possible confounding by physical activity, smoking, history of hypertension, and history of diabetes, but none of these altered the relative risks by more than a factor of 1.02.

Figure 1 shows the relative risk of endometrial cancer in ever users compared with never users of HRT at recruitment, according to the last type of HRT used. Compared with never users, risk of endometrial cancer was lower in women who reported last using continuous combined HRT (relative risk 0.71 [95% CI 0.56–0.90], p=0.005), not altered in those who reported last using cyclic combined HRT (1.05 [0.91–1.22], p=0.5), and higher in women who reported last using tibolone (1.79 [1.43–2.25], p<0.0001) and oestrogen-only HRT (1.45 [1.02–2.06], p=0.04). In the remaining women, the relative risk was 1.67 (0.98–2.84; p=0.06) for last use of other or mixed types of HRT (based on 14 cancers) and 1.12 (0.92–1.37; p=0.03) for women who did not specify the last type of HRT used (109 cancers). Relative risks differed between users of continuous and cyclic combined preparations (χ²=7.6; p=0.006) and between users of continuous combined, cyclic combined, tibolone, and oestrogen-only preparations (χ²=31.2; p<0.0001).

The analyses in figure 1 are adjusted for age, region, socioeconomic status, time since menopause, parity, oral contraceptive use, body-mass index, alcohol consumption, residence of region, and socioeconomic status. Includes users of tridestra (oestradiol for 70 days followed by medroxyprogesterone acetate for 14 days; Orion, Newbury, UK), among whom the relative risk was 1.12 (0.58–2.17), based on nine cases.

Among the 63 658 women who last used other or unknown HRT types, 58 796 did not specify the last HRT preparation used, and 4862 reported last using other or more than one HRT type. Their pattern of use of HRT differed from that of the users of the four main types: they were less likely to be current users at entry (19% vs 70%) and their average duration of use of any type of HRT was shorter (2.1–5–1 years; table 1).

Women without prior cancer or hysterectomy were followed up for cancer incidence over 2.44 million person-years, an average of 3.4 years per woman (table 1). During this follow-up period, 1320 incident invasive endometrial cancers were notified by the NHS Central Registers, diagnosed an average of 2.1 years after recruitment, with the median year of diagnosis being 2000.

Figure 2 shows various results for endometrial cancer in women who reported at recruitment that they had last used continuous combined HRT. 84% were current users at recruitment and no significant differences were found in the relative risks of endometrial cancer between current and past users, between women with short and long durations of use of HRT, or between users of preparations containing different progestagenic constituents. The statistical power of these comparisons is limited, and because past users had ceased use only an average of 1.2 years previously, this study cannot provide reliable information about the effect on endometrial cancer of past use of continuous

<table>
<thead>
<tr>
<th>Characteristic of women at recruitment</th>
<th>Cases/ population (1000s)</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current smoker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>26/29.5</td>
<td>1.05 (0.88–1.26)</td>
</tr>
<tr>
<td>No</td>
<td>196/70.5</td>
<td>1.12 (0.92–1.37)</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25 g/week</td>
<td>106/49.5</td>
<td>1.12 (0.92–1.37)</td>
</tr>
<tr>
<td>≥30 g/week</td>
<td>88/31.3</td>
<td>1.05 (0.88–1.26)</td>
</tr>
<tr>
<td>Body-mass index (kg/m²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>154/61.4</td>
<td>1.12 (0.92–1.37)</td>
</tr>
<tr>
<td>≥30</td>
<td>88/31.3</td>
<td>1.05 (0.88–1.26)</td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strenuous exercise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>135/51.3</td>
<td>1.11 (0.92–1.35)</td>
</tr>
<tr>
<td>No</td>
<td>103/39.0</td>
<td>1.04 (0.74–1.42)</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>26/29.5</td>
<td>1.05 (0.88–1.26)</td>
</tr>
<tr>
<td>No</td>
<td>196/70.5</td>
<td>1.12 (0.92–1.37)</td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strenuous exercise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>135/51.3</td>
<td>1.11 (0.92–1.35)</td>
</tr>
<tr>
<td>No</td>
<td>103/39.0</td>
<td>1.04 (0.74–1.42)</td>
</tr>
</tbody>
</table>
combined therapy. The consistency of the overall result for continuous combined HRT was examined, according to eight personal characteristics of the women (figure 2). The only significant heterogeneity was according to body-mass index, with the relative risk decreasing with increasing obesity ($\chi^2 = 11.8; p = 0.003$), reflecting the fact that obesity increases the incidence of endometrial cancer more strongly in never users than in HRT users (see below).

Figure 3 shows various results for endometrial cancer in women who reported at recruitment that they had last used cyclic combined HRT. Progestagens were added to oestrogen therapy, usually for 10–14 days per month, with 78% using preparations that added progestagens for 12 days per month. 65% were current users at recruitment, and among past users the average time since last use was 3.2 years. No significant differences in relative risks were found between current and past users, between those with short and long total durations of use, between women with different progestagenic constituents in the preparation last used, or between those using progestagens cyclically for different numbers of days. The consistency of the overall results for cyclic HRT was examined, according to various characteristics of the women (figure 3). There was some heterogeneity according to women’s age at recruitment ($\chi^2 = 9.2; p = 0.003$) and socioeconomic status ($\chi^2 = 5.9; p = 0.02$), and strong heterogeneity according to body-mass index ($\chi^2 = 16.7; p = 0.0002$), again reflecting the fact that obesity increases the incidence of endometrial cancer more strongly in never users than in ever users of HRT (see below).

Figure 4 shows various results for endometrial cancer in women who reported at recruitment that they had last used tibolone. 73% were current users at recruitment, past users had ceased use an average of 2.7 years previously, and 48% were classified as exclusive users of tibolone. No significant differences were found in the relative risk of endometrial cancer between current and past users or according to whether or not use of tibolone was exclusive. The average duration of use of tibolone reported at recruitment was 3.1 years, and the relative risk for endometrial cancer was greater among women who reported having used tibolone for more than 3 years compared with shorter durations ($\chi^2 = 5.4; p = 0.02$). When analyses were restricted to women who were likely to be exclusive users of tibolone, the relative risk of endometrial cancer remained significantly higher in those reporting use of tibolone for more than 3 years compared with shorter durations (33 cancers, 2.57 [95% CI 1.81–3.66] vs ten cancers, 0.96 [0.51–1.79]; $\chi^2 = 6.6; p = 0.01$). The consistency of the overall findings was examined, according to eight personal characteristics of the women (figure 4). Again, heterogeneity was noted according to body-mass index ($\chi^2 = 11.4; p = 0.003$).

Among the small group of 14,204 women who reported at recruitment that they last used oestrogen-only HRT, the relative risk of endometrial cancer in current users was 1.80 (95% CI 1.19–2.70), based on 24 cancers, and the corresponding figure for past users was 0.97 (0.50–1.97), based on nine cancers. The relative risks of endometrial cancer associated with last use of oestrogen-only HRT again tended to be higher in non-obese than in obese women: 1.69 (0.92–3.11) for those with a body-mass index of less than 25 kg/m², 2.10 (1.24–3.54) for those with an index of 25–29 kg/m², and 0.55 (0.21–1.49) for women with a body-mass index of 30 kg/m² or greater ($\chi^2 = 5.5; p = 0.02$). When analyses were restricted to women who last used tibolone compared with never users of HRT, stratified by age and adjusted, where appropriate, by time since menopause, parity, oral contraceptive use, body-mass index, alcohol consumption, region of residence, and socioeconomic status.

<table>
<thead>
<tr>
<th>Characteristics of women at recruitment</th>
<th>Cases/ population (1000s)</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All tibolone</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current user</td>
<td>38/17.3</td>
<td>1.71 (1.27–2.32)</td>
</tr>
<tr>
<td>Past user</td>
<td>38/10.7</td>
<td>1.81 (1.29–2.53)</td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>38/9.5</td>
<td>2.22 (1.57–3.14)</td>
</tr>
<tr>
<td>Low</td>
<td>43/13.0</td>
<td>1.71 (1.25–2.35)</td>
</tr>
<tr>
<td>Number of children</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>29/7.1</td>
<td>2.15 (1.45–3.17)</td>
</tr>
<tr>
<td>&gt;2</td>
<td>57/20.9</td>
<td>1.66 (1.26–2.19)</td>
</tr>
<tr>
<td>Past use of oral contraceptives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>47/10.7</td>
<td>1.80 (1.34–2.43)</td>
</tr>
<tr>
<td>Yes</td>
<td>38/17.1</td>
<td>1.72 (1.22–2.44)</td>
</tr>
<tr>
<td>Body-mass index (kg/m²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>36/12.6</td>
<td>2.99 (2.08–4.30)</td>
</tr>
<tr>
<td>25–29</td>
<td>29/10.1</td>
<td>1.77 (1.20–2.62)</td>
</tr>
<tr>
<td>&gt;30</td>
<td>18/4.2</td>
<td>1.19 (0.67–2.75)</td>
</tr>
<tr>
<td>Strenuous exercise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None/once a week</td>
<td>48/16.2</td>
<td>1.62 (1.20–2.19)</td>
</tr>
<tr>
<td>More than once a week</td>
<td>34/11.0</td>
<td>2.01 (1.40–2.90)</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10 g/week</td>
<td>69/21.3</td>
<td>1.77 (1.39–2.18)</td>
</tr>
<tr>
<td>&gt;10 g/week</td>
<td>17/6.8</td>
<td>1.81 (1.08–3.05)</td>
</tr>
<tr>
<td>Current smoker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>73/21.5</td>
<td>1.77 (2.39–2.65)</td>
</tr>
<tr>
<td>Yes</td>
<td>7/5.0</td>
<td>1.59 (0.73–3.49)</td>
</tr>
</tbody>
</table>

Figure 4: Relative risk* of endometrial cancer in women who last used tibolone, by reported pattern of use of tibolone at recruitment and by various characteristics of the women studied

Dotted line represents the relative risk for all women who last used tibolone. Totals are not always the same because of missing values. *For women who last used tibolone compared with never users of HRT, stratified by age and adjusted, where appropriate, by time since menopause, parity, oral contraceptive use, body-mass index, alcohol consumption, region of residence, and socioeconomic status.

Figure 5 shows standardised incidence rates for endometrial cancer per 1000 women over a 5-year period, subdivided both by type of HRT last used and by body-mass index. Among never users of HRT, incidence rates
Among HRT users, incidence rates at every level of body-mass index were lowest in women who last used continuous combined HRT and highest in women who last used tibolone (figure 5). However, when HRT users were compared with never users, findings varied depending on a woman’s body-mass index. Among HRT users who were not overweight (body-mass index <25 kg/m²), use of cyclic combined HRT or tibolone significantly increased the incidence of endometrial cancer compared with never users, and continuous combined HRT conferred no benefit. By contrast, among obese women (body-mass index ≥30 kg/m²), users of both continuous and cyclic HRT had a lower incidence of endometrial cancer than never users, and tibolone users had similar rates to never users.

Breast cancer incidence varies according to the type of HRT used, but unlike endometrial cancer, combined HRT preparations (both continuous and cyclic) cause a greater increase in incidence than does use of tibolone or oestrogen-only therapy. Breast cancer is much more common than endometrial cancer, and when the two cancers were considered together, the total incidence of (endometrial plus breast) cancer was dominated by breast cancer (figure 6). The total incidence of endometrial plus breast cancer was significantly greater in current users of the four specific types of HRT studied than in never users. Furthermore, when the totals were compared across the different types of HRT, incidence rates were higher in current users of combined preparations (both continuous and cyclic) than in current users of tibolone or oestrogen-only therapy. Both for endometrial and breast cancer, the HRT-associated absolute increase in incidence was greatest among women who were not overweight and least among obese women (table 2). There is insufficient information about incidence rates of endometrial cancer in past users of the commonly used HRT types to permit estimation of the effect on endometrial plus breast cancer many years after use ceases.

**Discussion**

In this large cohort study, the types of HRT commonly used in the UK were shown to have sharply contrasting effects on the risk of developing endometrial cancer. Compared with never users of HRT, the overall incidence of endometrial cancer was increased in users of tibolone and oestrogen-only therapy and decreased in users of continuous combined HRT. However, the effect of the commonly used types of HRT varied, depending on a woman’s body-mass index. Among women who were not overweight, use of tibolone, oestrogen-only HRT, and cyclic combined HRT significantly increased the incidence of endometrial cancer whereas use of continuous combined therapy conferred no benefit. By contrast, among obese women (who normally have substantially higher incidence of endometrial cancer than non-obese women), use of both continuous and
cyclic combined HRT significantly reduced the incidence of endometrial cancer, whereas use of dibolone and oestrogen-only HRT had little additional effect on incidence.

The increasing incidence of endometrial cancer with increasing obesity in postmenopausal never users of HRT (figure 5) is well known and believed to result from endometrial proliferation caused by oestradiol and related endogenous hormones.24–26 Oestradiol is produced by adipose tissue, circulating levels of oestradiol rise with increasing body-mass index in postmenopausal women, and the risk of endometrial cancer increases with increasing amounts of circulating oestradiol.24–27 Use of oestrogen-only HRT also increases the incidence of endometrial cancer. Incidence of endometrial cancer similar to, or slightly higher than, that seen in women who have not had a hysterectomy.

The prevalence of obesity and other characteristics of the study populations. †Calculated from published data. ‡For HRT users of each type of HRT. The definition of use and non-use of HRT varies from study to study, as does the background incidence of endometrial cancer in obese women suggests that the levels of circulating endogenous hormones may already be so high that exogenous oestrogens can have little additional effect.

Progestagens counteract the proliferative effects of oestrogens on the endometrium.24 Others have shown that the more days every month that a progestagen is used in a cyclic HRT regimen the lower is the incidence of endometrial cancer.9 Published evidence, including data from this study, indicates that use of combined HRT preparations that administer progestagens cyclically, or continuously, may slightly lower the incidence of endometrial cancer compared with never use of HRT and that cyclic addition of progestagens for 10 or more days per month results in an incidence of endometrial cancer similar to, or slightly higher than, that seen in menopausal women who used HRT.

Table 2: Standardised incidence rates for endometrial and breast cancer per 1000 women over a 5-year period in current users and never users of HRT at recruitment

<table>
<thead>
<tr>
<th>Type of HRT currently used at recruitment</th>
<th>Endometrium</th>
<th>Breast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclic combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tibolone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oestrogen only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never use of HRT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standardised incidence rates (95% CI) calculated for Million Women Study participants, taking never users of HRT as the standard and standardising by age, region of residence, socioeconomic status, time since menopause, parity, use of oral contraceptives, alcohol consumption, and where appropriate, body-mass index. *In women who have not had a hysterectomy.

Figure 7: Summary of published results on the relation between use of combined HRT and endometrial cancer

Dotted line represents the overall relative risk, for all studies combined. Relative risk for users compared with non-users of each type of HRT. The definition of use and non-use of HRT varies from study to study, as does the prevalence of obesity and other characteristics of the study populations. Calculated from published data. For HRT preparations where progestagens are added to oestrogens cyclically for about 10 or more days per month.
never users (figure 7). However, incidence rates depend on a woman’s body-mass index, and as far as we know, no previous study has presented results for combined HRT subdivided by this variable. Among women who were not overweight, use of continuous combined HRT conferred no benefit and cyclic administration of progestagens for about 12 days per month significantly increased the incidence of endometrial cancer, compared with rates in never users of HRT (figure 5).

By contrast, among obese women, use of both continuous and cyclic combined HRT (with progestagens administered for about 12 days each month) significantly reduced the incidence of endometrial cancer, the decrease being greatest with daily use of progestagens. Thus, the progestagens in combined HRT seem to counteract the carcinogenic effects of both endogenous and exogenous oestrogens on the endometrium. In non-obese women, the progestagens appear to act mainly by countering the effects of the exogenous oestrogens in the HRT preparations themselves, whereas in obese women they appear to act largely therapeutically, countering the effects of endogenously produced oestradiol.

No significant differences were found here with respect to the specific progestagen constituent of the combined HRT preparations used or the total duration of use of HRT. However, the ability to detect true differences is hampered by the fairly small number of incident cancers in every subgroup and also by some women’s use of other HRT preparations before and after recruitment.

Women who reported last using tibolone were found to be at an increased risk of endometrial cancer overall, the risk increasing with increasing duration of its use. These results accord with findings from another UK study. The associations persisted when analyses were restricted to women who were likely to have used tibolone exclusively. Adherence to use of tibolone was high during the course of this study, with 84% still reporting 2·8 years after entry that the last preparation used was tibolone. Users of this therapy had similar characteristics at recruitment to users of combined HRT; hysterectomy rates for conditions other than cancer were similar during follow-up (1·4% vs 1·8%); and no other explanation was apparent for the higher incidence of endometrial cancer in tibolone users compared with users of combined HRT.

Randomised trials are generally free from bias and confounding, but they have not been designed to compare the effect of different types of HRT and are too small to provide reliable information about effects in women of different body-mass indices on endometrial cancer. To minimise bias in this study, information on the use of HRT and other factors was collected before the diagnosis of cancer, and all participants were flagged on the NHS Central Registers so that details of any incident cancer were coded before they were notified to the study investigators. Women were excluded if they had cancer diagnosed previously since this might affect both use of HRT and the subsequent incidence of endometrial cancer. Hysterectomy before recruitment was also a reason for exclusion in analyses relating to endometrial cancer since this should eliminate a woman’s risk of developing the cancer. Since hysterectomy after recruitment was rare, but more common in HRT ever users than never users, the relative risk estimates quoted here for endometrial cancer in users compared with never users of HRT at recruitment are slightly underestimated.

Women were classified here according to self-reported information on use of HRT, recorded at entry into the study. Self-reported information at recruitment showed 97% agreement with prescription records for the type of HRT currently used and 95% agreement with self-reported information on ever use recorded 2·8 years after entry; use of the specific types of HRT studied here, except oestrogen-only therapy, did not change appreciably during follow-up (table 1). No information was gathered at recruitment on the name of HRT preparations used before the last one. A small proportion of women who reported at recruitment that they were never users of HRT reported 2·8 years later that they were current users. Changes in use of HRT over time would tend to dilute slightly the estimates of relative risk of endometrial cancer associated with use of the four main types of HRT studied and to dilute any comparisons between women with different patterns of use of HRT.

The characteristics of users of the four commonly used types of HRT were broadly similar. Furthermore, all analyses were routinely adjusted by eight potential confounding factors, and additional adjustment by other factors did not alter the findings. Thus, the qualitative differences reported here in the effect of different HRT types on endometrial cancer are unlikely to be due to confounding. The few significant interactions—other than by body-mass index—observed might be due to chance in view of the many comparisons made.

The overall incidence rates of endometrial cancer and breast cancer in this population (3 and 16 per 1000 over 5 years, respectively) are broadly comparable with rates among women of similar age at a similar time in England (3 and 15 per 1000 over 5 years, respectively, after allowing for the prevalence of hysterectomy in the general population). Because breast cancer is more common than endometrial cancer, the total incidence of endometrial plus breast cancer is dominated by breast cancer. Furthermore, since current use of combined HRT causes a greater increase in breast cancer than the other HRT preparations do, the total incidence of endometrial and breast cancer is greater with current use of combined HRT (both continuous and cyclic) than with
use of tibolone or oestrogen-only therapy. This general conclusion applies to all women with a uterus, both obese and non-obese, but the absolute increase in total cancer incidence is greater in non-obese women. There are insufficient data to provide reliable estimates of the incidence of endometrial cancer many years after use of each type of HRT ceases.

Million Women Study Collaborators
Analysis and Writing Committee—Valerie Beral, Diana Bull, Gillian Reeves.
Sterling Committee—Joan Austoker, Emily Banks, Valerie Beral, Ruth English, Julietta Patnick, Richard Peto, Gillian Reeves, Martin Vessey, Matthew Wallis.


