1. Medical education in the UK: building a firm foundation • EDITORIAL
   Page 607
   The Lancet

2. Technology can make patients safer • EDITORIAL
   Page 608
   The Lancet

3. Chocolate—more a food than a medicine • EDITORIAL
   Page 608
   The Lancet

4. Surgical resection improves outcome in metastatic epidural spinal cord compression • DISCUSSION
   Pages 609-610
   Martin J van den Bent

5. Genetic susceptibility to bladder cancer • DISCUSSION
   Pages 610-612
   Emanuela Taioli and Sara Raimondi

6. Neonatal hearing screens: Wessex re-visited • DISCUSSION
   Pages 612-613
   Patricia Mutton and Kenneth Peacock

7. Role of the global civil society • DISCUSSION
   Pages 613-615
   Ian Magrath

8. Does a NEAT difference in energy expenditure lead to obesity? • DISCUSSION
   Pages 615-616
   Gema Frühbeck

9. Health research: what's culture got to do with it? • DISCUSSION
   Pages 617-618
   Linda M Hunt
<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Type</th>
<th>Pages</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>Cerebral perfusion deficits in dysbaric illness</td>
<td>CORRESPONDENCE</td>
<td>632-633</td>
<td>JJW Sykes</td>
</tr>
<tr>
<td>23</td>
<td>Cerebral perfusion deficits in dysbaric illness</td>
<td>CORRESPONDENCE</td>
<td>633</td>
<td>Ramsay Pearson</td>
</tr>
<tr>
<td>24</td>
<td>Combination anti-hypertensives in WHO essential medicine list</td>
<td>CORRESPONDENCE</td>
<td>633-634</td>
<td>Madhur D Bhattarai</td>
</tr>
<tr>
<td>25</td>
<td>Adjuvant trastuzumab for HER2-positive breast cancer</td>
<td>CORRESPONDENCE</td>
<td>634</td>
<td>James Spicer, Mark Harries and Paul Ellis</td>
</tr>
<tr>
<td>26</td>
<td>Controlling the three “P”s in Africa</td>
<td>CORRESPONDENCE</td>
<td>634</td>
<td>Felix ID Konotey-Ahulu</td>
</tr>
<tr>
<td>27</td>
<td>Department of Error</td>
<td>ERRATUM</td>
<td>634</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Chemotherapy versus allogeneic transplantation for very-high-risk</td>
<td>ARTICLE</td>
<td>635-642</td>
<td>Adriana Balduzzi, Maria Grazia Valsecchi,</td>
</tr>
<tr>
<td></td>
<td>childhood acute lymphoblastic leukaemia in first complete remission:</td>
<td></td>
<td></td>
<td>Cornelio Uderzo, Paola De Lorenzo, Thomas</td>
</tr>
<tr>
<td></td>
<td>comparison by genetic randomisation in an international prospective</td>
<td></td>
<td></td>
<td>Klingebiel, Christina Peters, Jan Stary,</td>
</tr>
<tr>
<td></td>
<td>study</td>
<td></td>
<td></td>
<td>Maria S Felice, Edina Magyarosy, Valentino</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conter et al.</td>
</tr>
<tr>
<td>29</td>
<td>Direct decompressive surgical resection in the treatment of spinal</td>
<td>ARTICLE</td>
<td>643-648</td>
<td>Roy A Patchell, Phillip A Tibbs, William F</td>
</tr>
<tr>
<td></td>
<td>cord compression caused by metastatic cancer: a randomised trial</td>
<td></td>
<td></td>
<td>Regine, Richard Payne, Stephen Saris,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Richard J Kryscio, Mohammed Mohiuddin and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Byron Young</td>
</tr>
<tr>
<td>30</td>
<td>NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder</td>
<td>ARTICLE</td>
<td>649-659</td>
<td>Montserrat García-Closas, Núria Malats,</td>
</tr>
<tr>
<td></td>
<td>cancer: results from the Spanish Bladder Cancer Study and meta-</td>
<td></td>
<td></td>
<td>Debra Silverman, Mustafa Dosemeci, Manolis</td>
</tr>
<tr>
<td></td>
<td>analyses</td>
<td></td>
<td></td>
<td>Kogevinas, David W Hein, Adonina Tardón,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Consol Serra, Alfredo Carrato, Reina</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>García-Closas et al.</td>
</tr>
<tr>
<td>Article Number</td>
<td>Title</td>
<td>Type</td>
<td>Pages</td>
<td>Authors</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>31</td>
<td>Universal newborn screening for permanent childhood hearing impairment: an 8-year follow-up of a controlled trial</td>
<td>Article</td>
<td>660-662</td>
<td>Colin Kennedy, Donna McCann, Michael J Campbell, Lindsay Kimm and Roger Thornton</td>
</tr>
<tr>
<td>32</td>
<td>Obesity, cigarette smoking, and telomere length in women</td>
<td>Article</td>
<td>662-664</td>
<td>AM Valdes, T Andrew, JP Gardner, M Kimura, E Oelsner, LF Cherkas, A Aviv and TD Spector</td>
</tr>
<tr>
<td>33</td>
<td>Phaeochromocytoma</td>
<td>Review Article</td>
<td>665-675</td>
<td>Jacques WM Lenders, Graeme Eisenhofer, Massimo Mannelli and Karel Pacak</td>
</tr>
<tr>
<td>34</td>
<td>Buying results? Contracting for health service delivery in developing countries</td>
<td>Article</td>
<td>676-681</td>
<td>Benjamin Loevinsohn and April Harding</td>
</tr>
<tr>
<td>35</td>
<td>End-of-life: a Hindu view</td>
<td>Discussion</td>
<td>682-686</td>
<td>Shirley Firth</td>
</tr>
<tr>
<td>36</td>
<td>Benefits of experience—missing the perf</td>
<td>Discussion</td>
<td>687</td>
<td>JS Huntley</td>
</tr>
<tr>
<td>37</td>
<td>Placental site trophoblastic tumour arising from a partial hydatidiform mole</td>
<td>Short Communication</td>
<td>688</td>
<td>Carlo Palmieri, Rosemary A Fisher, Neil J Sebire, Iain Lindsay, J Richard Smith, W Glenn McCluggage, Phillip Savage and Michael J Seckl</td>
</tr>
</tbody>
</table>
Medical education in the UK: building a firm foundation

Last week saw a quiet revolution in postgraduate training for medical school graduates in the UK’s National Health Service (NHS)—quiet in the sense that the public, who pay for and use the service, will barely have known about it. However, almost 5000 trainees entered the new Foundation Programme last week.

Early in 2003, the UK Department of Health published *Modernising medical careers*, in which major reform of postgraduate medical training and education was proposed. The existing “apprenticeship model” of training junior doctors was seen as outdated and inflexible in that they had to make a specialist career choice too early. In the old model, new medical graduates spent a preregistration year as a house officer and a second (and usually subsequent years) as a senior house officer, before becoming a specialist registrar en route to a consultancy or entering training for general (ie, family) practice. The apprenticeship model was also seen as lacking team and multidisciplinary working. Later that year, a strategy group was convened to oversee the project, including the Department of Health, the General Medical Council, the Royal Colleges, and various other bodies involved in postgraduate medical education. Their work was summarised in *The next steps* document, published in April, 2004.

In the Foundation Programme, the first year is akin to the current preregistration year, and the second replaces the first year of senior house officer training. The programme is based on achieving nationally consistent sets of competencies rather than on time served, as in the apprenticeship model. The competencies are predefined observed behaviours, skills, and attributes, linked with the new knowledge being learned. Strategically, the Foundation Programme means an overall national training programme for the first 2 years after medical school. For junior doctors, it means placements through which they must progress. For postgraduate deans, who will still be running the training locally, the programme is a series of organised placements through which trainees can acquire and be assessed in the competencies required. What will be common throughout the UK will be a programme which is “managed, structured, assessed, approved, and quality-assured” and which provides “choice and variety to trainees”. The Foundation Programme has been set up to meet the legal requirements of training for the General Medical Council, and comply with the European Working Time Directive which limits working hours, including those for junior doctors.

The placements for trainees in the programme can be seen as “units of approval” that are educationally viable and manageable by deaneries. That means that each deanery has to provide the infrastructure for training, although a full commitment to the use of electronic training aids is lacking. Such e-training, which recent graduates generally like and want to use, is expensive and requires a computerised environment that is not nationally available. What is needed is an integrated system for electronic training aids (medical textbooks and databases, including access to bibliographic databases) that are geared to the Foundation Programme’s curriculum, electronic patients’ records held securely and in compliance with data-protection regulations, and virtual training suites. Training will be mostly in acute-care settings and will include mental health and general practice, although training in other settings is not excluded, depending on local needs. An important change is that competencies will include communications skills, decision-making, setting priorities, and teamworking, as well as diagnosis and treatment. A priority in the new programme is patients’ safety.