Million Women Study Coordinating Centre—Simon Abbott, Keith Baker, Emma Bailey, Emily Banks, Angela Balkwill, Valerie Beral, Amy Berrington de Gonzales, Judith Black, Anna Brown, Diana Bull, Barbara Crossley, Dave Ewart, Laura Gerrard, Adrian Goodhill, Jane Green, Elizabeth Hilton, Ann Hogg, Jo Hooley, Carol Keene, Nicky Langston, Bette Liu, Cecilia Magnusson, Gillian Reeves, Andrew Roddam, Emma Sherman, Moya Simmonds, Elizabeth Spencer, Alison Timajder.

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

Acknowledgments
We thank the women who participated in the Million Women Study. This research was funded by Cancer Research UK, the NHS Breast Screening Programme, and the Medical Research Council.

References
Effects of N-acetylcysteine on outcomes in chronic obstructive pulmonary disease (Bronchitis Randomized on NAC Cost-Utility Study, BRONCUS): a randomised placebo-controlled trial

Marc Decramer, Maureen Rutten-van Molkem, P N Richard Dekhuijzen, Thierry Troosters, Cees van Herwaarden, Riccardo Pellegrino, CP Onno van Schayck, Dario Olivieri, Mario Del Donno, Wilfried De Backer, Ida Lankhorst, Alfredo Ardia

Background Increased oxidative stress is important in the pathogenesis of chronic obstructive pulmonary disease (COPD). We postulated that treatment with the antioxidant N-acetylcysteine would reduce the rate of lung-function decline, reduce yearly exacerbation rate, and improve outcomes.

Methods In a randomised placebo-controlled study in 50 centres, 523 patients with COPD were randomly assigned to 600 mg daily N-acetylcysteine or placebo. Patients were followed for 3 years. Primary outcomes were yearly reduction in forced expiratory volume in 1 s (FEV₁) and the number of exacerbations per year. Analysis was by intention to treat.

Findings The yearly rate of decline in FEV₁ did not differ between patients assigned N-acetylcysteine and those assigned placebo (54 mL [SE 6] vs 47 mL [6]; difference in slope between groups 8 mL [9]; 95% CI –25 to 10). The number of exacerbations per year did not differ between groups (1·25 [SD 1·35] vs 1·29 [SD 1·46]; hazard ratio 0·99 [95% CI 0·89–1·10, p=0·85]). Subgroup analysis suggested that the exacerbation rate might be reduced with N-acetylcysteine in patients not treated with inhaled corticosteroids and secondary analysis was suggestive of an effect on hyperinflation.

Interpretation N-acetylcysteine is ineffective at prevention of deterioration in lung function and prevention of exacerbations in patients with COPD.

Introduction Oxidative stress has major importance in the pathogenesis of chronic obstructive pulmonary disease (COPD), as shown by: enhanced exhalation of hydrogen peroxide, nitric oxide, ethane, carbon monoxide, and isoprostane; increased concentrations of 4-hydroxyl-2-nonenal in the lungs; or by increased urinary excretion of isoprostane. Oxidative stress leads to inactivation of antiproteases, activation of proteases, and expression of interleukin 8 and tumour necrosis factor α through expression of nuclear factor κB, further enhancing recruitment of neutrophils and formation of isoprostane from the oxidation of arachidonic acid.

Antioxidants such as N-acetylcysteine could well reduce oxidative stress in patients with COPD. However, the effects of antioxidants on outcomes in COPD have not been studied in detail. Results from meta-analyses of between six and 23 studies have shown that N-acetylcysteine reduces exacerbation by 22–29%. However, most of these studies assessed chronic bronchitis and not COPD. A retrospective study of the risk of readmission to hospital for 1291 patients with COPD associated N-acetylcysteine with a dose-dependent decrease in readmission rate. The effects of N-acetylcysteine on outcomes and progression of the disease have not been studied in a large-scale prospective trial lasting long enough to assess potential effects on disease progression.

The Bronchitis Randomized on NAC study (BRONCUS) was designed as a randomised controlled trial of the effects of N-acetylcysteine on the progression of disease and exacerbation rate in patients with COPD who had frequent exacerbations (ie, at least two per year for 2 years). Specifically, the study was designed to assess whether N-acetylcysteine reduced the rate of decline in lung function, reduced exacerbation rate, and improved health status.

Methods Patients The design of the trial has been described in detail before. Briefly, the study was designed as a phase III, double-blind, randomised placebo-controlled parallel-group trial of 600 mg per day oral N-acetylcysteine versus placebo. This dose was chosen on the basis of efficacy in previous studies, which were later reviewed in meta-analyses. Patients were followed for 3 years and analysis was by intention-to-treat. Primary endpoints of the study were yearly reduction in lung function (measured by decline in forced expiratory volume in 1 s, FEV₁) and exacerbation rate; secondary endpoints were quality of life and cost-utility. Here, we discuss the primary endpoints and quality of life. Cost-utility will be reported in a separate publication. Patients were stratified according to use of inhaled corticosteroids at entry.
Patients were first seen at a screening visit and afterwards at 1 month, 3 months, and every 3 months until the end of the trial. At screening the following were recorded: inclusion and exclusion criteria; medical history; smoking history; number of exacerbations during the previous 2 years; physical examination; lung-function tests (including functional residual capacity [FRC] and diffusing capacity as measured by single-breath carbon monoxide transfer factor [TLCO]); chest radiograph; electrocardiogram; laboratory tests; the St George's respiratory questionnaire; and the Euroqol-5D questionnaire.

If the patient complied with the inclusion criteria, he or she was asked to sign the informed-consent form, was randomly allocated to a treatment group, and was asked to take the tablets for the next 3 years. At each subsequent visit, the following were recorded: physical examination; lung function; smoking habits; adverse events; patient diary card; and compliance. There was no run-in or optimisation of treatment period at the beginning of the trial, and we did not try to establish optimum treatment. Patients were strongly advised to quit smoking before the trial. Only ex-smokers (defined as patients who had stopped smoking for at least 6 months before the trial) or smokers who did not change their habits were enrolled. Smoking habits were verified by history only.

During exacerbations, diary cards were completed daily. The St George’s respiratory questionnaire and the Euroqol-5D were completed at 6, 12, 18, 24, 30, and 36 months. The study was approved by the ethics committee of the different centres, and written informed consent was obtained from all patients.

At every centre, patients were assigned in numerical order to 600 mg daily N-acetylcysteine or to placebo, on the basis of a predetermined computer-generated randomisation list provided by the data-management unit. This unit was responsible for the collection and quality control of data during the rest of the trial. This list was generated so that the two groups were balanced by centre and within centre by use of inhaled corticosteroids at entry. Each centre was provided with the following sets of tablets: 1–50 to be assigned to patients taking inhaled corticosteroids at entry and 51–99 to be assigned to patients who were not taking these steroids at entry. Supplies of tablets for every patient were identified by a four-digit number (the first two identifying the centre and the last two denoting the order number from one to 99). The randomisation list was based on a block size of four.

The study was done under double-blind conditions—ie, neither the investigator nor the patient knew to which group they were assigned. N-acetylcysteine tablets and placebo tablets looked identical, and were packaged and labelled so they could not be identified. The randomisation code for any patient was not to be broken by the investigator during the study, except in the circumstance of a serious life-threatening adverse event. Any opening of a sealed envelope that contained the randomisation code, whether deliberate or accidental, had to be carefully recorded on the case-report form. The steering committee and data-management unit were masked to the treatment allocations during the study.

Patients were compliant if they took at least 80% of tablets. Patients were instructed to bring their tablets with them to every visit, and compliance was assessed at every visit by counting of returned boxes. The compliance check was not to take place in the presence of the patient, and the number of returned tablets was recorded in the case-report form.

Recruitment ran from June 1, 1997, to Jan 15, 2000. The study was completed by Jan 16, 2003. Patients were included from 50 centres in ten European countries. Eligible patients were aged 40–75 years and had: smoking-related COPD; postbronchodilator forced FEV1 of 40–70% of predicted; FEV1 and vital capacity (VC) reversibility of less than 12% of predicted and of less than 200 mL 15 min after 400 μg of salbutamol (Ventolin, Glaxo Smithline, Stockley Park, UK) by metered-dose inhaler and spacer; a ratio of FEV1 to VC of less than 88% of predicted in men and less than 89% of predicted in women; and history of at least two exacerbations per year during the 2 years before enrolment.

Exclusion criteria were: history of intolerance of N-acetylcysteine; continuous treatment with oral corticosteroids; treatment with N-acetylcysteine for 3 months or longer; history of asthma, allergic rhinitis, or allergic eczema; likely to have long-term oxygen treatment within the next 3 years; α1-antitrypsin deficiency; cystic fibrosis; bronchiectasis; history of infection or active infection due to Mycobacterium tuberculosis (except for patients with small old tuberculous lesions); pneumoconiosis; pulmonary restriction because of other pulmonary disease; history of active peptic-ulcer disease or intestinal malabsorption; congestive heart failure class 2 or more of the New York Heart Association; reduced life expectancy because of other disease; evidence of illicit drug use or alcohol abuse; planned for lung transplantation at time of enrolment or expected to undergo transplantation within 3 years; enrolment in a rehabilitation programme participation in another trial within 3 months before the beginning of this study; and no compliance in taking tablets. Patients who used any type of bronchodilator were eligible if they met the inclusion criteria. Inhaled corticosteroids were allowed, but patients were recommended to use the same dose throughout the study. Concomitant use of other drugs was recorded in the case-report form. Patients were advised not to use vitamins or other dietary supplements. Preplanned subgroup analyses were defined by use of inhaled corticosteroids at entry and by disease severity (ie, FEV1, >50% and <50% predicted). Subgroup
analysis was later redefined by GOLD (global initiative on obstructive lung disease) stages II and III, which were similar to the predefined analyses.

Spirometry was done at each visit14 according to the European Respiratory Society standards.15 At every visit FEV₁ and slow inspiratory VC were recorded. All measurements were done before patients received 400 μg salbutamol and at 15 min after. Patients were asked to take tablets about 2 h before testing. Salbutamol doses were added to the dose of regular inhaled treatment to obtain a maximum postbronchodilator FEV₁. The ratio of FEV₁ to VC was calculated. The VC used was the slow VC, compatible with the European Respiratory Society definition of Tiffeneau index.16 All measurements reported in the text refer to postbronchodilator measurements.

Reversibility was expressed as a percentage of predicted FEV₁ and VC. At the screening visit and at the end of the study, diffusing capacity (TLCO) and FRC were measured in the lung-function laboratory. Thoracic-gas volume or FRC was measured by whole-body plethysmography or by dilution method. TLCO was assessed by the single-breath method of measuring carbon monoxide. At the end of the study these measurements were repeated on the same device used at the screening visit. All spirometric procedures were supervised by the quality-control committee.

Exacerbations were defined as an increase in dyspnoea, cough, or both associated with a change in quality and quantity of sputum, which led the patient to seek medical attention and which lasted for at least 3 days. The same definition was used for exacerbations occurring before and during the trial. The duration of exacerbations was defined as the duration of the medical intervention. The patient was instructed to call the investigator at the onset of the exacerbation, which thus lasted from the first telephone call until the end of medical treatment. Exacerbations leading to hospitalisation were defined as severe; those not needing hospitalisation were not severe. On a posthoc basis, exacerbations that were not severe were subdivided into moderate for those needing treatment with oral steroids or antibiotics and mild for those that did not.

Quality of life was assessed by the St George’s respiratory questionnaire (a disease-specific assessment) and by the Euroqol-5D (a generic method of assessment).17–19 Euroqol-5D utilities were calculated by the York-A1 tariff based on valuations from the UK public.21 Scores from the St George’s respiratory questionnaire were presented on a scale from 1 to 100 (increasing scores indicate worsening), and Euroqol-5D data are presented on a scale of 0 to 1 (increasing scores indicate improvement). Both questionnaires were administered every 6 months during the visits. To standardise conditions, questionnaires were completed in a separate room in the clinic before patients were examined. A nurse gave the patient instructions on how to complete the questionnaire, but did not interpret the questionnaire for the patient and left the room while the patient did this task. Scores were calculated according to the guidelines in the manuals, which state that St George’s respiratory questionnaires with more than 10 missing values and Euroqol-5D questionnaires with one or more of the five items missing are invalid and should not be included in analyses.

Statistical analysis
Analyses were done by intention to treat—ie, on all randomised patients who had at least one FEV₁ measurement after randomisation and on all patients who completed the trial. The presentation of results is confined to the analysis on all randomised patients.

Calculation of the sample size needed to detect a significant effect of N-acetylcysteine was based on its presumptive effect on decline in FEV₁. The following assumptions were made: the yearly decline in FEV₁ was 20 mL in healthy individuals and 60 mL in patients with COPD; N-acetylcysteine reduces this rate of yearly decline in patients with COPD to 40 mL; and the SD of the yearly decline is 55 mL. The assumptions on drop-out were that 35% of patients would be lost during 3 years’ treatment and that 5% would not be evaluable (ie, would have no FEV₁ measurement after randomisation). With these assumptions, 239 patients were needed in each treatment group, 478 overall, to detect a significant difference (p=0·05, two-sided) with 80% power. The number of patients needed to show a 20% reduction in yearly exacerbation rate or a four-point increase in the St George’s respiratory questionnaire was substantially smaller than that needed for assessment of FEV₁ decline (334 and 280 people, respectively).

For assessment of pulmonary function and health status over time, data analysis was done by mixed-effect models with random intercept and random slope,20 which allow use of all data. Analyses for FEV₁ were based on a mean of 9·5 measurements (SD 3·7) per patient in the N-acetylcysteine group and 8·8 measurements (4·1) in the placebo group. For St George’s respiratory questionnaire, analyses were based on 5·4 measurements (2·1) per patient assigned N-acetylcysteine and 5·1 measurements (2·3) per patients assigned placebo. Missing data were assumed to be missing at random—ie, that patients with incomplete data differ from those with complete data, but that the pattern of incompleteness is predictable from other variables in the database.25 In the models for pulmonary function (ie, FEV₁ and VC) baseline data were treated as a covariate and all measurements from the second visit onwards were treated as outcomes. Data from the first year were analysed separately from the period afterwards because health status improved during the first year and deteriorated thereafter.

In the models for the first year, baseline data were treated as covariates and all measurements during the
first year after the baseline visit were treated as outcomes. In the models for the period after the first year, baseline data were treated as covariates and all measurements from 12 months onwards were treated as outcomes. Other covariates in all models were: centre, GOLD classification (II or III), use of inhaled corticosteroids at entry (yes or no), smoking status (smoker or ex-smoker), and age (=60 years or >60 years). Centre was classed as a fixed effect. The time variable entered in all models was year. The treatment×time interaction tested for a differential effect of treatment on the rate of change in FEV1, VC, and in health status. Data are reported as a difference in slope between the N-acetylcysteine group and placebo group. For lung function, a negative value indicates a faster drop in the N-acetylcysteine group than in the placebo group, and a positive value indicates a faster decline in the placebo group than in the N-acetylcysteine group. For the St George’s Respiratory Questionnaire, a positive value during the first year indicates a slower improvement in the N-acetylcysteine group than in the placebo group, whereas a positive value during subsequent years indicates a faster deterioration in the N-acetylcysteine group. For Euroqol-5D, a positive value during the first year indicates a faster improvement in the N-acetylcysteine group than in the placebo group, whereas a positive value during the years thereafter indicates a slower deterioration in those assigned N-acetylcysteine. Prespecified subgroup analyses were done for those who had moderate or severe COPD and for patients who did or did not use inhaled corticosteroids at entry by including the interaction terms GOLD×treatment, GOLD×time, and GOLD×treatment×time or by inhaled corticosteroids×treatment, inhaled corticosteroids×time, and inhaled corticosteroids×treatment×time, respectively. The effect of treatment on FRC and diffusing capacity was assessed by a Wilcoxon rank sum test applied to change versus baseline. The effect of treatment on FRC was analysed by linear regression with the same covariates for FEV1, reduction with centre as a fixed effect. Comparison of the proportion of patients who withdrew was done by a Cochran-Mantel-Haenszel test, and the relation between treatment and time to withdrawal was analysed by the log-rank test. Exacerbations were analysed by a multiple-failure time approach as suggested by Andersen and Gill.28 We used a multiplicative-hazards model that calculates failure time for each patient. Events were assumed to be independent. Thus, for all patients the time of every exacerbation was recorded. Patients were censored at the time they dropped out of the study or did not have any further exacerbation. Time to exacerbation was calculated as the time from the beginning of the study to the first exacerbation for the first exacerbation, and as the difference between the times of two consecutive exacerbations for subsequent occurrences. For patients who finished the study, the last day of the study was the last follow-up time; for those who dropped out, their drop-out date was the last follow-up date. Results are reported as a risk ratio for an exacerbation in the N-acetylcysteine group versus the placebo group. A risk ratio less than one thus indicates a protective effect of N-acetylcysteine. Subgroup analysis was done, including the interaction terms GOLD×treatment and inhaled corticosteroids×treatment in the model. For a significant interaction, marginal proportional hazards were modelled. All hypothesis testing was two-sided, and p<0·05 was defined as significant. All analyses were done with SAS version 8.2 (Cary, NC, USA). Role of the funding source The funding source had no decisive role in the collection, management, analysis, or interpretation of the data; writing of the report; or the decision to submit the paper for publication. The corresponding author had full access to all the data in the study and had final responsibility to submit for publication. Results Table 1 shows baseline characteristics of patients. The number of patients enrolled exceeded that estimated and the true drop-out rate was less than assumed—70 (27%) of 256 in the N-acetylcysteine group and 99 (37%) of 267 in the placebo group (figure 1). Drop-out rate was significantly greater in the placebo group than in the N-acetylcysteine group (p=0·018), and time to withdrawal was significantly shorter in patients assigned placebo than in those assigned N-acetylcysteine (p=0·014). Drop-out gradually fell from year 1 to year 3. 

<table>
<thead>
<tr>
<th>N-acetylcysteine (n=256)</th>
<th>Placebo (n=267)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>53 (21%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>62 (8)</td>
</tr>
<tr>
<td>Current smokers</td>
<td>130 (51%)</td>
</tr>
<tr>
<td>GOLD II</td>
<td>190 (74%)</td>
</tr>
<tr>
<td>GOLD III</td>
<td>66 (26%)</td>
</tr>
<tr>
<td>Use of inhaled corticosteroids</td>
<td>182 (71%)</td>
</tr>
<tr>
<td>Average daily dose of inhaled corticosteroids (mg equivalent fluticasone)</td>
<td>579 (374)</td>
</tr>
<tr>
<td>Use of short-acting β2 agonists</td>
<td>182 (71%)</td>
</tr>
<tr>
<td>Use of short-acting anticholinergics</td>
<td>90 (35%)</td>
</tr>
<tr>
<td>Use of long-acting β2 agonists</td>
<td>161 (63%)</td>
</tr>
<tr>
<td>Use of theophylline</td>
<td>87 (34%)</td>
</tr>
<tr>
<td>VC (L)</td>
<td>3·41 (0·85)</td>
</tr>
<tr>
<td>FEV1 (L)</td>
<td>1·65 (0·38)</td>
</tr>
<tr>
<td>Predicted FEV1</td>
<td>57% (9)</td>
</tr>
<tr>
<td>FRC (L)</td>
<td>4·43 (1·30)</td>
</tr>
<tr>
<td>Reversibility (% predicted)</td>
<td>4% (4)</td>
</tr>
<tr>
<td>Yearly exacerbation rate before study (events)</td>
<td>2·4 (0·7)</td>
</tr>
<tr>
<td>St George’s respiratory questionnaire total score</td>
<td>39 (16)</td>
</tr>
<tr>
<td>Euroqol-5D score</td>
<td>0·76 (0·22)</td>
</tr>
</tbody>
</table>

Data are number (%), mean (SD), or % (SD).

Table 1: Baseline characteristics
Articles

Table 2 shows characteristics of drop-outs compared with completers. 230 (90%) of 256 patients assigned N-acetylcysteine were compliant with treatment compared with 225 (84%) of 267 those assigned placebo. Compliance with treatment was a mean of 94% of prescribed tablets in the N-acetylcysteine group and 92% for placebo.

During the trial, the decline in FEV₁, was close in the N-acetylcysteine group and placebo group (54 mL [SE 6] vs 47 mL [6]; mean difference in slope between groups –8 mL [95% CI –25 to 10, figure 3]). No difference in slope between the two groups was seen on analysis of subgroups defined on the basis of GOLD classification (GOLD×treatment×time interaction, p=0·688) or use of inhaled corticosteroids (inhaled corticosteroids×treatment×time interaction, p=0·097). For reduction in FEV₁, the mean difference in slope was –8 mL (SE 95% CI –25 to 10, p=0·393) for all patients; 2 mL (11) –20 to 24, 0·859) for patients taking inhaled corticosteroids; –2 mL (21) –43 to 39, 0·918) for those not taking inhaled corticosteroids; –12 mL (19) –49 to 25, 0·538) for those in GOLD stage II; and –13 mL (17) –47 to 21, 0·446) for patients in GOLD stage III.

Secondary analysis showed that the reduction in diffusion capacity during 3 years did not differ between the N-acetylcysteine group and the placebo group (–0·59 mmol/min/kPa [SD 2·21] vs –0·60 mmol/min/kPa [1·68]; p=0·493). FRC decreased significantly from 4·46 L (SD 1·24) to 4·09 L (1·23) in patients assigned N-acetylcysteine (difference –0·374 L [SD 1·03]; p<0·0001 for 120 patients who completed treatment).

A similar pattern was recorded for the drop in VC. However, for patients in GOLD stage III the difference in slope between N-acetylcysteine and placebo was 36 mL per year in favour of N-acetylcysteine (p=0·2863). The mean differences in slope were 0 mL (SE 17) for all patients; –2 mL (21) –43 to 39, 0·890) for those taking inhaled corticosteroids; 4 mL (26) –48 to 56, 0·890) for those not taking inhaled corticosteroids; –12 mL (19) –49 to 25, 0·538) for those in GOLD stage II; and 36 mL (34) –30 to 102, 0·289) for those in GOLD stage III.

Table 2: Characteristics of drop-outs

<table>
<thead>
<tr>
<th>Completed (n=354)</th>
<th>Drop-outs (n=169)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>82 (23%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>62 (8)</td>
</tr>
<tr>
<td>Current smokers</td>
<td>155 (44%)</td>
</tr>
<tr>
<td>Use of inhaled steroids</td>
<td>248 (70%)</td>
</tr>
<tr>
<td>GOLD II</td>
<td>273 (77%)</td>
</tr>
<tr>
<td>GOLD III</td>
<td>81 (23%)</td>
</tr>
<tr>
<td>VC (L)</td>
<td>3·43 (0·87)</td>
</tr>
<tr>
<td>FEV₁ (L)</td>
<td>1·66 (0·39)</td>
</tr>
<tr>
<td>Predicted FEV₁</td>
<td>57% (9)</td>
</tr>
<tr>
<td>FRC (L)</td>
<td>4·33 (1·23)</td>
</tr>
<tr>
<td>St George’s respiratory questionnaire total score</td>
<td>38 (15)</td>
</tr>
<tr>
<td>Euroqol-5D score</td>
<td>0·78 (0·23)</td>
</tr>
</tbody>
</table>

Data are number (%), mean (SD), or % (SD).

Figure 2: Kaplan-Meier plot of drop-out during the trial
compared with from 4·33 L [1·16] to 4·34 L [1·20] in those assigned placebo (difference 0·008 L [SD 0·943]; p=0·99 for 107 patients who completed treatment). When analysed per protocol, the difference in FRC between the N-acetylcysteine group and the placebo group was significant for 227 (64%) of 354 patients who completed the trial (p=0·008). Similarly, regression analysis then showed a significant effect of treatment on FRC (p=0·003).

Table 3 shows the exacerbation rate in both groups and risk ratios for exacerbations in the two groups. The total number of events was 1845, 494 (27%) of which were censored. Residuals were skewed, especially with the interaction inhaled corticosteroids×treatment and with marginal results. However, no particular outliers could be identified. Yearly exacerbation rates did not differ between patients allocated N-acetylcysteine and those allocated placebo (1·25 [SD 1·35] vs 1·31 [1·39]; HR 0·99 [95% CI 0·89–1·10, p=0·85]). In patients who were not taking inhaled corticosteroids (n=155), risk of exacerbation was lower for those assigned N-acetylcysteine than for those assigned placebo (0·96 [SD 1·36] vs 1·29 [1·46]; HR 0·79 [95% CI 0·631 to 0·989], p=0·040). The interaction inhaled corticosteroids×treatment was significant (p=0·024). Similarly, patients who were not taking inhaled corticosteroids and were assigned N-acetylcysteine had fewer moderate or severe exacerbations than did those who were assigned placebo (0·76 [1·14] vs 1·11 [1·39]; HR 0·762 [95% CI 0·595–0·976, p=0·032]); inhaled corticosteroids×treatment interaction was also significant (p=0·037). No effect on exacerbation rate was recorded in smokers or non-smokers, or in patients in GOLD stage II or III.

The number of patients included in the health-status analyses was 445 for the St George’s respiratory questionnaire and 403 for the Euroqol-5D questionnaire. This number is lower than anticipated because a translation of the St George’s respiratory questionnaire was not available for Estonia, and translations of Euroqol-5D were not available for Portugal, Poland, and Estonia. Furthermore, only 22 (37%) of 59 patients from Poland completed the St George’s respiratory questionnaire at screening, because the questionnaire was not made available in time.

All questionnaire scores recorded significant improvements in health status during the first year of the study in patients assigned N-acetylcysteine and those assigned placebo (p<0·002 for the St George’s respiratory questionnaire scores and p=0·024 for Euroqol-5D). All improvements in the first year recorded by St George’s respiratory questionnaire scores were close to or greater than the four units’ threshold for the minimum clinically important difference (−3·76 [95% CI −5·55 to −1·98] for N-acetylcysteine and −4·95 [−6·74 to −3·16] for placebo; difference between groups 1·18 [−1·34 to 3·7], p=0·358). For Euroqol-5D, first-year improvements were 0·018 (−0·015 to 0·05) for N-acetylcysteine and 0·037 [0·004 to 0·069] for placebo (difference between groups 0·019 [−0·065 to 0·027], p=0·414).

After the first year, scores from both questionnaires significantly worsened in both groups, except for the St George’s respiratory questionnaire symptom score. Yearly drop in St George’s respiratory questionnaire total score was 1·45 (0·52 to 2·38) for N-acetylcysteine and 1·24 (0·28 to 2·21) for placebo (difference between groups 0·21 [−1·13 to 1·55], p=0·763); yearly decrease in St George’s respiratory questionnaire symptom score was 0·62 (−0·72 to 1·95) and 0·14 (−1·25 to 1·53), respectively (difference 0·47 [−1·45 to 2·39], p=0·631, figure 4). Yearly decrease in Euroqol-5D was −0·02

![Figure 3: Reduction in FEV, in the two groups](image_url)

![Figure 4: Change in health status, measured with St George’s respiratory questionnaire, during the trial](image_url)
Articles and health status, we investigated potential effects of steroids affect lung function, yearly exacerbation rate, N-acetylcysteine per day for patients with COPD did not This study found that treatment with 600 mg oral Discussion and nine patients in each group died. and 49 assigned placebo had exacerbation due to COPD, those admissions, 42 patients assigned N-acetylcysteine the placebo group who were admitted were observed. Of those admissions, 42 patients assigned N-acetylcysteine and 49 assigned placebo had exacerbation due to COPD, and nine patients in each group died.

We do not know why no effects on health status with N-acetylcysteine were noted, or whether more focused symptom scores might be affected by N-acetylcysteine, although improvement in symptoms have been recorded with this compound. In this study, St George’s respiratory questionnaire scores did not decrease as much as they did in the ISOLDE (Inhaled Steroids in Obstructive Lung Disease in Europe) trial: total score in the St George’s respiratory questionnaire decreased by 3–2 units per year in the placebo group compared with 2–0 units per year in the fluticasone-propionate group in ISOLDE. In BRONCUS, the St George’s respiratory questionnaire total score decreased by about 1–2–1·5 units per year in both the placebo and treatment groups during the last 2 years of the trial. The difference between the two trials is probably related to the fact that our study included patients with less severe COPD, whose lung function was not restored optimally with a course of oral prednisolone before randomisation, as was done in the ISOLDE trial.

Although the results of subgroup analysis showing a decreased risk for exacerbation in patients not taking inhaled corticosteroids assigned N-acetylcysteine are only suggestive, the magnitude of the reduction was consistent with the results of meta-analyses and with the reductions recorded with patients taking inhaled steroids, tiotropium, or combination drugs. Calverley and co-workers’ study showed no additive effect of a combination treatment compared with inhaled steroid alone. Calverley and co-workers later reported that in patients with severely reduced FEV1, budesonide combined with formoterol had greater effects on exacerbation rate than did either agent alone. However, because formoterol did not affect exacerbations, budesonide was less effective than fluticasone in the earlier study, and the overall effect was much the same in both studies, the later study confirms rather than contradicts the results of the previous study. Whether the effects of these drugs on yearly exacerbation rate are additive or not remains unanswered, even though in general they are of importance for the treatment of COPD. Indeed, the importance of exacerbations for the progressive deterioration of pulmonary function and health status is well recognised. The data from the present study and from that of Calverley and co-workers suggested that the effects of the drugs are not additive.

The present data cannot be used to make conclusions about the mechanism of a potential reduction in exacerbation rate, although an antioxidant and anti-inflammatory effect, as well as a reduction in bacterial adherence to the bronchial mucosa, might be implicated. Moreover, an effect on mucus clearance might be beneficial for COPD exacerbations, which are related to chronic hypersecretion, FEV1 decline, and admission. However, we did not identify any beneficial outcomes.

We do not know why no effects on health status with N-acetylcysteine were noted, or whether more focused symptom scores might be affected by N-acetylcysteine, although improvement in symptoms have been recorded with this compound. In this study, St George’s respiratory questionnaire scores did not decrease as much as they did in the ISOLDE (Inhaled Steroids in Obstructive Lung Disease in Europe) trial: total score in the St George’s respiratory questionnaire decreased by 3–2 units per year in the placebo group compared with 2–0 units per year in the fluticasone-propionate group in ISOLDE. In BRONCUS, the St George’s respiratory questionnaire total score decreased by about 1–2–1·5 units per year in both the placebo and treatment groups during the last 2 years of the trial. The difference between the two trials is probably related to the fact that our study included patients with less severe COPD, whose lung function was not restored optimally with a course of oral prednisolone before randomisation, as was done in the ISOLDE trial.

Although the results of subgroup analysis showing a decreased risk for exacerbation in patients not taking inhaled corticosteroids assigned N-acetylcysteine are only suggestive, the magnitude of the reduction was consistent with the results of meta-analyses and with the reductions recorded with patients taking inhaled steroids, tiotropium, or combination drugs. Calverley and co-workers’ study showed no additive effect of a combination treatment compared with inhaled steroid alone. Calverley and co-workers later reported that in patients with severely reduced FEV1, budesonide combined with formoterol had greater effects on exacerbation rate than did either agent alone. However, because formoterol did not affect exacerbations, budesonide was less effective than fluticasone in the earlier study, and the overall effect was much the same in both studies, the later study confirms rather than contradicts the results of the previous study. Whether the effects of these drugs on yearly exacerbation rate are additive or not remains unanswered, even though in general they are of importance for the treatment of COPD. Indeed, the importance of exacerbations for the progressive deterioration of pulmonary function and health status is well recognised. The data from the present study and from that of Calverley and co-workers suggested that the effects of the drugs are not additive.

The present data cannot be used to make conclusions about the mechanism of a potential reduction in exacerbation rate, although an antioxidant and anti-inflammatory effect, as well as a reduction in bacterial adherence to the bronchial mucosa, might be implicated. Moreover, an effect on mucus clearance might be beneficial for COPD exacerbations, which are related to chronic hypersecretion, FEV1 decline, and admission. However, we did not identify any beneficial outcomes.

We do not know why no effects on health status with N-acetylcysteine were noted, or whether more focused symptom scores might be affected by N-acetylcysteine, although improvement in symptoms have been recorded with this compound. In this study, St George’s respiratory questionnaire scores did not decrease as much as they did in the ISOLDE (Inhaled Steroids in Obstructive Lung Disease in Europe) trial: total score in the St George’s respiratory questionnaire decreased by 3–2 units per year in the placebo group compared with 2–0 units per year in the fluticasone-propionate group in ISOLDE. In BRONCUS, the St George’s respiratory questionnaire total score decreased by about 1–2–1·5 units per year in both the placebo and treatment groups during the last 2 years of the trial. The difference between the two trials is probably related to the fact that our study included patients with less severe COPD, whose lung function was not restored optimally with a course of oral prednisolone before randomisation, as was done in the ISOLDE trial.
Long-term studies of deterioration in pulmonary function always have several drawbacks, which might hinder interpretation of data. The main difficulty is drop-out, which we assumed would be 40% overall whereas it was only 32%. Consequently, we retained sufficient power to draw conclusions about the primary endpoint of FEV₁, decline. Power, however, is not sufficient to conclude with certainty about lung-function decline in subgroups. Only one difference proved to be fairly large—ie, the effect of N-acetylcysteine on VC decline for patients in GOLD stage III, which was 36 mL per year. This difference might be clinically important but was not significant. Several findings in our study indicate that drop-out might not be completely random, but that the most severely ill patients drop-out first. Drop-out was greater in the placebo group than in the active-treatment group, and similar findings have been reported in other long-term studies of COPD. However, our analysis of lung function and health status applied mixed-effect models using baseline values as covariates and to some extent, these models do account for not completely random drop-out.

In conclusion, we noted that 600 mg oral N-acetylcysteine per day in patients with COPD did not affect the rate of decline in FEV₁ or VC, exacerbation rate, or health status. Subgroup analysis suggests that N-acetylcysteine might reduce exacerbation rate in patients not taking inhaled steroids. Secondary analysis of FRC data suggests that N-acetylcysteine might reduce hyperinflation. Because N-acetylcysteine is very well tolerated, higher doses such as 1200 mg or 1800 mg per day could be assessed in future trials.

Conflict of interest M Decramer, M Rutten-van Mólen, R Dekhuijzen, T Troosters, C van Herwaarden, O van Schaayk, D Olivier, I Lankhorst, and A Ardia, who were on the Steering Committee, have designed the study, reviewed the analysed data and wrote the article. R Pellegrino did quality control, reviewed the analysed data, and helped with the writing of the article. M Del Donno and W De Backer have been consultants for Zambon. I Lankhorst has been employed as Medical Manager of Zambon, but is now an independent consultant for Zambon. A Ardia was employed by Zambon as a statistician, but is retired now. A Peviani and G Corvasce are employed by Zambon.

Acknowledgments We thank Antonella Peviani and Giuseppina Corvasce for their technical help. The study was funded by Zambon Group Spa, Bresso-Milan, Italy.

References


Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis

Charles H King, Katherine Dickman, Daniel J Tisch

Summary

Background Schistosomiasis is one of the world’s most prevalent infections, yet its effect on the global burden of disease is controversial. Published disability-adjusted life-year (DALY) estimates suggest that the average effect of schistosome infection is quite small, although this is disputed. To develop an evidenced-based reassessment of schistosomiasis-related disability, we did a systematic review of data on disability-associated outcomes for all forms of schistosomiasis.

Methods We did structured searches using EMBASE, PUBMED, and Cochrane electronic databases. Published bibliographies were manually searched, and unpublished studies were obtained by contacting research groups. Reports were reviewed and abstracted independently by two trained readers. All randomised and observational studies of schistosomiasis morbidity were eligible for inclusion. We calculated pooled estimates of reported disability-related effects using weighted odds ratios for categorical outcomes and standardised mean differences for continuous data.

Findings 482 published or unpublished reports (March, 1921, to July, 2002) were screened. Of 135 selected for inclusion, 51 provided data for performance-related symptoms, whereas 109 reported observed measures of disability-linked morbidities. Schistosomiasis was significantly associated with anaemia, chronic pain, diarrhoea, exercise intolerance, and undernutrition.

Interpretation By contrast with WHO estimates of 0-5% disability weight assigned to schistosomiasis, 2–15% disability seems evident in different functional domains of a person with schistosomiasis. This raised estimate, if confirmed in formal patient-preference studies, indicates a need to reassess our priorities for treating this silent pandemic of schistosomiasis.

Introduction

Chronic schistosomiasis, caused by parasitic blood flukes Schistosoma haematobium, S mansoni, S intercalatum, S mekongi, or S japonicum, is one of the most prevalent infectious diseases. More than 200 million people are infected worldwide, most of whom live in sub-Saharan Africa. A mature, patent, schistosome infection is associated with chronic tissue inflammation. The range and potential severity of symptoms and pathological changes associated with schistosomiasis have been well described. By contrast, the disabling effects of these clinical manifestations have not been adequately quantified. In particular, population-based studies have not provided a clear estimate of the effect of chronic schistosomiasis on performance status or overall quality-of-life.

In the past two decades, national and international programmes have come to depend on cost-effectiveness analysis for their allocation of health-care resources. This approach needs fairly precise estimates of the disutility (eg, death and disability) of the health conditions that are to be prioritised, and controversy persists over what is the best means to measure and compare the burden of different diseases. In 1995–96, as a first approach to effecting this comparison, the Global Burden of Disease Programme developed disability-adjusted life-year (DALY) estimates for many diseases. Since then, WHO, the World Bank, and other agencies have used DALY estimates to rank the effect of different diseases on world health. DALY estimates were based on a weighting scale developed by a panel of experts convened in Geneva. However, studies have cast doubt on the ability of experts to assess outlook for patients with chronic diseases, characterised by low mortality, especially for patients with comorbidities. For example, the age-specific DALY weights for schistosomiasis, assigned by the Global Burden of Disease Programme, ranged from 0.005 to 0.006, which is similar to those for disorders such as moderate discolouration of the face (facial vitiligo). Although this process was reproducible in terms of the experts’ ordinal ranking of disability caused by different disorders, the magnitude of the weights for schistosomiasis and other helmintic infections, as derived by this Delphi method, seemed exceptionally low to many specialists.

Over the past 80 years, several workers in observational and interventional studies have independently attempted to estimate the effect of schistosome infection on the health and performance status of infected individuals. Others have quantified the
schistosomiasis-specific risk of morbidities known to be linked to disability. These studies have yielded conflicting conclusions about the functional significance of schistosome infection.\textsuperscript{14–16} Our aim was to critically review the available evidence on disability caused by chronic schistosomiasis, and to use summary data from the abstracted studies to develop a working, evidence-based estimate of the disability-weight associated with chronic schistosome infection.

**Methods**

**Search strategy**

We identified published studies using the National Library of Medicine’s MEDLINE computer database and Elsevier’s EMBASE database with the key words listed in the panel. Database searches were restricted to the period from when the earliest electronic records were available (Jan 1, 1966) to when we obtained data for this study (July 31, 2002). Additional studies published before and during this period were identified from the reference lists in the reports, by hand searches of published abstracts of old literature,\textsuperscript{17} from personal collections at Case Western Reserve University, from the Cochrane databases, and from a working paper that estimated global number of cases of schistosomiasis.\textsuperscript{18} Unpublished study results were obtained from international schistosomiasis researchers via e-mail through a pathogen-specific interest group sponsored by the US National Institutes of Health. Webtables 1 and 2 show details of studies that were included in and excluded from our meta-analysis.

**Selection criteria**

We included all published and unpublished studies in which schistosomiasis-associated morbidity was quantified, between March 1, 1921, and July 31, 2002. Data without quantifiable measures of human schistosomiasis morbidity were excluded (figure 1). Reports in languages other than English were assessed against inclusion and exclusion criteria after translation by the authors or by our colleagues fluent in the language at Case Western Reserve University School of Medicine. We identified studies for potential inclusion that were published in Portuguese, French, Chinese, and Japanese. Identified reports were retrieved and collected in binders. Two experienced, independent reviewers (CK, KD) assessed every report against the inclusion and exclusion criteria according to a written study protocol. The reports were logged and indexed, then assessed for their content on disability-outcome data.