The Foundation Programme has been generally welcomed but also criticised by the British Medical Association, which worries that existing workforce problems have not been dealt with or might be worsened. The programme will not create enough training places for the increasing numbers graduating from UK medical schools, says the Association, and some deaneries are too poor to provide educational approval to the training posts demanded locally by hospital trusts. Unapproved posts do not count towards the Foundation Programme’s requirements.

This training revolution for UK junior doctors may have slipped in somewhat quietly, but will patients notice any difference? Time will tell. A training programme for junior doctors that includes a high priority on patients’ safety and communication and that should lead to joined-up training nationally certainly looks good on paper. Time will also tell whether the Foundation Programme will prove to be a blueprint for medical education elsewhere in the world.
Technology can make patients safer

Misidentification of patients, leading to inappropriate drug administration or surgery, is a common problem. In a busy clinical environment it is all too easy to prescribe, dispense, or administer the wrong drug or dose. However, innovative technology is beginning to deliver systems that should improve patient safety, reducing human errors to a minimum.

Wristbands have been used in hospitals for decades to identify patients. Electronic wristbands, which contain detailed information about the patient’s history and treatment requirements, are being introduced. For example, one electronic wristband, launched last week, alerts a surgical team if the patient’s incision site has not been marked. The wristband, which costs US$2-50, is disposable and contains electronics similar to those used in anti-theft chips in shops. Another prototype wristband, designed by a student from Brunel University in the UK, could ensure that patients receive the correct medication by matching an electronic chip in the bracelet with an electronic chip in the drug’s packaging.

Harnessing innovative technologies like these will be essential if health-care providers are to reduce the high incidence of medical errors. Indeed, earlier this year the US Institute of Medicine published a report entitled Building a better delivery system: a new engineering/health care partnership. As the report points out, health-care providers have been slow to adapt some of the systems-engineering tools, which have been used successfully in the commercial sector, in a medical environment.

This needs to change, especially considering the scope of the problem. The Institute of Medicine estimates that medical errors cause up to 98 000 deaths a year in hospitals in the USA. Other developed countries have reported similar statistics. Encouragingly, however, there is considerable political will to address patient safety. For example, in October, 2004, WHO launched the World Alliance for Patient Safety to help countries develop safety policies and practices. These initiatives will need to invest in cost-effective and simple technological interventions if medical errors are to be reduced.

Chocolate—more a food than a medicine

Does eating chocolate prevent or even treat cardiovascular and cerebrovascular disease, or do the calories, and especially the high fat and sugar content, do more harm than good? The health benefits or harms of chocolate have intrigued researchers and provided fodder for media frenzies for decades with every positive study of the effects of cocoa flavanols (a flavonoid subclass) making the headlines. It would suit many a chocolate lover if it were proven beyond doubt that chocolate offers substantial health benefits.

It would also suit, naturally, chocolate manufacturers, not all of whom are sitting back and waiting for others to prove their case. Mars, Incorporated, is one manufacturer who has supported research into the health effects of cocoa for more than 15 years. At a Mars-sponsored meeting last month, scientists from Mars reported that they had synthesised flavanols from natural cocoa flavanols, and that the company is now seeking pharmaceutical partners for a licensing or joint venture agreement to develop some of these compounds as prescription drugs.

So how robust is the evidence that cocoa flavanols are beneficial? Much of the epidemiological data on flavonoid-rich products have focused on red wine and tea, and there is certainly some evidence that high intakes of some flavonoids are associated with a low risk of coronary heart disease and stroke but the precise role of cocoa is more difficult to determine. Preliminary research indicates that cocoa flavanols have antioxidant effects, decrease LDL-cholesterol oxidation, reduce platelet aggregation, and enhance endothelial function. But huge variability in cocoa processing and flavanol content makes it hard, at present, to determine how far these early findings translate into tangible clinical benefits. What is clear is that few confectionary chocolate products contain doses of flavanols that are even likely to be beneficial. Dark chocolate may be better than milk chocolate, and certainly better than white chocolate, but cocoa-rich products in which flavanols have been preserved need to be developed and tested in trials before any clear-cut health benefit can be ascribed to them. At least trial recruitment should be straightforward.
Surgical resection improves outcome in metastatic epidural spinal cord compression