**Outcome assessments**

Because very few reports specifically addressed health-related quality of life, we abstracted and compared the following subcomponents, which are generally included in assessments of health-related quality of life: increased health-care needs and deficits in individual performance—ie, exercise or play, work yield, housework, personal care, religious activity, school attendance, school performance, cognition, and exercise intolerance. Additionally, we looked at prevalence or severity of morbidities that could be logically linked to physical, mental, or social disability—ie, abdominal pain, diarrhoea, dysuria, infertility, oxygen uptake ($\text{VO}_2\text{max}$) deficit, and anaemia—as well as undernutrition measures, including deficits in weight, height, weight for height, skin-fold thickness, percentage of body fat, and serum protein concentrations.

In the first phase of assessment, disability-related outcomes were analysed separately. As such, a specific research study could contribute more than one disability outcome for inclusion in our survey; however, for each

---

**Panel: Key words**


---

**Figure 1: Flow diagram, showing selection of studies**
disability category, only one outcome could be included per study (eg, only one outcome was abstracted to the study database if multiple anaemia outcomes were reported, as in studies recording haemoglobin and haematocrit [packed cell volume], as well as the numbers of children with anaemia). We recorded the study-specific definition for each disability measure because of the variation in published anaemia definitions (eg, according to age and region) and the complexity of abstracting aggregate data. We categorised disability outcomes by type using the outcome categories, study size, location, and year of study. The reported disability was also classified as a correlate of infection prevalence, infection intensity, or response to specific antischistosomal treatment. Potential sources of heterogeneity in the study results were investigated, where possible, with predefined stratification criteria—ie, by region, study period, type of schistosomiasis, intensity of schistosome infection, and study design (observational, interventional, or randomised controlled trial).

Statistical methods
Results of individual reviews of reports were entered into a Microsoft Access (Microsoft, Redmond, WA, USA) database developed for this project. Reviewers were asked to include the identifying author and journal citations, study design, inclusion and exclusion status, reasons for exclusion (if appropriate), study date and location, age range, sex ratio, type of treatment given (if any), and a summary of the authors’ primary interpretation of their disability-linked outcomes. Each citation entry was linked to a separate table cataloguing disability outcomes by type of outcome, the association with infection, infection intensity, and treatment.

To further develop a working consensus estimate for schistosomiasis-associated effects, we combined and assessed evidence for individual performance deficits or disability-linked morbidity outcomes associated with schistosomiasis infection. For categorical outcomes, odds ratios were listed and compared for the association between specific disability outcomes and infection (present or absent), or for low-intensity versus high-intensity infection. We then derived a separate estimate of schistosomiasis-specific effects on the outcomes studied, by weighted averaging of the results from individual studies. Weighting was based on the reported size of study subgroups. Validity of individual study methods was not formally quantified in the analysis because of the subjective nature of such analyses and the difficulty of detecting effect differences related to quality.19 For our analysis, we assumed that every report represented the best available information for the decision to submit for publication.

For continuous data outcomes, group-specific differences were compared with actual or normalised data from individual studies, and then a separate estimate was derived by weighted averaging of the results from individual studies. Analysis for heterogeneity of outcomes was done and adjusted for with the use of DerSimonian and Laird random-effects modelling, coded with the metan function in STATA (version 7.0).20 Results were displayed visually in forest plots with the “ci.plot” function of S-plus (version 2), adapted for use on the Windows operating system.21 In the various tables and figures presented, odds ratios greater than one and standardised mean differences less than zero indicated that morbidity or disability were associated, respectively, with infection, with greater infection intensity, or with non-treatment.

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.

Figure 2: Forest plot of the effect of schistosomiasis infection on anaemia, measured by haemoglobin concentration
Standardised mean differences less than zero indicate lower concentrations of haemoglobin in individuals with infection than in individuals without infection. Diamonds represent pooled estimates across grouped studies (eg, across species of infection). The lowest diamond indicates the overall pooled estimate of the effect of schistosomiasis infection on anaemia. See webtable 1 for full references.
The included studies represented every major human schistosomiasis species and endemic region. 118 investigated single-species infections: 14 were of patients with *S japonicum*, 60 focussed on those with *S mansoni* alone, and 44 assessed those with *S haematobium*. 17 studies provided data for patients with both single and multiple species infections. Of the *S japonicum* studies, six were done in China, six in the Philippines, and one in Japan. Of the *S mansoni* studies, six were undertaken in the Caribbean, one in central Africa, 14 in eastern Africa, 15 in Egypt or Sudan, one in Yemen, 16 in Brazil, five in South Africa, and two in western Africa. Of the *S haematobium* studies, eight were from central Africa, 18 were from eastern Africa, five were from Egypt or Sudan, ten were from southern Africa or Madagascar, and three were from western Africa.

Publication and selection biases were not detected for the various outcomes reported, either visually with funnel plots or statistically with the Begg test. We identified heterogeneity in several analyses using the χ² test, especially in observational studies. However, this heterogeneity generally diminished with stratified subanalysis based on schistosomiasis species and region of infection, or both. For example, overall, the 16 studies that investigated the effect of schistosomiasis infection on weight showed statistical heterogeneity (p<0·001), whereas results of analyses undertaken according to parasite species subsets did not (p=0·362 for *S mansoni*, 0·328 for *S haematobium*, and 0·272 for *S japonicum*). The χ² test for heterogeneity was significant for all pooled studies assessing anaemia according to infection status, as well as for results pooled by species (figure 2). Furthermore, results from this analysis pooled by region showed significant heterogeneity in studies from east Africa and northern Africa (p<0·001), but not from South Africa (p=0·201). Results from the randomised control trials (figure 3) did not show significant heterogeneity when pooled or when categorised by species or region. To keep to a minimum the effects of detected and undetected heterogeneities on study outcomes, all reported summary estimates were those
infected with all three species. When all studies were analysed irrespective of species, individuals with infection had significantly lower haemoglobin concentrations than did non-infected individuals (figure 2). A significant difference was not recorded for S haematobium infection per se, although heavy S haematobium infection was significantly associated with lower haemoglobin concentrations than was light infection status (figure 4). The only included study of S japonicum reported significantly reduced mean haemoglobin in infected individuals, 22 as did a multicentre treatment study that combined patients infected with all three species. 23 When all studies were analysed irrespective of species, individuals with infection had significantly lower haemoglobin concentrations than did those without infection (figure 2), an effect size suggesting that 18% of infected individuals were below the normal range of their respective local (uninfected) control groups. 24

The difference across all studies for infection intensity was also significant (figure 4). This difference, when analysed according to species, was significant in the subgroup infected with S haematobium. We recorded a stronger, but non-significant association in the aggregate studies of individuals infected with S mansoni than with S haematobium. This result for S mansoni was borderline, but became significant when an outlier study was removed (–0·63; 95% CI –0·95 to –0·30). Combined infection with S mansoni and S haematobium also contributed to greater anaemia in heavy than in light infections in one study. 25 Treatment significantly improved measures of anaemia in randomised controlled trials of individuals with S haematobium, S japonicum, and in studies in which all three infections were assessed (figure 3). Overall, across all treatment studies, specific antischistosomal treatment significantly improved anaemia (figure 3), indicating an infection-related effect size nearly equivalent to that estimated from observational studies.

Few studies reported other quantitative measures of morbidity or disability. As noted in table 1, summary estimates from available studies reporting measures of growth and nutrition indicated significant deficits in weight, weight for height, and skin-fold thickness occurring with schistosomiasis infection, although the extent of these effects differed between analyses done according to infection prevalence, infection intensity, or response to treatment. Infection status was also significantly associated with reduced exercise tolerance. No consistent effect was shown on linear growth (height), on maximum exercise capacity (VO2 max), on school performance, or on serum protein concentrations (data not shown).

The combined data for self-reported symptoms and other functional deficits for the included studies indicated that schistosomiasis infection was significantly associated with a subjective history of diarrhoea, a history of abdominal or genitourinary pain, and with fatigue or exercise intolerance (table 2). There was no

<table>
<thead>
<tr>
<th>How measured</th>
<th>Obsenational</th>
<th>Interventional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaemia</td>
<td>Haemoglobin</td>
<td>RCT placebo vs active schistosomicidal agent</td>
</tr>
<tr>
<td></td>
<td>Infected vs uninfected</td>
<td>Heavy vs light infection</td>
</tr>
<tr>
<td></td>
<td>(n=20)</td>
<td>(n=11)</td>
</tr>
<tr>
<td></td>
<td>–0·25 (–0·40 to –0·11)</td>
<td>–0·25 (–0·36 to –0·15)</td>
</tr>
<tr>
<td></td>
<td>(n=16)</td>
<td>(n=5)</td>
</tr>
<tr>
<td></td>
<td>–0·19 (–0·45 to 0·07)</td>
<td>–0·64 (–1·08 to –0·20)</td>
</tr>
<tr>
<td></td>
<td>(n=5)</td>
<td>(n=5)</td>
</tr>
<tr>
<td></td>
<td>–0·05 (–0·31 to 0·22)</td>
<td>–0·11 (–0·23 to 0·01)</td>
</tr>
<tr>
<td></td>
<td>(n=8)</td>
<td>(n=4)</td>
</tr>
<tr>
<td></td>
<td>–0·17 (–0·65 to 0·31)</td>
<td>–0·63 (–1·37 to 0·12)</td>
</tr>
<tr>
<td></td>
<td>(n=2)</td>
<td>(n=3)</td>
</tr>
<tr>
<td></td>
<td>–1·36 (–2·59 to –0·13)</td>
<td>–0·46 (–1·22 to –0·10)</td>
</tr>
<tr>
<td></td>
<td>(n=5)</td>
<td>(n=5)</td>
</tr>
</tbody>
</table>
|                               | Weight for height              | –2·25 (–5·37 to 0·87)                    | –2·25 (–5·37 to 0·87) (n=2) –
|                               | (n=5)                          | (n=2)                                    |
|                               | Skin fold thickness            | –0·66 (–1·22 to –0·10)                   |
|                               | (n=7)                          | (n=5)                                    |
|                               | Reduced fitness                | –0·66 (–1·22 to –0·10)                   |
|                               | Exercise duration              | –0·05 (–0·24 to 0·17)                    |
|                               | (n=4)                          | (n=4)                                    |
|                               | VO2 max                        | –2·25 (–5·37 to 0·87)                    |
|                               | (n=2)                          | (n=2)                                    |
|                               | Educational                    | –0·73 (–2·52 to 1·05)                    |
|                               | School performance             | –0·73 (–2·52 to 1·05)                    |
|                               | (n=2)                          | (n=2)                                    |

Data are standardised mean difference (95% CI) and number of studies. *Meta-analysis of religious activity, personal care, healthcare needs, cognition, housework deficit, and work yield deficit was restricted because of small sample size (ie, number of available studies for inclusion), or because of inconsistent measures used to assess morbidity outcomes.

Table 2: Summary estimates of the effect of schistosomiasis or heavy schistosomiasis by disability-related outcomes

<table>
<thead>
<tr>
<th>How measured</th>
<th>Infected vs uninfected subjects</th>
<th>Heavy vs light infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic symptoms</td>
<td>Diarrhoea history</td>
<td>1·59 (1·23–2·06) (n=27)</td>
</tr>
<tr>
<td></td>
<td>Pain history</td>
<td>1·50 (1·25–1·78) (n=30)</td>
</tr>
<tr>
<td>Reduced fitness</td>
<td>Exercise intolerance</td>
<td>2·25 (1·31–3·85) (n=13)</td>
</tr>
<tr>
<td>Education</td>
<td>Poor school performance</td>
<td>1·18 (0·72–1·93) (n=6)</td>
</tr>
<tr>
<td>Infertility</td>
<td>Obstetric history</td>
<td>1·45 (0·96–2·19) (n=3)</td>
</tr>
</tbody>
</table>

Data are odds ratios (95% CI) and number of studies.
significant association with poor school performance or infertility, although few studies reported these outcomes.

Meta-analysis was not possible for some important performance outcomes because of the low number of studies addressing such outcomes and because of the broad differences in the measures used to assess them. Of nine studies reporting work output,26–34 six indicated some infection-associated deficit in performance, and of four assessing cognitive ability of schoolchildren,35–38 three indicated some infection-associated deficit in multiple skills tested. One detailed case-control study showed a significant deficit in personal care and in religious activity associated with heavy S mansoni infection,30 and one survey study suggested an association between schistosomiasis and increased use of health-care services.39

Discussion
Our results show a significant association between human schistosome infection and the symptoms of diarrhoea, pain, and fatigue, as well as the objective findings of haemoglobin deficit, undernutrition, and reduced exercise tolerance. Assessment of results, which were consistent across many observational and interventional studies, suggests that there is a significant, infection-associated loss of performance in a person with schistosomiasis, which can be improved, at least in part, through intervention with specific antischistosomal treatment. Although in most cases schistosomiasis is unlikely to progress to a lethal outcome,46 it is a prominent, chronic, recurring infection in endemic areas, and its effect on the health status of infected individuals is clearly not negligible.

The outcomes selected for inclusion and analysis in our study were chosen to represent, as much as possible, their average effect on the performance status of all infected individuals.46 Therefore, we focused on symptoms of physical and social discomfort, along with measured outcomes that could be directly related to physical and intellectual function. Likely pathways for disease formation in schistosomiasis are summarised in figure 5, which displays the potential complexity of disease and the related clinical syndromes that can be associated with chronic schistosomiasis. We did not include morbidity outcomes of the objective findings of hepatosplenomegaly and radiograph or ultrasound imaging abnormalities, as has been done before,14,18 because their effect on daily performance status is unclear. Similarly, we did not include schistosomiasis-related mortality in our analysis. Reviewers41,42 have critically assessed the data on schistosomiasis-related mortality, and lend support to the notion that estimates of cause-specific mortality based on available data should be viewed with caution. There is consensus that schistosomiasis-specific mortality is related to advanced disease, which occurs in a small percentage of infected individuals and needs several decades of infection to develop. Consequently, the years of life lost to schistosomiasis probably represent a very small proportion of the overall DALYs lost to schistosomiasis compared with the years lost to disability.7,43 However, the true effect of schistosomiasis-related mortality is unknown.

Our interpretation contrasts with those of two previous extended reviews of schistosome-associated morbidity that have suggested only a minimum effect (or importance) for schistosomiasis infection on public health.15,16 The difference in our respective interpretations of the morbidity data is based on differences in study methods used for estimation of effect size. Whereas early studies tallied or voted the reported results of many population-based studies (on the basis of observed statistical significance of each individual study’s outcomes), we obtained aggregate estimates of parasite effects on the basis of combined results of multiple studies. In retrospect, the tally approach did not identify several moderate but significant schistosomiasis-associated deficits because individual studies were often underpowered to reliably detect the significance of group-wise differences in morbidity that in this study have proven significantly related to infection.43

There are limitations to our study. The use of observational studies and the inclusion of select subpopulation surveys (eg, school-age children) allows possible confounding effects on the observed results, thus obscuring the assessment of attributable risk due to schistosomiasis. Schistosomiasis is inevitably associated with other potential causes for morbidity and disease, especially with restricted access to safe water supplies and with co-infection by other parasites. Several studies have reported that age, sex, socioeconomic status, and diet can significantly modify the risk for schistosomiasis-associated morbidity. In nine of ten
individual surveys that adjusted for some or all of these cofactors, the effect of schistosomiasis on measured disability outcomes has remained significant. However, details of these potentially modifying factors were not available in most of the studies included in our analysis, and so adjustment was not attempted in the estimation of our summary statistics.

In view of our study objectives, we thought that it was important to include the broadest range of data (ie, over many periods and from multiple locations) from field-based studies to develop a more general, inclusive assessment of schistosome infection-related disability. There are few randomised controlled trials for schistosomiasis in which confounding effects were likely to be balanced between study groups, and these studies often reported only short-term follow-up of a single round of treatment. Whether a single treatment, as studied in randomised controlled trials, will effectively reverse all forms of chronic schistosome-related morbidity in 6–12 months is unknown. Thus the conclusion could be that schistosomiasis was not associated with certain morbidities that did not respond to single therapy. For this reason, we thought comparison of results of observational and interventional studies was important with respect to the types of morbidities associated with schistosomiasis. Although our estimates cannot be regarded as truly definitive, they are based on extensive study of diverse infected populations, and we were reassured that the observed trends for schistosomiasis-related morbidity were generally the same for both observational and the nominally more balanced randomised controlled trials.

Generalised mean differences and odds ratios in our study will probably not apply to all areas, depending in part on the distribution and intensity of schistosomiasis and the underlying health of the local population.44 Previous research has shown that nutritionally at-risk subpopulations are more likely to manifest signs of undernutrition or anaemia with schistosomiasis infection than are those with adequate protein-calorie and micronutrient intake.45 For example, populations with high iron intake, such as those in southern Africa, might not manifest anaemia with schistosomiasis, whereas those in Sahel areas (desert) can have profound undernutrition or anaemia with schistosomiasis.46

We noted that across all studies reporting haemoglobin concentrations, the weighted mean haemoglobin was reduced by 4 g/L in the presence of detectable schistosome infection. The net result was that more than half of reported schistosome-infected individuals were anaemic by WHO standards (haemoglobin concentration <120 g/L). Such levels of anaemia have been associated with 3–33% reductions in field productivity, 60% reductions in peak exercise capacity, and 8–20% reductions in peak work endurance.47 These objective effects, which have been validated for other disease states,48 combined with the subjective effects of schistosomiasis-related symptoms and other performance deficits, undoubtedly contribute to disease-specific disability from schistosome infection.

Frequencies of schistosomiasis-attributable anaemia—eg, undernutrition and disabling symptoms within affected populations—might be derived from the odds ratios and standardised mean differences in our quantitative review. These estimates, combined with knowledge of the general disability estimates for anaemia,49–53 undernutrition,45–49 diarrhoea,49 and chronic abdominal or pelvic pain,54 led us to calculate that, even after adaptation, the global cost to a Schistosoma-infected individual is at least 2–15% chronic disability, dependent on the reference point. Our assessment is consistent with the disability estimates of early reviews of soil-transmitted helminthiasis.55–58 Furthermore, preliminary patient preference studies of support have estimated up to 15% disability in S mansoni infection,59 although these estimates need to be confirmed by further testing. We take our derived estimates as working approximations of schistosomiasis-related disability values, which can be used to regauge the probable cost-effectiveness of treatment and preventive measures.60 We recognise that formal research on quality-of-life outcomes in schistosomiasis and other helminthiasis is clearly needed to provide greater precision in these assessments.

Intensity of infection is an important aspect of all helminthiasis. Since these multicellular parasites do not reproduce within the human body, burden of the infection is determined by the frequency of an individual’s exposure to infection.61 Disease formation is related, in part, to the number of parasite eggs that become trapped in host tissues. In our analysis, we identified a significant association between heavy (or high-intensity) infection and risk for diarrhoea, anaemia, and calorie undernutrition. Previous work has established the association between higher individual egg output and risk for hepatosplenomegaly with S mansoni and S japonicum,62–64 and for structural urinary tract disease with S haematobium.65 However, the correlation is not perfect. Emerging evidence indicates that immune regulation of inflammatory response, independent of infectious burden, is associated with risk for some forms of morbidity due to schistosomiasis infection.66 Therefore, light infection should not be overlooked as a cause for disability, and its prevention should be integrated as part of the design for population-based schistosomiasis control programmes.