In today’s Lancet, Roy Patchell and colleagues1 present the long awaited outcome of a randomised trial of decompressive surgical resection in patients with metastatic epidural spinal cord compression (MESCC), which will have a major effect on the therapeutic approach to these patients. MESCC is a great threat for all cancer patients with vertebral body metastases, which can lead to paraplegia. The current standard of care consists of a high level of suspicion toward cancer patients with back pain, a low threshold for performing MRI, and immediate radiotherapy in cases of spinal cord compression. Once paraplegia develops the outlook for the patient is grim. Not only does paraplegia severely affect quality of life, but the functional response to radiotherapy mainly depends on the ambulatory status of the patient at the start of treatment: most ambulant patients will remain ambulant, but non-ambulant patients rarely recover to an ambulant status and have a very limited survival.2

Previous randomised or retrospective studies did not identify a benefit of surgical procedures before radiotherapy, but these studies mainly addressed posterior surgical approaches (laminectomy).3,4 However, in MESCC the compression of the spinal cord usually originates from the vertebral body, which is difficult to reach by a posterior approach. Furthermore, posterior surgery might actually contribute to mechanical instability of the spinal column by further disrupting its stability, which can add to the spinal cord compression. Several recent uncontrolled studies have shown interesting outcomes of anterior and lateral approaches to the spinal column in MESCC.5 Not only did these studies report good pain control, but also substantial neurological improvements in patients with severe motor deficits and with limited morbidity. The arguments against these surgical procedures were mainly the absence of evidence from randomised trials, the assumed morbidity, and extensive hospital stay for patients with a poor prognosis.

The randomised study by Patchell and colleagues seems to pave the way for a widespread acceptance of decompressive spinal column surgery in selected patients with a single area of MESCC from a solid tumour. Patchell and colleagues compared an immediate circumferential decompression of the spinal cord followed by $10 \times 3$ Gy radiotherapy with the same radiotherapy regimen without surgery. The primary endpoints were ability to walk and ambulatory time after treatment. 200 patients had been foreseen for the trial, but the study was closed after accrual of 123 patients when superiority of the surgical treatment was shown at interim analysis. The outcome of the study is convincing: the percentage of patients retaining the ability to walk after treatment was higher after surgery than after radiotherapy only (84% vs 57%, p=0.001). Most striking, in the surgical group ten of 16 paraplegic patients at the time of study entry regained the ability to walk, by contrast with only three of 16 in the radiotherapy only group (p=0.012). Both morphine and steroid use were less in the surgical group, and this group survived longer (median survival time 126 days vs 100 days, p=0.033). Surgery did not increase the hospital stay and 30-day morbidity was worse in the radiotherapy only group. In short, surgery patients tended to remain ambulant without excessive morbidity, whereas radiotherapy patients tended to become paraplegic.

These results show that anterior decompressive surgery is an option for many MESCC patients currently managed with radiotherapy only. Patient selection, however, is an important consideration: only patients with a satisfactory medical status (medical status sufficient to undergo surgery and an expected survival of 3 months) and with...
only a single area of MESCC were eligible, in this trial. Moreover, it took this multicentre trial 10 years to accrue 123 patients, which might suggest that not all eligible patients were actually considered for the study. Despite this, median survival in the study was just over 4 months in the surgical group, and even with this limited survival a clear benefit was obtained. It will, however, be a great clinical challenge to select patients for this type of intervention and to identify those patients in whom the improved outcome outweighs the efforts and costs of surgical intervention. An even greater challenge will be to organise the surgical care of MESCC patients. These time-consuming operations are emergency procedures, which even in the present study manifest themselves all too often on Friday afternoon. Such operations demand cooperation and dedication from several medical disciplines, and will disturb regular (surgical) programmes. Still, the improved outcome observed after surgery necessitates concerted action in specialised centres to make such interventions possible.

Martin J van den Bent
Daniel den Hoed Oncology Center, Neuro-Oncology Unit, 3008 AE Rotterdam, Netherlands
m.vandenbent@erasmusmc.nl
I declare that I have no conflict of interest.