The disability estimates derived from our study are between four and 30 times greater than the age-specific disability weights attributed to schistosomiasis in the Global Burden of Disease Project in 1996.7 The disability estimates derived from our study are between four and 30 times greater than the age-specific disability weights attributed to schistosomiasis in the Global Burden of Disease Project in 1996.70 Unfortunately, the Global Burden of Disease analysis separated specific causes of disease, such as helminthic infection, from their morbidities, such as anaemia and undernutrition.71 The result has been to de-emphasise the disability caused by chronic helminthic infections,
and to devalue the expected effect of cause-specific treatment, which has resulted in a substantial lowering of priorities for helminth parasite control. The approach of syndromic treatment of anaemia or undernutrition will not necessarily succeed without addressing significant local causative factors. Nutritional supplementation is not effective in overcoming the effects of chronic helminthic infection, and combined treatment and nutritional supplementation strategies will probably be needed to prevent or effectively reverse parasite-associated morbidity in this respect.27,28

Even small schistosomiasis-associated deficits could represent a substantial public-health burden in view of the large numbers of individuals with schistosomiasis (about 200 million), and the longlasting chronicity of infection.16 Our new estimates of schistosomiasis-related disability, combined with recent information on the global prevalence of schistosome infection, could provide a more useful working definition of the effect of schistosome infection on the world burden of disease.

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

Acknowledgments
We thank the many reference librarians of the Cleveland Health Sciences Library and the Allen Memorial Library for their assistance in retrieval of the articles and books surveyed for this study. This project was supported by US National Institutes of Health (NIH) Research Grant # R01 TW001543 funded by the Fogarty International Center, National Institute of Allergy and Infectious Diseases, and the National Institute of Environmental Health Sciences, and by NIH Research Grant # AI 45473 of the International Collaborations in Infectious Diseases Research program of the National Institute of Allergy and Infectious Diseases.

References
18 van der Werf MJ, De Vlas SJ. Morbidity and infection with schistosomes or soil-transmitted helminths. Rotterdam: Erasmus University, 2001: 1–103.
Non-HLA transplantation immunity revealed by lymphocytotoxic antibodies

Gerhard Opelz for the Collaborative Transplant Study*

Summary

Background The presence of panel-reactive antibodies (PRA) against HLA antigens before transplantation is associated with early rejection of kidney grafts from cadaver donors. Transplants from HLA-identical sibling donors do not provide a target for antibodies to HLA antigens and should therefore not be affected by PRA.

Methods Data from the Collaborative Transplant Study were used to examine the influence of PRA on graft survival. Uncensored graft survival and death-censored graft survival were calculated, and the data were analysed by multivariate Cox’s regression methods.

Findings Among recipients of HLA-identical sibling transplants, 3001 patients with no PRA had significantly higher 10-year graft survival (72·4% [SE 1·1]) than 803 patients with 1–50% PRA (63·3% [2·5]; p=0·0006) or 244 patients with more than 50% PRA (55·5% [4·0]; p<0·0001). The effect of PRA became apparent after the first post-transplant year and was, therefore, strikingly different from the early steep decline in graft survival during the first year associated with PRA in recipients of cadaver kidneys. We could not discern whether graft loss was a direct effect of non-HLA humoral sensitisation or whether PRA served as an indicator of heightened immunity against non-HLA transplantation antigens.

Interpretation PRA reactivity is strongly associated with long-term graft loss in kidney transplants from HLA-identical sibling donors.

Relevance to practice Our findings suggest that non-HLA immunity has a much stronger role in clinical transplantation than previously thought. In contrast to immunity against HLA mediated by antibodies present before transplantation, which leads to early acute graft rejection, non-HLA immunity is associated with chronic graft loss. The possibility of identifying recipients at increased risk of late graft loss before transplantation could be used to devise specific immunosuppressive strategies for these patients.

Introduction

Terasaki and colleagues first reported 30 years ago that lymphocytotoxic antibodies before transplantation were at increased risk of graft rejection. Their finding was confirmed in many subsequent studies, and now laboratories use commercial kits with frozen serum samples contain antibodies against lymphocytes from 56 random donors. There is no binding convention about the size of the cell panel used for testing, although most laboratories use commercial kits with frozen lymphocytes from 56 random donors. There is general acceptance that the results of the test system are suboptimally reproducible. Nevertheless, risk of rejection appears to rise as serum reactivity against the random lymphocyte test panel increases. PRA-positive serum samples contain antibodies against HLA antigens on lymphocytes, and graft survival in preimmunised recipients is assumed to be lower as the result of insufficient sensitivity in the pretransplant lymphocytotoxic cross-match test against donor lymphocytes. Much effort has therefore been spent on increasing the sensitivity of the cross-match assay so that weak anti-HLA sensitisation can be detected, and the use of new techniques for pretransplant antibody testing based on highly sensitive, strictly HLA-specific ELISAs has been encouraged. Patients cannot form antibodies against their own HLA antigens; therefore they cannot form anti-HLA antibodies to lymphocytes of an HLA-identical sibling donor. In distinction from genetically identical twins, who share all genes and therefore do not require immunosuppression when tissues are transplanted between them, the common definition of HLA-identical siblings is that they are matched for both HLA chromosomes but mismatched at other chromosomes; thus, they can be of different sex, eye colour, and so on, as well as age. Since HLA chromosomes are inherited according to mendelian rules, the likelihood that two siblings will inherit identical HLA chromosomes from their parents is 25%. Although many other factors influence the outcome of kidney transplantation, transplants from HLA-identical sibling donors are recognised as a special category. They have significantly better success rates than transplants from HLA-mismatched donors, and they are the standard against which the results of transplantation from other donor sources are compared. However, since rejection of HLA-identical sibling grafts commonly occurs if no immunosuppression is given, these recipients are
treated with immunosuppressive drugs, albeit at lower doses than recipients of grafts from cadaver donors. This need for immunosuppressive treatment shows that, aside from HLA, there must be other antigen systems that can cause transplant rejection. HLA-identical sibling transplants do not provide a target for anti-HLA, and PRA reactivity before transplantation should therefore not influence their success rate.

We studied the influence of pretransplant PRA status on the long-term outcome of kidney grafts from HLA-identical sibling donors.

**Methods**

**Patients**

Kidney transplants reported by 245 centres to the international Collaborative Transplant Study were analysed. Transplants were carried out between 1982, the year the study was initiated, and 2002. The centres included in this analysis provided written assurance of compliance with local ethical and consent guidelines and of patients’ consent for the use of data for scientific analysis. When consent from patients could not be obtained, care was taken to ensure that all data processing was carried out anonymously.

**Procedures**

A transplant was classed as HLA-identical if recipient and donor were reported to have identical HLA A, B, and DR antigens. 3681 first transplants and 367 retransplants from HLA-identical sibling donors formed the main study population for this analysis. 160,486 cadaver-donor transplants were analysed for comparison. HLA typing and testing for PRA was done at participating laboratories and reported to the study centre. The PRA reactivity of the last pretransplant serum sample was analysed. For 16 patients (five with no PRA, five with 1–50% PRA, and six with more than 50% PRA), positive pretransplant lymphocytotoxic cross-matches against lymphocytes of the kidney donor were reported, whereas all other patients had negative cross-match results. All 16 patients with positive anti-donor cross-matches also had positive cross-match results against their own (autologous) lymphocytes, indicating that the positive cross-matches were the result of autoantibodies.

Clinical follow-up was recorded at 3 months, 6 months, and 12 months, and yearly thereafter.

**Statistical analysis**

Graft survival was calculated by Kaplan-Meier analysis. Statistical significance was assessed with the log rank test. Transplant outcome in relation to pretransplant antibody status was analysed by weighted regression analysis, in which the dependent variable was survival proportion, the independent variable was the degree of antibody reactivity, and the weight factor was the number of patients who had the given antibody reactivity. In the analysis of graft survival, all graft failures irrespective of cause (including death of the patient) were counted as failures. In the analysis of functional graft survival, deaths were censored. The Kruskal-Wallis and Mantel-Haenszel tests were used to estimate the statistical association between PRA reactivity and number of pretransplant blood transfusions, recipient’s sex, and proportion of retransplants. Immunosuppressive treatment (analysed by intention to treat) of patients with transplants from HLA-identical sibling donors included ciclosporin in

---

**Figure 1:** 1-year graft survival analysis of kidney transplants from cadaver donors or HLA-identical sibling donors in relation to PRA

---

<table>
<thead>
<tr>
<th></th>
<th>No PRA</th>
<th>1–50% PRA</th>
<th>&gt;50% PRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grafts surviving (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time (months)</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Cadaver kidney transplants</td>
<td>100</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>No PRA</td>
<td>115</td>
<td>103</td>
<td>93</td>
</tr>
<tr>
<td>1–50% PRA</td>
<td>131</td>
<td>115</td>
<td>99</td>
</tr>
<tr>
<td>&gt;50% PRA</td>
<td>102</td>
<td>95</td>
<td>92</td>
</tr>
<tr>
<td>HLA-identical sibling transplants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grafts surviving (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time (months)</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>No PRA</td>
<td>3001</td>
<td>2914</td>
<td>2894</td>
</tr>
<tr>
<td>1–50% PRA</td>
<td>203</td>
<td>177</td>
<td>176</td>
</tr>
<tr>
<td>&gt;50% PRA</td>
<td>244</td>
<td>235</td>
<td>229</td>
</tr>
</tbody>
</table>
2712 (67%) of the patients, tacrolimus in 122 (3%), and regimens without calcineurin inhibitors in 1214 (30%); there were no significant differences in success rates between the groups treated with these regimens.

Multivariate Cox’s regression analysis was used to ascertain the effect of the covariates: transplant number (first or retransplant); year of transplantation; immunosuppressive regimen (calcineurin inhibitor or not); age, sex, and race of recipient and donor; original disease leading to endstage renal failure; number of pretransplant blood transfusions; and geographical location of the transplant centre (continent). The software packages SPSS (version 11.5) and SAS (version 8.2) were used.

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

Results
The differential effect of PRA on survival of transplants from cadaver donors or HLA-identical sibling donors during the first year after transplantation is shown in figure 1. As expected, a significant effect of PRA on graft survival was evident in cadaver transplants (p<0·0001), but no significant effect was noted in transplants from HLA-identical sibling donors (p=0·0831). PRA reactivity in cadaver-transplant recipients was associated with immunological graft loss rather than death of the patient. The calculation of death-censored functional graft survival, which provides an approximation of the rate of immunological graft rejection, gave almost the same result as that shown in figure 1 (p<0·0001). By contrast, PRA in HLA-identical sibling transplants affected neither graft nor functional survival.

When the analysis was extended to 10 years of follow-up, however, a significant effect of PRA on graft survival became apparent in the analysis of HLA-identical sibling transplants (figure 2). At 10 years, the proportion of grafts surviving was 72.4% (SE 1·1) in patients with no PRA, 63·3% (2·5) in those with 1–50% PRA, and 55·5% (4·0) in those with more than 50% PRA (regression p<0·0001). The result was significant in both first transplants (p=0·0002) and retransplants (p=0·0001), and subset analysis showed significant associations for the transplant periods 1982–90 (p=0·0003) and 1991–2002 (p=0·005). Analysis of death-censored functional survival showed that the PRA effect was due to graft loss and not death of patients. The 10-year functional graft survival proportions were 82·5% (1·0), 75·3% (2·2), and 63·1% (4·0), for patients with no PRA, 1–50% PRA, and more than 50% PRA, respectively (p<0·0001; figure 3). The proportions of patients surviving at 10 years were 86·2% (0·9), 81·7% (2·0), and 81·9% (3·1) respectively (p=0·0266; data not shown). Multivariate Cox’s regression analysis showed that, compared with patients who were PRA-negative before transplantation, recipients with 1–50% PRA had a significantly increased risk of graft loss (relative risk 1·29 [95% CI 1·09–1·53], p=0·0033) and the risk was even higher in patients with more than 50% PRA (1·87 [1·47–2·37], p<0·0001). In the analysis of functional
graft survival, the risk was significantly increased for patients with 1–50% PRA (1·26 [1·02–1·57], p=0·0282), and the result was highly significant for those with more than 50% PRA (2·3 [1·7–2·9], p<0·0001).

The long-term evolution of the success of cadaver and HLA-identical sibling transplants differed substantially. PRA affected the survival proportion of cadaver-donor transplants primarily in the first few months after transplantation (figure 4). By contrast, transplants from HLA-identical siblings were not affected during the first year and the influence of PRA developed continuously during the 10-year follow-up, which suggests that different mechanisms were involved (figure 2).

Owing to the intrafamilial pathway of inheritance of HLA chromosomes, the definition of HLA-matched transplants is not difficult in sibling settings. However, owing to the complexity of the HLA system and imperfect characterisation of HLA alleles in clinical typing for renal transplantation, identification of perfectly matched transplants in a registry series of cadaver kidney donors is impossible. At the level of allelic definition of HLA antigens, almost all cadaver transplants included in this analysis must be deemed HLA mismatched. With the available data, an analysis of perfectly matched cadaver transplants was therefore not possible. Nevertheless, one would assume that the non-HLA effect of PRA described in HLA-identical sibling grafts must also have a role in HLA-mismatched cadaver transplants, in addition to the anti-HLA effect shown in figure 1. We addressed this issue by doing a separate analysis of graft survival for the period after the first post-transplant year, when grafts subject to early anti-HLA immunity had already been rejected (figure 1).

Indeed, when the graft-survival analysis was restarted at 100% at 1 year after transplantation, a long-term effect of PRA became apparent in cadaver transplants (figure 5). We ruled out the possibility that the degree of HLA matching, which is known to affect long-term survival of cadaver transplants, might have caused the separation of the survival curves. The mean number of HLA A, B,
than HLA, controlled by genes on other chromosomes, determine minor histocompatibility antigens but serve as an indicator of immunogenic epitopes able to elicit a T-cell response against transplanted cells or tissue. Influential in marrow/stem-cell transplantation but until now no proven influence on organ transplant survival.

Minor histocompatibility antigens

These antigens, less well defined than HLA, controlled by genes on other chromosomes, determine immunogenic epitopes able to elicit a T-cell response against mismatched HLA antigens on donor tissue. These antibody reactions were probably weak, because they went undetected in the pretransplant cross-match assay and did not lead to immediate graft failure from hyperacute rejection. Irreversible rejection, however, occurred within a few weeks or months as shown by the early decline of graft survival curves in presensitised patients (figure 1). The survival curves for HLA-identical-sibling transplants, by contrast, declined very slowly, indicating a much later immunological event (figure 2). This differentiation in early and late graft loss also argues against the unlikely possibility that graft failure in the sibling group might have been due to incorrect HLA typing; had HLA mismatches been present among the HLA-identical siblings, early graft losses as a result of antibody-mediated rejection would have been expected. All sibling donors and recipients analysed in this study were identical at the HLA A, B, and DR loci, the histocompatibility loci established to be influential in clinical kidney transplantation. Although the presence of incompatibilities at other loci within the HLA region (eg, DQ, DP) cannot be excluded, the likelihood that such incompatibilities existed can be estimated at less than 3%. Thus, the chance that the results of this study were influenced to an important extent is very small.

The targets for antibodies causing late rejections could be so-called minor histocompatibility antigens, which are not coded for in the HLA genetic region and have been shown to influence the outcome of bone-marrow transplantation from HLA-identical sibling donors.20 Antibodies against these antigens might not lead to acute rejection of kidney grafts but to protracted chronic rejection. Since decreased graft survival was associated with PRA reactivity, two possible hypotheses are that antibodies to minor histocompatibility antigens occur frequently together with anti-HLA, or that cross-reactivity of anti-HLA with epitopes on minor histocompatibility antigens might induce late graft rejection.

Another possibility is that PRA reactivity does not signal the existence of antibodies against minor histocompatibility antigens but serves as an indicator of a generally increased state of immune responsiveness due to allogeneic preimmunisation, and that

### Table: Association of preformed PRA with various characteristics of patients

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Preformed PRA</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>None (n=3001)</td>
<td>1–50% (n=803)</td>
</tr>
<tr>
<td>Proportion female</td>
<td>1014 (34%)</td>
<td>363 (45%)</td>
</tr>
<tr>
<td>Proportion with retransplant</td>
<td>159 (5%)</td>
<td>112 (14%)</td>
</tr>
<tr>
<td>Mean (SE) pretransplant blood transfusions</td>
<td>3·47 (0·15)</td>
<td>6·01 (0·43)</td>
</tr>
</tbody>
</table>

Discussion

This study showed a highly significant association between the presence of lymphocytotoxic antibodies before transplantation and the outcome of kidney grafts from cadaver donors and HLA-identical sibling donors. Such an association was previously noted in cadaver kidney transplantation, in which it has been attributed to unrecognised antibody reactivity against mismatched HLA antigens. However, the finding of a similar association for transplantation of organs between HLA-identical siblings was unexpected because the HLA-antigen profile of recipient and donor is identical. In this special setting, no effect would be expected because PRA reactivity is generally believed to be directed against HLA antigens. The presence of weakly reactive antibodies to HLA antigens that went undetected in the pretransplant cross-match assay cannot explain the lower graft success rate observed in recipients of transplants from HLA-identical siblings. An association of graft rejection with the presence of lymphocytotoxic antibodies before transplantation has been noted previously in bone-marrow transplantation from HLA-identical sibling donors.20

Apart from the main issue that antibodies to HLA antigens should not influence the outcome of transplants between HLA-identical siblings, the differing evolution of the PRA effect on graft survival suggests that cadaver and HLA-identical-sibling transplants were affected by different types of antibodies or different mechanisms of immunological rejection. The result for cadaver transplants is compatible with the concept that HLA antibodies present in the recipient’s circulation at the time of transplantation reacted with mismatched HLA antigens on donor tissue. These antibody reactions were probably weak, because they went undetected in the pretransplant cross-match assay and did not lead to immediate graft failure from hyperacute rejection. Irreversible rejection, however, occurred within a few weeks or months as shown by the early decline of graft survival curves in presensitised patients (figure 1). The survival curves for HLA-identical-sibling transplants, by contrast, declined very slowly, indicating a much later immunological event (figure 2). This differentiation in early and late graft loss also argues against the unlikely possibility that graft failure in the sibling group might have been due to incorrect HLA typing; had HLA mismatches been present among the HLA-identical siblings, early graft losses as a result of antibody-mediated rejection would have been expected. All sibling donors and recipients analysed in this study were identical at the HLA A, B, and DR loci, the histocompatibility loci established to be influential in clinical kidney transplantation. Although the presence of incompatibilities at other loci within the HLA region (eg, DQ, DP) cannot be excluded, the likelihood that such incompatibilities existed can be estimated at less than 3%. Thus, the chance that the results of this study were influenced to an important extent is very small.

The targets for antibodies causing late rejections could be so-called minor histocompatibility antigens, which are not coded for in the HLA genetic region and have been shown to influence the outcome of bone-marrow transplantation.21 Antibodies against these antigens might not lead to acute rejection of kidney grafts but to protracted chronic rejection. Since decreased graft survival was associated with PRA reactivity, two possible hypotheses are that antibodies to minor histocompatibility antigens occur frequently together with anti-HLA, or that cross-reactivity of anti-HLA with epitopes on minor histocompatibility antigens might induce late graft rejection.

Another possibility is that PRA reactivity does not signal the existence of antibodies against minor histocompatibility antigens but serves as an indicator of a generally increased state of immune responsiveness due to allogeneic preimmunisation, and that
incompatibilities for minor histocompatibility antigens might exert a strong but protracted graft-damaging influence in patients with heightened immunity. Aside from acute rejections of cadaver kidneys mediated by antibodies to HLA antigens, delayed rejections would not be directly mediated by the antibodies detected in pretransplant serum, thus explaining the differential time profiles of rejection due to “major” HLA incompatibilities (cadaver transplants with HLA mismatches) or “minor” non-HLA incompatibilities (a proportion of HLA-identical-sibling and cadaver transplants). In cadaver-transplant recipients with heightened immunological reactivity against both major and minor incompatibilities, the grafts would fail early. The finding that a substantial proportion of sibling transplants failed even in the absence of any detectable antibody reactivity (figure 2) shows that the process described here is the cause of only some, not all, graft failures.

The results presented here have important fundamental and practical implications. The study has shown that non-HLA immunity contributes substantially to long-term kidney-transplant failure. When the long-term results for kidney recipients with PRA were examined over 10 years of follow-up, the influence of non-HLA-directed immunity was of similar magnitude to that of antibodies against HLA (figures 2 and 4). The study shows the importance of characterising the non-HLA antigens that bring about graft loss. However, because it was based on registry data, serum and cells from recipients and donors are not available for in-vitro studies. Prospective collection of biological material from donors and recipients of HLA-identical sibling transplants will be needed for further progress on this issue. If future research is successful, many of the late graft failures attributable to non-HLA effects might be avoidable. For the time being, the fact that HLA-identical siblings at increased risk of late graft loss can be identified before transplantation could be used to devise specific immunosuppressive strategies for these patients.

Another interesting point arising from these results is the implication that the current worldwide trend towards replacement of pretransplant antibody testing in the complement-dependent lymphocytotoxicity assay with strictly HLA-specific ELISA testing might be counterproductive. ELISA assays have the advantage of better reproducibility than the complement-dependent cytotoxicity method, and ELISA testing at increased sensitivity with strict anti-HLA specificity is widely expected to lead to better clinical transplant results.6,10 Direct comparison of transplant outcome in relation to pretransplant antibody testing with lymphocytotoxicity or HLA-specific ELISA, however, did not show a convincing advantage for either method and provided evidence that some serum samples contained ELISA-non-reactive antibodies that were associated with kidney graft rejection.22 The results of this study suggest that an important part of the antibody range might be missed if pretransplant serum testing were limited to HLA-specific ELISA. Although the introduction of ELISA techniques has greatly improved ability to characterise antibodies to HLA, the additional non-HLA reactions detected in the classic lymphocytotoxicity assay seem to be clinically highly relevant. Although the exact immunological mechanisms involved remain to be discovered, caution should be used in modifying pretransplant screening procedures without further knowledge about the effect of lymphocytotoxic antibodies on long-term outcome. For the time being, use of both ELISA and cytotoxicity assays in parallel for pretransplant testing seems wise to allow a separation of anti-HLA from anti-non-HLA activities. Our results also suggest that the introduction of highly sensitive, strictly HLA-specific ELISA-based pretransplant cross-match assays21 has only a limited potential for improving transplant outcome. Although these tests can be expected to lower further the incidence of HLA-antibody-mediated rejections, they will not affect the substantial rate of late rejections attributable to immunity unrelated to HLA.

Contributors
Gerhard Opelz initiated and coordinated the study and wrote the report.

Participating centres
Argentina—Buenos Aires (3); Mar del Plata; Rosario. Australia—Adelaide; Brisbane; Heidelberg; Hobart; Melbourne; Newcastle; Perth; Sydney (7). Austria—Graz; Innsbruck; Linz; Vienna. Belgium—Leuven (2); Liege. Brazil—Belo Horizonte (2); Maceio; Pato Branco; Porto Alegre (2); Ribeirao Preto; Rio de Janeiro (2); Sao Paulo (4). Canada—Quebec; Toronto; Vancouver; Winnipeg; Chile—Santiago (2); Valdivia. Colombia—Medellin. Croatia—Rijeka; Zagreb. Egypt—Cairo (2). India—Mumbai; New Delhi; Jammu; Kuala Lumpur; Lahore; Karachi. Ireland—Helsinki. France—Lille; Lyon; Nancy; Nantes; Paris; Poitiers; Reims; Rennes; St Etienne; Toulouse. Germany—Aachen; Augsburg; Berlin (2); Bochum; Bonn; Bremen; Cologne (2); Dusseldorf; Erlangen; Frankfurt; Freiburg; Fulda; Giessen; Göttingen; Halle; Hann-Muenenden; Hannover; Heidelberg; Homburg-Saar; Jena (2); Kaiserslautern; Kiel; Lübeck; Mainz; Mannheim; Marburg; Münster; Munich; Regensburg; Rostock; Stuttgart; Tübingen; Ulm; Würzburg. Greece—Thessaloniki. Hong Kong—Hong Kong (8). Hungary—Budapest; Debrecen; Pecs; Szeged. India—New Delhi. Iran—Shiraz; Tehran. Ireland—Dublin. Israel—Petach Tikva. Italy—Brescia; Cagliari; Florence; Genova (2); Milan (4); Padova (2); Treviso; Turin; Udine; Varese; Verona. Japan—Kyöto. Korea—Seoul. Lithuania—Kaunas; Vilnius. Mexico—Guadalajara (2); Mexico City. Netherlands—Nijmegen. New Zealand—Auckland; Christchurch; Hamilton; Wellington. Pakistan—Islamabad; Karachi. Peru—Lima. Philippines—Manila. Poland—Katowice. Russia—Moscow; Slovakia—Martin. Slovenia—Ljubljana. South Africa—Cape Town (5). Spain—Badalona; Barcelona (6); Madrid (5); Oviedo; Pamplona; Santander; Valencia (3). Sweden—Gothenburg. Switzerland—Basel; Bern; Geneva; Lausanne; St Gallen; Zurich. Turkey—Ankara (2); Antalya; Izmir (2). UK—Aberdeen; Belfast; Birmingham; Brighton; Bristol; Cambridge; Cardiff; Carshalton; Coventry; Dresden; Dundee; Edinburgh; Glasgow (2); Kent; Leeds; Leicester; Liverpool; London (9); Manchester; Newcastle upon Tyne; Nottingham; Oxford; Plymouth; Portsmouth; Sheffield. USA—Cincinnati; Dallas (3); Grand Rapids; Kansas City (3); New Orleans (2); Omaha; Orlando; Portland (2); Shreveport; Stanford; Valhalla. Venezuela—Caracas; Maracaibo.

Conflict of interest statement
I declare that I have no conflict of interest.
Acknowledgments
I thank staff members at 245 transplant centres who provided data for this study for their generous support, and Bernd Doehler for valuable assistance with statistical analysis.

References
Abdominal aortic aneurysms cause 1·3% of all deaths among men aged 65–85 years in developed countries. These aneurysms are typically asymptomatic until the catastrophic event of rupture. Repair of large or symptomatic aneurysms by open surgery or endovascular repair is recommended, whereas repair of small abdominal aortic aneurysms does not provide a significant benefit. Abdominal aortic aneurysm is linked to the degradation of the elastic media of the atheromatous aorta. An inflammatory cell infiltrate, neovascularisation, and production and activation of various proteases and cytokines contribute to the development of this disorder, although the underlying mechanisms are unknown. In this Seminar, we aim to provide an updated review of the pathophysiology, current and new diagnostic procedures, assessment, and treatment of abdominal aortic aneurysm to provide family practitioners with a working knowledge of this disorder.

Abdominal aortic aneurysms are a substantial burden on health care in more developed countries, occurring mostly among men older than 65 years of age. The disorder is the thirteenth leading cause of death in the USA.1 Although some patients have vague symptoms, such as back pain or abdominal pain, most abdominal aneurysms are asymptomatic until rupture, which leads to death in 65% of patients.2 An increased awareness of the characteristics of abdominal aortic aneurysm by first-contact practitioners might reduce the risk of a fatal outcome with this disorder. Here, we review the key aspects of this disorder, including epidemiology, pathogenesis, diagnosis, and treatment.

Terminology

Aneurysm derives from the Greek ανευρυσμα (aneurysma), meaning widening, and can be defined as a permanent and irreversible localised dilatation of a vessel. Although an aneurysm occurring in any portion of the infradiaphragmatic aorta could be termed an abdominal aortic aneurysm, common practice restricts this definition to an aneurysm of the infrarenal aorta. Aneurysms involving the renal ostia (intrarenal, suprarenal aorta) are also included under this term. The normal diameter of the abdominal aorta varies with age, sex, and bodyweight,3 and decreases progressively from its entry into the abdominal cavity to the iliac bifurcation. In elderly men, the infrarenal abdominal aortic diameter is between 15 mm and 24 mm.4 McGregor and colleagues5 proposed the definition of an abdominal aortic aneurysm as an aorta with an infrarenal diameter greater than 30 mm. In 1991, the Society for Vascular Surgery and the International Society for Cardiovascular Surgery Ad Hoc Committee on Standards in Reporting proposed a criterion that the infrarenal diameter should be 1·5 times the expected normal diameter.6 There is no definite consensus on the definition of abdominal aortic aneurysm; however, the disorder is conventionally diagnosed if the aortic diameter is 30 mm or more. This dilatation affects the three layers of the vascular tunic; otherwise, the dilatation is called a pseudoaneurysm. Most aneurysms are fusiform since the whole circumference of the artery is affected, whereas an aneurysm that includes only a part of the circumference is termed saccular. An inflammatory aneurysm is characterised by extensive perianeurysmal and retroperitoneal fibrosis and dense adhesions to adjacent abdominal organs.

Epidemiology

The incidence of abdominal aortic aneurysms has increased during the past two decades, due in part to the ageing of the population, the rise in the number of smokers, the introduction of screening programmes, and improved diagnostic tools. Rupture of these aneurysms causes about 8000 deaths per year in the UK and roughly 15 000 per year in the USA.1,2 The disorder is more common in men than in women, with prevalence rates estimated at between 1·3% and 8·9% in men and between 1·0% and 2·2% in women.4–12 However, since smoking is one of the most important risk factors for abdominal aortic aneurysm13 and the number of female smokers is rising,14 the sex ratios for prevalence of the disorder will probably change in the future.15–17 Most aneurysms discovered by screening are of small size and do not need immediate surgical repair.15–17 However, they can become enlarged with time at a mean rate that is initially slow and then increases exponentially.18 In general, the risk of rupture increases as the diameter of the aneurysm enlarges.14,16,19 The overall mortality rate for patients with ruptured abdominal aortic aneurysms is between 65% and 85%,20 and about half of deaths attributed to rupture occur before the patient reaches the surgical room.21,22

Search strategy and selection criteria

The primary source of references included MEDLINE searches for recent literature with many keywords for both clinical and basic research topics. We largely selected publications in the past 5 years, but did not exclude commonly referenced and highly regarded older publications. We also reviewed books and review articles pertaining to abdominal aortic aneurysm.
Aetiology and risk factors

There are many causes of aneurysmal dilatation, but few abdominal aortic aneurysms are the direct consequence of specific causes such as trauma, acute infection (brucellosis, salmonellosis), chronic infection (tuberculosis), inflammatory diseases (Behçet and Takayasu disease), and connective tissue disorders (Marfan Syndrome, Ehlers-Danlos type IV). Thus, most abdominal aortic aneurysms are called non-specific. Moreover, because this disorder is invariably associated with severe atherosclerotic damage of the aortic wall, it has been traditionally regarded as a consequence of atherosclerosis. This conventional view has been increasingly challenged in recent years. Clinical and basic research studies indicate that aneurysms arise through pathogenic mechanisms that differ, at least in part, from those responsible for athero-occlusive disease. Much published work lends support to this concept. Defawe and colleagues showed that two physiological inhibitors of proteases (TIMP-2 and PAI-1) were expressed less in abdominal aortic aneurysms than in athero-occlusive disease, suggesting a significant role for protease inhibitors during the divergent evolution of the initial atherosclerotic plaque towards either abdominal aortic aneurysm or athero-occlusive disease. Moreover, since not all patients with atherosclerosis develop an abdominal aortic aneurysm, even if atherosclerosis does have a role in the pathogenesis of the disorder, additional factors are probably involved in aneurysm development.

There is a strong clinical association between tobacco smoking and aneurysm development. The prevalence of abdominal aortic aneurysms in tobacco smokers is more than four times that in life-long non-smokers. A report that compared relative risks for smokers is more than four times that in life-long non-smokers is three-fold greater than the risk for developing abdominal aortic aneurysms. Majumder and co-workers, in a segregation analysis based on 91 probands and including 13 familial cases, showed that the most likely genetic model was an autosomal locus with a recessive inherited gene for the disorder. An analysis undertaken in 313 pedigrees, by Verloes and colleagues, provided evidence for a single autosomal dominant inheritance. A multinational study identified 233 families with 653 affected members; the inheritance mode was autosomal recessive in 72% of families and autosomal dominant in 25% of families. Linkage between aortic aneurysm growth and a 4G/5G polymorphism in the plasminogen activator inhibitor (PAI-1) promoter has been reported. Several candidate genes are present in this region of chromosome 19, such as LDL receptor-related protein 3 (LRP3), which is particularly relevant since conditional knockout mice for LRP1, another member of this gene family, developed atherosclerosis and arterial aneurysm.

As discussed by Powell, familial clustering of abdominal aortic aneurysms is probably not due to chance alone. An underlying cause could be particular genetic background, as mentioned before, but probably in conjunction with environmental factors. In comparison, clear evidence was reported for the interaction between smoking and polymorphic variation in the nitric oxide synthase gene for the development of carotid artery disease. Familial clustering of abdominal aortic aneurysms could also result from exposure to common environmental factors, such as tobacco smoke. Parental smoking has been suggested to underlie familial clustering. Additionally, women who smoke tend to have infants of low birthweight who have a high risk of developing coronary diseases later in life.

Pathophysiology

The development of abdominal aortic aneurysms is clearly associated with alterations of the connective tissue in the aortic wall. Elastic fibres and fibrillar collagens are the main determinants of the mechanical properties of the aorta. Elastin and associated proteins form a network of elastic fibres responsible for the viscoelastic properties. Elastin is stabilised by cross-links between the molecules and can be degraded by specific proteases that display elastase activity. Elastic fibres associated with smooth muscle cells are most abundant in the media of the aortic wall. Collagen, in polymeric form, is also a significant component of the media and
the surrounding fibrous adventitia. Two specific types of fibrillar collagen (types I and III) provide tensile strength and help maintain the structural integrity of the vascular wall. Beside elastic and collagen fibres, proteoglycans are also implicated in the organisation of the aortic wall.50

One of the most important histological features of aneurysmal tissue is the fragmentation of the elastic fibres and a decreased concentration of elastin during aneurysm growth until the time of rupture.51–53 The loss of elastic fibres seems to be an early step in aneurysm formation.54 Although elastin fragmentation and medial attenuation are the most important characteristics of the wall of an aneurysm, the adventitial tissue, in which collagen is predominant, is responsible for the resistance of the aorta in the absence of medial elastin. According to Dobrin and Mrkvicka,54 collagen degradation is the ultimate cause of rupture. Increased collagen turnover has been reported in abdominal aortic aneurysms in human beings,55 suggesting the existence of a repair process as shown in animal models.56 An imbalance between collagen degradation and its synthesis could create the catabolic conditions that lead to rupture.

The alteration of elastin and collagen in the aortic wall is dependent on production of proteases by resident vascular wall cells (medial smooth muscle cells and adventitial fibroblasts) and by the cells of the lymphomonocytic infiltrate. These inflammatory cells in the media and adventitia come from the aortic blood and from a medial neovascularisation, which characterises abdominal aortic aneurysms.57–59 Leucocyte recruitment into the aortic wall is promoted by elastin degradation fragments as well as proinflammatory cytokines, chemokines, and prostaglandin derivates produced by both the resident mesenchymal cells and the inflammatory cells themselves.60–63 Immunity has been suggested to play a part in the development of abdominal aortic aneurysm.64,65 Elastic and collagen fibres are degraded by proteolytic enzymes mostly represented by matrix metalloproteinases (MMP) locally activated by either other MMP or by plasmin generated by plasminogen activators.66–79 The role of MMP and plasmin in the development of abdominal aortic aneurysms has been confirmed in animal models.70,76,77,80–84 The tissue inhibitors of matrix metalloproteinases (TIMP) are also increased in the wall of the aneurysm.85 However, the balance between proteases and antiproteases seems to be in favour of proteolysis.81,83,86 The importance of this imbalance in aneurysm development is reinforced by experimental studies in which the antiproteases are overexpressed or genetically inactivated.87–90

Besides rarefaction of its extracellular matrix, the elastic media also undergoes a reduction in the density of smooth muscle cells, which is regarded as a key event in the development of abdominal aortic aneurysms.91 Smooth muscle cells participate in vascular wall remodelling through localised expression of various extracellular matrix proteins as well as proteases and their inhibitors. Additionally, smooth muscle cells have a protective role against inflammation and proteolysis.92 In-vitro aortic smooth-muscle cells have been shown to produce less monocyte chemotactic protein-1 (MCP-1), a major inflammatory mediator in abdominal aortic aneurysms, under cyclic stretching than with static culture, which lends support to the notion of a protective paracrine function of smooth muscle cells.93 The development of abdominal aortic aneurysms is also associated with a mural thrombus in most patients. By contrast with arterial occlusive diseases, blood flow is maintained in aortic aneurysms resulting in a persistent remodelling activity of the components of the thrombus. Evolution of aneurysmal diameter has been reported to correlate with plasma markers of fibrin formation and degradation94 as well as with the circulating complex plasmin-α2-anti-plasmin95 potentially related to thrombus turnover. The role of an adherent thrombus in
Aneurysmal degeneration has also been investigated. Although the thrombus can substantially reduce aneurysmal wall stress, its increasing thickness leads to local hypoxia at the inner layer of the media, which can induce increased medial neovascularisation and inflammation. The implication of the thrombus in aneurysmal evolution in terms of a source of proteases has also been proposed after an initial report of an enrichment of MMP-9 (gelatinase B) in the thrombus. Furthermore, Fontaine and colleagues have provided evidence of polymorphonuclear neutrophils (PMN) trapping and storing MMP-9 within the aneurysmal thrombus. They also showed that plasminogen and its activator (u-PA) are present in the thrombus in the aneurysm wall, which might result in local generation of plasmin, an activator of MMP.

Methods of diagnosis

The examination for a pulsatile mass should be done by bimanual palpation of the supraumbilical area. Sensitivity of abdominal palpation for detection of abdominal aortic aneurysms increases with the diameter of the lesion: 61% for aneurysms 3.0–3.9 cm, 69% for those 4.0–4.9 cm, and 82% for those 5.0 cm and larger. The palpation sensitivity also depends inversely on the size of the abdominal waistline.

Abdominal standard radiography can incidentally be diagnostic, mainly in the transverse view, if calcifications are present in the aortic wall, which allows visualisation of dilatation. However, standard radiography is not the method of choice for the diagnosis of abdominal aortic aneurysms. Ultrasonography is the simplest and cheapest diagnostic procedure and can accurately measure the size of the aorta in longitudinal as well as in anteroposterior and transverse directions (figure 1) with an accuracy of 3 mm. Ultrasonography is largely used not only for the initial assessment and the follow-up

Figure 2: CT images of abdominal aortic aneurysms


Figure 3: MRI of an abdominal aortic aneurysm

A: 2D T1-weight post contrast MRI. B: Gadolinium-enhanced MRA in the same patient showing tortuous aorta and iliac arteries.
surveillance, but also for population screening. If the diameter of the aneurysm is such that surgical procedure is contemplated, CT is the next step to help determine which treatment should be used (endovascular or open surgery) (figure 2). Serial CT scans can be used to visualise the proximal neck (the transition between the normal and aneurysmal aorta), the extension to the iliac arteries, and the patency of the visceral arteries. They can also measure the thickness of the mural thrombus. Venous anomalies that can be hazardous during the access to the neck are also clearly indicated (left vena cava, posterior left renal vein). CT can also display the presence of blood within the thrombus (crescent sign), which has been regarded by some groups as a predictive marker of imminent rupture. In case of inflammatory aneurysm, CT allows estimation of the thickness of the aortic wall outside of the calcified deposits and visualisation of the presence of para-aortic fibrosis potentially associated with ureterohydronephrosis (figure 2 D). Extravasation of contrast material is diagnostic of aneurysm rupture. With three-dimensional imaging, helical CT and CT angiography can provide additional anatomical details, especially useful if endovascular procedure is considered. MRI, combined with magnetic resonance angiography (MRA) (figure 3), is of little harm since non-nephrotoxic contrast material (eg, gadolinium) is used, whereas conventional arteriography uses nephrotoxic contrast material, which can lead to renal failure and distal embolisation. Because of the steady development of MRA and CT angiographies, there will be hardly any place left for conventional aortography during preoperative assessment of the disorder. The use of conventional aortography is mainly restricted to the placement of endovascular devices or when a horse-shoe kidney is diagnosed.