Genetic susceptibility to bladder cancer

The study from Montserrat García-Closas and colleagues1 in today’s Lancet brings new evidence in favour of a role for metabolic gene polymorphisms in bladder cancer risk. Bladder cancer is among the most frequent cancers in the developed world, and the incidence is rising in Europe and the USA.2 Incidence and mortality increase with age; with males being more affected than females.3 The known risk factors are smoking and occupational exposure to aromatic amines. Although the association with occupational exposure is reported to be stronger than the association with smoking, the high prevalence of smokers gives a high population attributable risk, and therefore makes this risk factor more relevant at a population level.1 The risk of bladder cancer associated with aromatic amines derived from both occupational exposure and smoking has directed attention to genes encoding proteins that are on the pathways of amine metabolism such as GSTM1 (glutathione S-transferase M1) and NAT2 (N-acetyltransferase-2). The issue is whether patients carrying unfavourable polymorphisms in these genes are at a higher risk of bladder cancer than patients with more favourable combinations of genetic polymorphism, given the same level of environmental exposure.

The present study shows a significantly increased risk of bladder cancer with a GSTM1 null genotype or with the NAT2 slow acetylator genotype, a weak interaction between the two genetic polymorphisms, and an interaction between NAT2 acetylator genotype and smoking status. The results are strengthened by an update of a previous meta-analysis and pooled analysis.4 Two interesting results arise from García-Closas’ current study: one is the distinct bladder cancer risk ascribed to the presence of one or two copies of GSTM1, an issue that has not been considered in the past, because patients with one functional copy are usually treated as having both GSTM1 alleles present. These patients represent 40% of the controls in the present study, and have their own excess risk for bladder cancer, and probably for other environmentally-related cancers as well. The other interesting point presented is the presence of a plateau in bladder cancer risk for smokers in the category of those who smoke over 20–29 cigarettes a day, irrespective of their acetylator status. This finding seems to indicate an effect of other environmental factors, such as occupational exposure, or the interaction of NAT2 with other genetic polymorphisms.

Gene-environment and gene-gene interactions are becoming more and more important in epidemiology,
but require large sample sizes to be tested. Large sample sizes can be obtained by prospectively collecting data according to a common protocol, through a consortium, or by pooling individual data from ongoing studies. This procedure was the basis of GSEC, an International Collaborative Study on Genetic Susceptibility to Environmental Carcinogens, which started in 1997, and now contains individual data on about 70 000 cancer cases and controls. The study gives an opportunity to look at several gene-disease associations (table). The results on bladder cancer, GSTM1, and NAT2 are the strongest associations found in the dataset, and are confirmed by the results published by García-Closas and co-workers. In addition to the interaction with smoking, an interaction between NAT2 and occupational exposure was observed in the GSEC pooled analysis.6

It would be interesting to be able to look at the simultaneous presence of both the GSTM1 null and NAT2 slow acetylator genotypes, smoking, and occupational exposure. The sample size for such a study would probably need to be almost 5000 cases and a similar number of controls, considering for example the study of a simple two-way interaction between smoking and the combination of GSTM1 and NAT2 genotype, assuming an interaction effect of 1.5. Before engaging in such an effort, an effective preventive strategy targeting more susceptible individuals should be in place, together with efficacious preventive measures against smoking in the general population.

*Emanuela Taioli, Sara Raimondi
Fondazione Policlinico IRCCS, Molecular and Genetic Epidemiology Unit, 20122 Milan, Italy
taioli@policlinico.mi.it

We declare that we have no conflict of interest.