Clinical presentation

Unruptured abdominal aortic aneurysms

Non-ruptured aneurysms are generally asymptomatic in most patients. They are essentially diagnosed incidentally during extensive clinical examination, especially in patients who complain of coronary, peripheral, or cerebro-vascular diseases, or during population screening. Fleming and co-workers reported that population screening in men aged 65–74 years significantly reduces mortality related to the disorder.

Non-ruptured aneurysms might exceptionally be diagnosed after complications, such as distal embolisation and, even more rarely, acute thrombosis. Minor and less specific symptoms include chronic vague abdominal and back pain, which can result from direct pressure or distension of adjacent structures. Recent onset of severe lumbar pain has been deemed to indicate impending rupture. Ureterohydronephrosis might also take place, especially if the aneurysm is inflammatory or involves the iliac bifurcation (figure 4).

Ruptured abdominal aortic aneurysms

Rupture of abdominal aortic aneurysms is heralded by the triad of sudden-onset pain in the mid-abdomen or flank (that may radiate into the scrotum), shock, and the presence of a pulsatile abdominal mass. However, the degree of shock varies according to the location and size of the rupture and the delay before the patient is examined. Rupture from the anterolateral wall into the peritoneal cavity is usually dramatic and most often associated with death at the scene. Most patients with a rupture who reach the clinic alive have a rupture of the posterolateral wall into the retroperitoneal space; a small tear can temporarily seal the rupture and the initial blood loss might be small. This initial event is systematically followed within hours by a larger rupture. This biphasic evolution emphasises the importance of the intermediate period after the initial event, which should be used for medical transfer and emergency repair.

Anecdotally, the first episode of rupture could be definitely contained and become a chronic pulsatile extra-aortic haematoma. Very rarely, the aneurysm might spontaneously rupture into the duodenum (figures 5 and 6); an incidence rate at necropsy of 0.04% to 0.07% has been reported. More often, aortoduodenal fistula can occur after previous repair, with an incidence rate of 0.5% to 2.3%. Rupture into the vena cava can also take place with an apparent pattern of lower extremity oedema erroneously attributed to cavoiliac thrombophlebitis (figure 4). However, the development of high output congestive heart failure and the perception of continuous
abdominal noise is pathognomonic. The overall prevalence of aortocaval fistula is 3% to 6% of all ruptured aortic aneurysms.

**Indications for treatment**

Although surgical treatment of non-ruptured abdominal aortic aneurysms relies on specific rare indications, such as distal embolisation, ureteral compression, or contained retroperitoneal haemorrhage, treatment of intact abdominal aortic aneurysm is essentially prophylactic and aimed at prevention of fatal rupture. Indication for surgical treatment is deduced from the estimated risk of rupture, the estimated risk of the surgical procedure, and the estimated life expectancy of the patient. Figure 7 shows a proposed management plan for asymptomatic abdominal aortic aneurysms.

**Risk of rupture**

The size of the aneurysm is a universally recognised factor to forecast rupture, and the general consensus is that patients with a large aneurysm should undergo surgery. The real controversy surrounds the management of small aneurysms. A study was undertaken in which patients with small aneurysms (diameter between 4·0 cm and 5·5 cm) were randomly assigned to two groups that underwent either early elective surgery or delayed repair after the diameter of the aneurysm had reached or exceeded 5·5 cm. The results show closely similar survival curves for the two groups of patients. A US Veterans Administration study led to similar findings despite a lower operative mortality (2·7% vs 5·8%) with early than with delayed repair. The conclusions of these two studies were similar: rigorous surveillance of infrarenal aortic aneurysms smaller than 5·5 cm in diameter is safe, whereas early surgery is not associated with improved long-term survival.

Rapid expansion of the aortic diameters preceding fissuration and rupture has been observed in abdominal aortic aneurysms independently of their initial size, which suggests that the size of the aneurysm, whatever its practical significance, is probably not the sole useful determinant for risk of rupture. Active investigations have been and still are being done to identify markers other than size that would predict a risk of rupture. A possible candidate is the level of serum MMP-9, which has been directly implicated in the proteolytic degradation of the extracellular matrix of the aortic wall. The amount of circulating MMP-9 has not only been reported to be significantly higher in patients with abdominal aortic aneurysm, but has also been significantly associated with the size and expansion rate of these aneurysms. Another factor that has been investigated as a potential serum marker is the reduced level of α1-antitrypsin (α1-AT) since it is the most abundant serum inhibitor of proteases. However, the importance of this marker for the prognosis of abdominal aortic aneurysms has not been defined because of contradictory findings.

Family history represents a risk factor for aneurysm rupture. A study of 313 pedigrees showed a four-fold higher rate of rupture in familial cases than in sporadic cases. Additionally, a significantly earlier age at rupture (65 years vs 75 years) was also reported in these familial cases. Another potential risk factor for rupture could be...

---

**Figure 5: Different possible sites of rupture of an abdominal aortic aneurysm**

1. Anterolateral free rupture in the abdominal cavity. 2. Retroperitoneal rupture. 3. Rupture of retroperitoneal sac. 4. Rupture into the duodenum. 5. Rupture into the inferior vena cava.

**Figure 6: Rupture of an abdominal aortic aneurysm in the duodenum**

CT displaying a known large aneurysm leaking into the duodenum (A, arrow), and presenting an air bubble in the parietal thrombus, which is a sign of aorta enteric fistula (B, arrow). The 92-year-old patient and her family had refused any intervention.
related to the sex of the patient. A report from the UK Small Aneurysm Trial has shown that the risk of rupture in women was four-times higher than in men.16

Preliminary data obtained by PET imaging of abdominal aortic aneurysms have shown focal uptake of 18-fluorodeoxyglucose (18F-FDG) within the aneurysm wall in patients with either large, rapidly expanding, or painful aneurysms (figure 8).114 The uptake of 18F-FDG is regarded as a functional image of the inflammatory infiltrate and thus as a potential non-invasive technique to identify unstable aneurysms that are prone to rupture.

**Risk of elective aneurysm repair**

Reported mortality rate related to elective aneurysm repair varies among hospitals and surgeons.115 Mean 30-day mortality rate has been reported at between 1·1% and 7·0%.116–121 Between 1998 and 2003, 453 patients were admitted to the University Hospital of Liège, Belgium, for elective repair of an abdominal aortic aneurysm (397 by open surgery and 56 by endovascular repair); the overall 30-day mortality, regardless of the risk factors, after open surgery and endovascular repair was 2·7% and 1·8%, respectively. Most deaths resulting from the repair occurred in the so-called high-risk patients. Factors of increased operative risk are renal failure, chronic obstructive pulmonary disease, and, most importantly, myocardial ischaemia. If these patients are excluded, 30-day mortality rate of elective repair should be expected to be as low as 2% in most hospitals.121–123

Several reports have shown the high incidence (between 40% and 60%) of coronary artery disease in patients with abdominal aortic aneurysm, which could be explained by common risk factors (eg, tobacco smoking and hypertension).124–127 Complications related to coronary artery disease are the main cause of the operative mortality of aneurysm surgery.126,128–131 However, the benefit of coronary artery revascularisation before surgery remains controversial. Simultaneous aneurysm repair and coronary artery revascularisation have been recommended in selected patients scheduled for elective or urgent repair.132–134 However, a large randomised study did not show any significant difference in the long-term outcome when coronary-artery revascularisation was undertaken before elective surgery.111 The investigators therefore suggested the restriction of preoperative coronary revascularisation to patients with unstable cardiac symptoms. As yet, there is no consensus on the optimum strategy for preoperative cardiac management in patients scheduled for major elective vascular surgery.136,137

As far as emergency repair for ruptured abdominal aortic aneurysms is concerned, mortality depends on the haemodynamic status of the patient at the time of surgery. By contrast with the progress in elective repair mortality, no improvement in operative mortality of ruptured aneurysms has been reported during the past decades, remaining as high as 30–70%.2,22 If the mortalities occurring at the scene of rupture, during transfer, shortly after admission to the emergency department, and during surgery are combined, then only 18% of patients with ruptured aortic aneurysms survive.138 Prance and co-workers138 suggested five preoperative risk factors to predict the mortality rate of ruptured abdominal aortic aneurysms: (1) age older than 76 years; (2) creatinine higher than 190 μmol/L; (3) haemoglobin below 9g/dL; (4) loss of consciousness; and (5) ECG evidence of ischaemia. In their study, the mortality rate was 100% when the patient had three or more risk factors and decreased to 48%, 28%, and 18% when the risk factor number decreased to two, one, or zero, respectively.138

**Management**

**Open surgical treatment**

During open surgical treatment, the abdomen is entered either through a long midline or a wide transverse incision. A retroperitoneal approach has been
Seminar

recommended in patients with chronic obstructive pulmonary disease. Disadvantages of this approach include: first that the intraperitoneal content cannot be inspected; and second, access to the right iliac artery can frequently be difficult, especially if there is a large right iliac aneurysm. Once the abdominal cavity is opened, the neck of the aneurysm needs to be identified to control it. In the cases of a suprarenal or intrarenal neck, a clamp above the renal arteries might be needed briefly. The iliac arteries are controlled in much the same way. The inferior mesenteric artery is tied close to the aortic wall to keep it collateral to the superior mesenteric artery; in some instances, encircling the inferior mesenteric artery with a rubber to reimplant it on the aortic prosthesis via a Carrell patch could be wise to maintain direct flow for sigmoid and rectum.

The vascular graft is a knitted synthetic textile sealed with collagen or albumin. The upper anastomosis is of the end-to-end type and the distal anastomosis is located on the aortic bifurcation, the iliac bifurcations, or the common femoral arteries depending on the extent of aneurysmal transformation and the patency of the external iliac arteries. Care is taken to preserve at least one internal iliac artery and to detect perioperatively a potential left colonic-ischaemia. In sexually active male patients, the recommendation is not to dissect the lateral left aortic wall and the common left iliac artery. Whenever possible, iliac anastomosis should be the preferred choice instead of common femoral anastomosis because anastomosis in the groin is more prone to infection. Specific morbidities linked to surgery are left colon ischaemia and renal failure (eventually due to thrombo-embolic events in renal arteries). Postoperative paraplegia, a huge concern with thoraco-abdominal surgery, is infrequent with abdominal aortic aneurysm surgery. The incidence of paraplegia after endovascular repair or open surgery has been reported to be 0·21% and 0·25–0·9%, respectively.

Since the very beginning of surgery for abdominal aortic aneurysms, survival after successful elective aneurysm repair has been reported to be less than the survival of the matched population for age and sex; 5-year survival after abdominal aortic aneurysm repair was about 70%, whereas the expected survival of a matched population was close to 80%.

Coggia and co-workers showed, in a preliminary study, the feasibility of repair by total laparoscopic surgery. Even if this technique is minimally invasive and reduces surgical trauma, more experience and further assessment are needed to ensure that the real benefit of this technique is realised compared with open repair.

**Endovascular repair**

Introduced by Parodi in 1991, endovascular repair consists basically of the placement of a graft across the aneurysm and the fixation to the normal aortic and iliac wall with stents at both ends. The aortomonoiliac approach consists of the insertion of a stent-graft, which is a tube of conventional graft fabric containing at least two stents. The delivery system consists of a sheath with some type of haemostatic mechanism and an obturator for sheath installation and stent-graft extrusion. Many CT angiogram-derived data are necessary to choose the best-adapted device for the patient. The endograft is composed of fabric and metal stents and comes loaded in a delivery system. Under fluoroscopic guidance, this introducer system is fed through the iliac arteries by means of catheters and guidewires until the endograft is positioned correctly at the top and bottom of the aneurysmal segment. Removal of the introducer system allows the fixing devices to attach to the aortic wall and hold the graft firmly in place, excluding blood flow from the aneurysm sac and removing pressure from the aneurysm wall. Excellent results are characterised by a perfectly canalised blood flow and later by a completely retracted aneurysm wall around the endograft (figure 9). Occasionally, an abdominal aortic aneurysm might rupture, despite the presence of the endograft, if there is still pressure in the aneurysmal sac via endoleaks (figure 9). Some 40–80% of abdominal aortic aneurysms could be amenable to endovascular grafting. However, morphological contraindications for endovascular repair...
are diversely proposed in published work. The most accepted contraindication is a proximal neck either shorter than 15 mm or absent.

Two randomised trials that compared conventional and endovascular repair showed a lower operative mortality rate for endovascular repair and less frequent complications than with conventional techniques. However, as discussed by Lederle, the question “Is endovascular repair preferable to open repair?” cannot be answered yet since there has been no long-term follow-up to determine whether the early promise of endovascular repair is sustained. Two large European registries have reported a failure rate for endovascular repair of 3% per year (1% by rupture and 2% by a required conversion to open repair) versus a failure rate of 0.3% for open repair. Long-term results after endovascular repair have also been reported to be worse for large aneurysms, which are the most in need of repair. Ouriel and colleagues noted that 2 years after endovascular repair, 6-1% of patients with an abdominal aortic aneurysm larger than 5-5 cm died from aneurysm-related causes and 8-2% needed open conversion. Additionally, a 4-year postoperative rupture rate was reported to reach 10% for large abdominal aortic aneurysms (diameter ≥6-5 cm).

Non-invasive prevention of growth and rupture

Patients with an abdominal aortic aneurysm are asked to stop smoking tobacco completely to allow a reduction of the growth rate of the aneurysm. The central role of MMP in aneurysm development and rupture has led to research interest in the pharmacologic inhibitors of these proteases. Tetracyclines provide a potentially effective treatment. Doxycycline, a synthetic tetracycline derivative, was shown to prevent MMP-mediated aneurysmal growth in animal models. Moreover, findings from a clinical study suggested that doxycycline treatment prevents aneurysm growth in human beings. The use of synthetic inhibitors of MMP activity, such as batimastat (BB-94), has also been shown to suppress the expansion of experimental abdominal aortic aneurysms. However, Defawe and colleagues recently showed that MMP can alter matrix remodelling independently of their proteolytic function, which suggests that the role of MMP might be more complex than mediation of a degradation process. Another attractive option of aneurysmal pharmacotherapy is to target the inflammatory response and interfere with the MMP pathway. Non-steroidal anti-inflammatory drugs, such as indometacin, are known to prevent development of abdominal aortic aneurysms in animal models. In the past decade, various substances have been proposed for the treatment of asymptomatic abdominal aortic aneurysms. The use of β-blocking agents (e.g., propranolol) seems to reduce the growth rate of large (>5 cm) aneurysms and even to lessen the size of experimental aneurysms. However, a randomised trial reported that propranolol does not have a significant effect on the growth rate of small aneurysms. Statins (hydroxymethylglutaryl coenzyme A reductase inhibitors), besides their cholesterol lowering effects, reduce the expression of various inflammatory molecules, including MMP. The addition of cerivastatin to tissue organ cultures of abdominal aortic aneurysms has been shown to down-regulate the production of MMP-9, which suggests that members of the statins family could prevent elastolysis in patients with this disorder. Long-term statin use in patients who underwent successful surgery was also associated with reduced mortality. Several innovative experimental studies have shown the potential of cell therapies (e.g., to seed cells overexpressing antiproteases in aneurysmal walls) for prevention of aneurysmal progression and rupture in animal models.

Conclusion

The past five decades have been marked by continuing progress in diagnosis, management, timing of interventional treatments, and assessment of endovascular repair versus conventional surgery. Hopefully, the upcoming decades will provide preventive treatments that can be applied to selected groups of individuals identified as high risk for abdominal aortic aneurysm by genomic or imaging technology, or both. The first goal in abdominal aortic aneurysm history was the prevention of rupture; the next aim will be the prevention of abdominal aortic aneurysm growth.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

We thank B V Nusgens and Ch M Lapierre (Laboratory of Connective Tissues Biology) for their help in the preparation of this Seminar. The experimental work reported here was supported by the Belgian Fonds Scientifique de la Recherche Médicale (FRSM 3-4785-05), the Fondation L. Fredericq, and the Fonds d’Investissement de la Recherche Scientifique de la Recherche Médicale. The funding sources had no role in the preparation of this Seminar.

References


Seminar


It was the morning of my 67th birthday, Jan 12, 2004. At 8.30 sharp, I felt the beginning of a low intensity, continuous pain in my lower left abdomen that irradiated to both flank and inguinal regions. I knew that I had hypertension, and that control of my weight was far from adequate. I had also previously suffered from chronic lumbar pain, which I thought was of spinal origin. But the onset of abdominal pain of an entirely new character and its location suggested otherwise—either renal colic (my left kidney has always been in an abnormal, pelvic position for many years) or the much more serious possibility of a leaking aortic aneurysm (previous ultrasonography had shown a slight dilatation of the abdominal aorta of 3 cm). So, I asked for an immediate abdominal ultrasound scan, which showed a large aortic aneurysm (9.5 cm). CT showed no signs of rupture (figure).

At 1 o’clock I consulted a vascular surgeon, who decided to admit me for an elective open repair. He did not regard mine as an emergency, preferring to do an angiography to visualise the aneurysm itself and the vasculature of the ectopic kidney. The pain was moderate, and I insisted on staying alone in my room where I could get some sleep.

At 3.30 the next morning, I was awoken by an excruciating abdominal pain. I can only describe its intensity as inhuman, evoking dreaded images of horror films in which the victim is perforated by an industrial drill. This pain went on for more than an hour. In the operating room, the anaesthesia brought me two different sensations: the relief of the end of pain and the certainty of no return. Yet, only a few hours later I found myself in the recovery room, awake and fully oriented. I learned that I had had a high perforation of the aneurysm, just below the origin of the right renal artery, with a huge retroperitoneal haematoma. During the procedure I had received ten red-cell units and two plasma units. The operation had been done via a transperitoneal approach and, because of my poor condition, limited to the substitution of the ruptured area by a Dacron segment.

I returned home on the sixth postoperative day. There were no important physical problems, though I had trouble eliminating excess fluid, which caused nocturnal and minimal effort dyspnoea and took at least a week to clear. I also experienced serious post-traumatic stress, waking every hour with nightmares, in which I relived the terrible pain that had preceded the operation. On Feb 14, I was able to resume my full professional duties.

From the time of hospital discharge to the present, my blood pressure control has been very strict, and I have lost at least 25 kg in weight. I knew that a second operation was essential after a minimum delay of 6 months, time needed for the surgeon to manoeuvre safely in a retroperitoneum dissected by the haematoma. So, until July, I had not dared to leave my home town. After another CT and MRI, the surgeon decided to use the same transperitoneal approach for an open reconstruction, instead of inserting an endoprosthesis.

On July 19, an aortobifemoral prosthesis was inserted in continuity with the previous Dacron segment. The surgeons anastomosed my ectopic kidney to the aortic prosthesis. This time, even though the operation was much longer (close to 7 hours), no blood transfusion was required, and I was discharged on the fifth postoperative day. My weight had increased by 8.5 kg, which I lost using an oral diuretic for 4 days. Fortunately, the nightmares did not recur and retrospectively I can tell that my general condition was radically better than after the previous surgery.

I learned important lessons from this dreadful experience: the ectopic kidney misled me since I had interpreted its presence as the cause of a pulsatile mass that I had noticed in the previous months. I should have had it checked since an aneurysm was a real possibility. I had been a smoker until the late 1970s so I was in a high-risk group; having had a known small dilatation of the abdominal aorta I should have had periodic screening. Even mild pain should be regarded as an indication of an ongoing rupture and signal immediate surgery. The consequences of doing otherwise are increased risk of death, especially if the rupture is open to the peritoneum instead of the retroperitoneum. Surgery presents the only possibility of survival. In my case, it has meant that I can now enjoy an extended, pleasant normal life.

**Figure:** Preoperative CT scan
Contrast differentiates true lumen from parietal thrombus.
Multiplicity in randomised trials I: endpoints and treatments

Kenneth F Schulz, David A Grimes

Multiplicity problems emerge from investigators looking at many additional endpoints and treatment group comparisons. Thousands of potential comparisons can emanate from one trial. Investigators might only report the significant comparisons, an unscientific practice if unwitting, and fraudulent if intentional. Researchers must report all the endpoints analysed and treatments compared. Some statisticians propose statistical adjustments to account for multiplicity. Simply defined, they test for no effects in all the primary endpoints undertaken versus an effect in one or more of those endpoints. In general, statistical adjustments for multiplicity provide crude answers to an irrelevant question. However, investigators should use adjustments when the clinical decision-making argument rests solely on one or more of the primary endpoints being significant. In these cases, adjustments somewhat rescue scattershot analyses. Readers need to be aware of the potential for under-reporting of analyses.

Many analytical problems in trials stem from issues related to multiplicity. Investigators usually address the issues responsibly; however, others ignore or remain oblivious to their ramifications. Put colloquially, some researchers torture their data until they speak. They examine additional endpoints, manipulate group comparisons, do many subgroup analyses, and undertake repeated interim analyses. Difficulties usually manifest at the analysis phase because investigators add unplanned analyses. Literally thousands of potential comparisons can emanate from one trial, in which case many significant results would be expected by chance alone. Some statisticians propose adjustments in response, but unfortunately those adjustments frequently create more problems than they solve.

Multiplicity problems stem from several sources. Here, we address multiple endpoints and multiple treatments. In the next article we address subgroup and interim analyses. The perspectives on multiplicity are contentious and complex. In proposing approaches to handle multiplicity, any position alienates many. Multiplicity issues stir hot debates.

The issue

Multiplicity portends troubles for researchers and readers alike for two main reasons. First, investigators should report all analytical comparisons implemented. Unfortunately, they sometimes hide the complete analysis, handicapping the reader’s understanding of the results. Second, if researchers properly report all comparisons made, statisticians proffer statistical adjustments to account for multiple comparisons. Investigators need to know whether they should use such adjustments, and readers whether to expect them.

Multiplicity can increase the overall error in significance testing. The type 1 error (α), under the hypothesis of no association between two factors, indicates the probability of the observed association from the data at hand being attributable to chance. It advises the reader of the likelihood of a false-positive conclusion. The problem emerges when multiple

Panel 1: Divergent views on statistical adjustments for multiplicity

Some statisticians favour adjustments for multiple comparisons, whereas others disagree.

* Several recent publications show that the multiple comparisons debate is alive and well. I . . . observe that it is hard to see views such as the following being reconciled . . . .*
* No adjustments are needed for multiple comparisons. *
* Bonferroni adjustments are, at best, unnecessary and, at worst, deleterious to sound statistical inference. *
* . . . Type I error accumulates with each executed hypothesis test and must be controlled by the investigators. *
* Methods to determine and correct type 1 errors should be reported in epidemiologic and public health research investigations that include multiple statistical tests. *
independent associations are tested for significance. If d=the number of comparisons, then the probability that at least one association will be found significant is 1–(1–α)^d. Frequently, investigators in medical research set α at 0.05. Thus, if they test ten independent associations, assuming the universal null hypothesis of no association in all ten, the probability of at least one significant result is 0.40, ie, (1–[1–0.05]^10). Stated alternatively, the cumulative chance of at least one false-positive result out of the ten comparisons is 40%. Nevertheless, the probability of a false positive for every single comparison remains 0.05 (5%) whether one or a million are tested.

**A proposed statistical solution**

Most statisticians would recommend reducing the number of comparisons as a solution to multiplicity. Given many tests, however, some statisticians recommend making adjustments such that the overall probability of a false-positive finding equals α after making d comparisons in the trial. Authors usually attribute the method to Bonferroni and simply state that making d comparisons in the trial. Authors usually attribute the method to Bonferroni and simply state that to test comparisons in a trial at α, all comparisons should be performed at the α/d significance level, not at the α level. Thus, for an α of 0.05, with ten comparisons, every test would have to be significant at the 0.005 level. Analogously, some investigators retain the same individual α threshold but multiply every observed p value by d. Thus, with ten comparisons, an observed p=0.02 from a trial would yield an adjusted p=0.20. Of note, the Bonferroni adjustment inflates β error thereby reducing statistical power.

Bonferroni adjustment, however, usually addresses the wrong hypothesis. It assumes the universal null hypothesis which, simply defined, tests that two groups are identical for all the primary endpoints investigated versus the alternative hypothesis of an effect in one or more of those endpoints. That usually poses an irrelevant question in medical research. Clinically, a similar idea would be: “...the case of a doctor who orders 20 different laboratory tests for a patient, only to be told that some are abnormal, without further detail.” Indeed, Rothman wrote: “To entertain the universal null hypothesis is, in effect, to suspend belief in the real world, and thereby to question the premises of empiricism.”

Drug regulation with the need for clear dichotomous answers appropriately drives much of the activity in multiplicity adjustments. Adjustments fit the hypothesis-testing paradigm—approval or no approval—needed for drug regulation. In most published medical research, however, we encourage the presentation of interval estimation (eg, relative risks with confidence intervals) for effects rather than just hypothesis testing (just a p value). Moreover, we suggest that the decision-making intent in most medical research discourages multiplicity adjustments.

**Multiple endpoints**

Although the ideal approach for the design and analysis of randomised controlled trials relies on one primary endpoint, investigators typically examine more than one. The most egregious abuse with multiplicity arises in the data-dredging that happens behind the scenes and remains unreported. Investigators analyse many endpoints, but only report the favourable significant comparisons. Failure to note all the comparisons actually made is unscientific if unwitting and fraudulent if intentional. “Post hoc selection of the end-point with the most significant treatment difference is a deceitful trick which invariably overemphasizes a treatment difference.” Investigators must halt this deceptive practice.

Researchers should restrict the number of primary endpoints tested. They should specify a priori the primary endpoint or endpoints in their protocol. Focusing their trial increases the simplicity of implementation and the credibility of results. Furthermore, they should follow their protocol for their analysis. Deviations for data-dredging can be condemned, but should be clearly labelled as explorations and fully reported. Disappointingly, trial reports frequently contain examinations of endpoints not included in the trial protocol but ignore planned primary analyses from the protocol. Safeguards to ensure that investigators have followed the protocol (such as The Lancet's protocol acceptance track and asking for protocols for all randomised controlled trials) provide assistance, but more extensive registering and publishing of protocols makes sense. Lastly, investigators must report all the comparisons made.

Statistical adjustments for multiple endpoints might sabotage interpretation. For example, suppose investigators undertook a randomised controlled trial of a new antibiotic compared with a standard antibiotic for prevention of febrile morbidity after hysterectomy. They designated fever the primary outcome, and the results showed a 50% reduction (relative risk 0.50 [95% CI 0.24–0.97]; p=0.048). Note the significant result. Alternatively, suppose they had designated two primary endpoints: wound infection and fever. As typically happens in trials, the endpoints are highly correlated. So in addition to the 50% reduction in fever, the trial also found a 52% decrease in wound infection (0.48 [0.24–0.97]; p=0.041). From some statisticians’ viewpoints, investigators should correct for multiple comparisons by, for example, multiplying every p value by the number of comparisons made—ie, 0.048×2=0.096 and 0.041×2=0.082. Both p values adjust to >0.05; thus the trial would be indeterminate (negative).

Seasoned clinical trialists, however, look at these results quite differently. The wound infection result enhances rather than debases the first result on fever. Clinicians understand biologically that the two endpoints are highly related. Adding the second endpoint on wound infection and observing similar results lends credence to the observed reduction in febrile morbidity. That adjustments
would abolish the basic finding defies logic. Doing so would somewhat resemble a doctor finding an abnormally low amount of haemoglobin in a patient but no longer judging it worthy of treatment because they also obtained an abnormal packed-cell volume (haematocrit).

Indeed, some statisticians would agree with not using formal adjustments for multiplicity in the aforementioned example. Even those predisposed to such adjustments recommend against them under certain delineated clinical decision-making scenarios. If an investigator proposes to claim treatment effect if all the endpoints are significant or if most (defined in the protocol) are significant, then they assert that no adjustment for multiple endpoints is necessary.

Furthermore, the Bonferroni adjustment, advocated most frequently for multiplicity, is an overcorrection at best. Moreover, it can be a severe overcorrection when the endpoints are associated with one another, which is generally the case. Overcorrecting for p values hampers interpretation of results. The adjustment for multiple comparisons “mechanizes and thereby trivializes the interpretive problem, and it negates the value of much of the information in large bodies of data”. Clinical insights remain important. Investigators need to focus on the smallest number of endpoints that makes clinical sense and then report results on all endpoints tested. If more than one primary endpoint exists, they should discuss whether additional endpoints reinforce or detract from the core findings. Formal adjustments for multiplicity frequently obscure rather than enhance interpretation.

Composite endpoints
Composite endpoints alleviate multiplicity concerns. A composite endpoint happens if any one of the prospectively defined components of the composite takes place. For example, a composite cardiovascular endpoint would happen if myocardial infarction, stroke, or cardiovascular death arose. If designated a priori as the primary outcome, the composite obviates the multiple comparisons associated with testing of the separate components. Moreover, composite outcomes usually lead to high event rates thereby increasing power or reducing sample size requirements. Not surprisingly, investigators frequently use composite endpoints.

However, interpretational difficulties sometimes arise. For example, aspirin produced an 18% reduction (relative risk 0.82 [95% CI 0.70-0.96]) in the above-defined composite endpoint of cardiovascular events (myocardial infarction, stroke, or cardiovascular death), a seemingly worthwhile result. However, a secondary look at the separate components revealed a 44% decrease in myocardial infarction, a 22% increase in stroke, and virtually no effect on cardiovascular death. That 18% reduction seems meaningless in view of the lack of beneficial effect on the relatively more important outcomes of death and stroke. Composite endpoints frequently lack clinical relevancy. Thus, composite endpoints address multiplicity and generally yield statistical efficiency at the risk of creating interpretational difficulties.

Multiple treatments (multiam trials)
Addressing multiplicity from multiple treatments is a more tractable problem than from multiple endpoints. First, investigators can avert multiple tests by one global comparison A vs B vs C in a three-arm trial—or by modelling a dose-response relation. Second, and perhaps most importantly, researchers have less opportunity to data-dredge on many treatments and not report them. While they easily can add more endpoints for analysis, they would have difficulty adding treatments in a trial. They theoretically could implement a multigroup trial and then only report the favourable group comparisons, but little evidence exists for that practice. We suspect that readers of a trial report usually see all the treatments implemented. Indeed, multiam trials have an important role in medical research (panel 2).

Panel 2: A role for multiam trials in medical research
Multiam trials are fairly common in the medical literature. A search of parallel designed randomised controlled trials indexed on PubMed in 2000 revealed 25% with more than two arms. Of those, 62% had three arms, 26% had four, and 12% had more than four (Altman DG, personal communication).

The preponderance of material in clinical trial textbooks addresses two-arm trials. Furthermore, eminent researchers have strongly recommended against more than two-arms: “A positive result is more likely, and a null result is more informative, if the main comparison is of only two treatments, these being as different as possible.” The argument against multiam trials mainly centres on trial power. Published trials typically have inadequate power. Given a finite number of potential participants, the argument holds that adding arms only further dilutes power. Although we sympathise with this argument, multiam trials might not only be attractive in some circumstances but also be more efficient.

For example, imagine an instance whereby a standard treatment exists and two new potentially effective therapies have materialised. A two-arm approach dictates a comparison of a new with standard and then probably an additional trial of the other new with a group from the first trial. In general, the overall study size and cost would be greater with this sequential two-arm approach than with one multiam trial. Multiam trials sometimes make sense. Furthermore, multiam trials do not necessarily raise methodological concerns. They can eliminate selection bias just like two-arm trials. Although they tend to be more complex to undertake and analyse, that complexity frequently yields commensurate gains in information.

Panel 2: A role for multiam trials in medical research
Multiam trials are fairly common in the medical literature. A search of parallel designed randomised controlled trials indexed on PubMed in 2000 revealed 25% with more than two arms. Of those, 62% had three arms, 26% had four, and 12% had more than four (Altman DG, personal communication).

The preponderance of material in clinical trial textbooks addresses two-arm trials. Furthermore, eminent researchers have strongly recommended against more than two-arms: “A positive result is more likely, and a null result is more informative, if the main comparison is of only two treatments, these being as different as possible.” The argument against multiam trials mainly centres on trial power. Published trials typically have inadequate power. Given a finite number of potential participants, the argument holds that adding arms only further dilutes power. Although we sympathise with this argument, multiam trials might not only be attractive in some circumstances but also be more efficient.

For example, imagine an instance whereby a standard treatment exists and two new potentially effective therapies have materialised. A two-arm approach dictates a comparison of a new with standard and then probably an additional trial of the other new with a group from the first trial. In general, the overall study size and cost would be greater with this sequential two-arm approach than with one multiam trial. Multiam trials sometimes make sense. Furthermore, multiam trials do not necessarily raise methodological concerns. They can eliminate selection bias just like two-arm trials. Although they tend to be more complex to undertake and analyse, that complexity frequently yields commensurate gains in information.
However, the situation is not entirely sanguine. What readers of a journal article might not see are all the different comparisons among the treatment groups. For example, with a three-arm trial, at least seven possible analyses emerge (figure). With more than three arms the potential comparisons explode. Obviously, investigators should specify a priori the comparisons intended.

With multiarm trials, as mentioned earlier, a frequently recommended approach entails undertaking one global test across all treatments. However, some methodologists believe such tests are of limited use because they do not identify which treatments are different and because of limited power to detect genuine differences. Many multiarm trials are designed for direct comparison with controls. Thus, investigators should plan the comparisons intended, limit the number, and document them in the protocol.

Adjustments for multiple comparisons generally need not have a role in multigroup trials. Similar to the above argument for multiple endpoints, clinicians usually find the addition of a group to a trial enhances rather than diminishes informativeness. For example, in the randomised controlled trial described earlier, comparing a new antibiotic with standard treatment for prevention of fever after hysterectomy, investigators might add a treatment group with a 300 mg dose to a trial of a 200 mg dose of the same antibiotic. The results showed a 40% reduction for the 200 mg dose (relative risk 0.60 [95% CI 0.37–0.98]; p=0.044). Note the significant result. The 300 mg dose expectedly yields a similar result, a 45% decrease in fever (0.55 [0.31–0.98]; p=0.041). The simple adjustment approach for multiple comparisons involves multiplying every p value by the number of comparisons made—ie, 0.044×2=0.088 and 0.041×2=0.082. With adjustment, the effects become non-significant at the 0.05 level and thus indeterminate (negative).

Again, however, trialists interpret these results quite differently. The result for 300 mg augments rather than degrades the result for 200 mg on fever. Clinicians expect similar results biologically. They would seriously distrust adjustment that abolishes those significant results. Adjusting p values, particularly with related treatment groups, does not aid in interpreting the results of the trial.

With multiple treatments, investigators sometimes use a prioritised sequence of tests. For example, investigators might decide on the 300 mg new antibiotic versus standard treatment as the priority test and, if that comparison is significant, only then proceed to the 200 mg comparison. Such procedures address multiplicity without adjustments. Again, formal adjustments for multiplicity usually complicate rather than enlighten.

### The role of adjustments for multiplicity

Sometimes formal adjustments for multiplicity are inescapable. An obvious example would arise with certain decision-making criteria in submissions to a regulatory agency for drug approval. If the sponsor specifies more than one primary endpoint and proposes to claim treatment effect if one or more are significant, investigators should adjust for multiplicity. Furthermore, the same principle extends to all investigators whose decision-making intent is to claim an effect based on any one of a number of endpoints being positive.

Adjustments might also be indicated in a multiarm trial in which investigators plan a scattershot analysis. For example, in a four-arm trial (treatments A, B, C, and D), they intend on claiming an effect for A if any one of the following comparisons yielded significant results: A versus B, A versus C, A versus D, A versus B+C, A versus B+D, A versus C+D, or A versus B+C+D. The best recourse might be a multiplicity adjustment.

In general, when prudence indicates multiplicity adjustments, trials tend to be poorly and diffusely designed. An adjustment for multiplicity merely partly salvages credibility. Moreover, even when adjustment becomes appropriate, implementation becomes difficult. Bonferroni adjustments are generally recommended, usually because of their simplicity. However, other adjustment strategies sometimes perform better. Depending on the correlation among the endpoints, simulation experiments display wide variability in α error and power of various multiplicity adjustment strategies. These comparative assessments help, but still clear-cut choices prove elusive. The adjustments usually provide crude answers.

### What readers should look for

Readers should expect the researchers to report all the endpoints analysed and treatments compared. Assessing whether they reported them all is usually difficult.
Access to the protocol would be helpful but is usually impossible. We urge greater access to protocols. Poor, incomplete reporting, however, frequently renders readers helpless to know the complete analysis undertaken by the investigators. Reporting according to the CONSORT statement obviates these difficulties.16,17

Readers should expect the primary endpoint or endpoints to be specified, with other analyses being labeled as exploratory. In lieu of direct statements, search for indirect indications. If the primary endpoint remains unclear, hopefully the authors provided a statistical power analysis that indicates the primary endpoint.

Readers should expect some interpretation if authors make multiple comparisons. If authors went overboard and reported results on 15 endpoints with one being significant they should display appropriate caution. If multiple comparisons yield multiple effects, authors should address the internal consistency of the results. Most importantly, transparent reporting of all comparisons allows readers to come to their own interpretations.

If a trial report specifies a composite endpoint, the components of the composite should be in the well-known pathophysiology of the disease. The researchers should interpret the composite endpoint in aggregate rather than as showing efficacy of the individual components. However, the components should be specified as secondary outcomes and reported beside the results of the primary analysis.18

In general, readers need not expect corrections for multiplicity. For most trials, adjustments for multiplicity lack substance and prove unhelpful. An exception might include a medical research article with an argument that rests solely on one or more of the primary endpoints being significant, essentially the test of the universal null hypothesis. An adjustment for multiplicity somewhat rescues such scattershot analyses.

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

Acknowledgments
We thank David I. Sackett, Douglas G Altman, and Willard Cates for their helpful comments on an earlier version of this report.

References
A 56-year-old woman with a history of well-controlled insulin-dependent diabetes mellitus and hypertension donated platelets in July, 2004. She had donated platelets approximately 50 times in the previous 5 and a half years. We did routine testing for bacterial contamination, and cultured *Streptococcus bovis* from the units she donated, 11·4 hours after incubation. We contacted the woman to discuss this result, and she reported no additional medical history. She had never had a colonoscopy, and we encouraged her to see her doctor for follow-up. We destroyed the units from her donation, and told her that she could not donate until she was cleared of not having an obvious source of *S bovis* bacteraemia. Her doctor ordered two sets of blood cultures to look for persistent bacteraemia. These cultures were negative. The woman then had a colonoscopy that showed a friable mass, 3 cm diameter, in the sigmoid colon. Sigmoidectomy was done in September, 2004, for moderately differentiated adenocarcinoma. The woman tolerated surgery well, and had no complications.

The major cause of platelet contamination is usually normal skin flora,¹ which can be introduced at the time of blood collection. Donor bacteraemia is less likely, though it can occur after dental manipulation such as tooth extraction, tooth brushing, or use of gum irrigation devices.¹ *S bovis* is not typical skin flora or a ubiquitous environmental agent. *S bovis* bacteraemia has been found in association with gastrointestinal neoplasia (ranging from colonic polyps to cancer), extracolonic malignancies, liver disease, endocarditis, cholangitis, meningitis, and diabetes mellitus.²³ The American Society of Hematology and American Association of Blood Banks has published guidelines on the regulation of bacterial contamination of blood components,¹ recently updated to include measures to detect and limit bacterial contamination in all platelet components.¹ Bacterial sepsis related to transfusion is the second most frequent transfusion-related cause of fatalities in the USA.¹

Specific identification of bacterial organisms found in testing of platelet donations can have value to the donor, in addition to the obvious benefit for any potential recipient or for quality control. Bacterial culture in this apparently healthy blood donor led to a timely diagnosis of malignancy, and a good clinical outcome.

References