### Table: Odds ratios and 95% CI for association between several types of cancer and genetic polymorphisms

<table>
<thead>
<tr>
<th>Cancer Type</th>
<th>Gene</th>
<th>Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>GSTM1 A7</td>
<td>1.15 (0.95–1.39)</td>
</tr>
<tr>
<td></td>
<td>NAT2</td>
<td>1.09 (0.99–1.21)</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>0.92 (0.76–1.11)</td>
</tr>
<tr>
<td></td>
<td>mEH</td>
<td>1.00 (0.82–1.21)</td>
</tr>
<tr>
<td></td>
<td>GSTP1</td>
<td>1.04 (0.67–1.61)</td>
</tr>
<tr>
<td>Breast</td>
<td>GSTM1 A7</td>
<td>1.11 (0.87–1.41)</td>
</tr>
<tr>
<td></td>
<td>NAT2</td>
<td>1.11 (0.87–1.41)</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>0.93 (0.62–1.38)</td>
</tr>
<tr>
<td></td>
<td>mEH</td>
<td>1.15 (0.86–1.53)</td>
</tr>
<tr>
<td>Head and neck</td>
<td>GSTM1 A7</td>
<td>1.25 (1.00–1.57)</td>
</tr>
<tr>
<td></td>
<td>NAT2</td>
<td>1.25 (1.00–1.57)</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>1.15 (0.86–1.53)</td>
</tr>
<tr>
<td></td>
<td>mEH</td>
<td>1.15 (0.86–1.53)</td>
</tr>
<tr>
<td>Colorectal</td>
<td>GSTM1 A7</td>
<td>0.98 (0.75–1.29)</td>
</tr>
<tr>
<td></td>
<td>NAT2</td>
<td>0.98 (0.75–1.29)</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>1.15 (0.86–1.53)</td>
</tr>
<tr>
<td></td>
<td>mEH</td>
<td>1.15 (0.86–1.53)</td>
</tr>
<tr>
<td>Bladder</td>
<td>GSTM1 A7</td>
<td>1.20 (1.20–1.64)</td>
</tr>
<tr>
<td></td>
<td>NAT2</td>
<td>1.20 (1.20–1.64)</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>0.98 (0.75–1.29)</td>
</tr>
<tr>
<td></td>
<td>mEH</td>
<td>1.15 (0.86–1.53)</td>
</tr>
</tbody>
</table>

CYP1A1 =cytochrome P450A1, GSTT1=glutathione S-transferase T1, GSTP1=glutathione S-transferase P1, mEH=microsomal epoxide hydrolase.
Comment


Neonatal hearing screens: Wessex re-visited

In today’s Lancet, Colin Kennedy and colleagues¹ provide a timely follow-up to the 1993–96 Wessex birth cohort,¹ which evaluated the effectiveness of neonatal hearing screens for detecting deafness in babies compared with the effectiveness of detecting deafness at a later date via the Health Visitor Distraction Test or other means. Kennedy and colleagues reviewed a group of children aged between 7 and 9 years who had undergone electrophysiological screens of hearing soon after birth, and compared them with a group who had only the Health Visitor Distraction Test at 7–8 months of age. Importantly, we note a substantial number of babies diagnosed as deaf before 6 months of age in the non-screened group. This difference might reflect increased awareness of deafness in young infants generated by the research project itself, in health-care workers, and the wider community. However, despite this factor, there remains an impressive, statistically significant difference in numbers detected by early screening.

We look forward to another chapter in Kennedy and colleagues’ study, hopefully evaluating the language, educational, and social outcomes for deaf children who were identified early. However, it is quite disconcerting that despite the big increase achieved by the researchers in early identification of deaf infants (before 6 months of age), the management (usually involving the fitting of hearing aids and starting early intervention) did not occur in roughly half of cases until the child was over 18 months of age. We gather that this situation was being rectified after 1996, but the work of Yoshinaga-Itano et al¹ and Moeller¹ suggests that delay in appropriate management translates into less benefit for those children who are managed late.

One of the fears that besets workers in hearing screens is the substantial number of deaf babies who “pass” neonatal screens, but who emerge later on and might be diagnosed very late because of a false sense of security generated in their families and even their physicians by the apparently normal neonatal screen.

Are false negatives due to a problem in the instrumentation or the screener? Was the deafness truly acquired after the screen? Kennedy and colleagues reflect on this question, and using neonatal screening data, Health Visitor Distraction Test data, and audiometric information about children whose hearing has clearly deteriorated over time, they suggest that a significant number of deaf children acquire their hearing loss after the neonatal period, perhaps as many as one quarter. This estimate is supported by other research,¹ and is a disturbing finding because it suggests that in an ideal programme there should be further measures to screen hearing in preschools or elementary schools. This additional cost might not be accepted happily by those who control the purse-strings of the public-health budget.

Why do a substantial number of children turn out to have hearing loss in childhood (after neonatal screens)? In our experience in Sydney, major causes contributing to this group are: intra-uterine infection with cytomegalovirus; inner ear malformations, such as large vestibular aqueduct syndrome; and progressive genetic hearing losses. Whereas we might have a history to warn us with the last group, we rarely do in the first two conditions.

Kennedy and colleagues have accomplished an important first step in persuading us that hearing screens are an important public-health measure. However, just as screening of hearing in infants has some challenges, parents face challenges in adjusting to the diagnosis of hearing loss and the recommendation to fit hearing aids in an apparently perfect infant. The potential for a family to reject the results of a computerised test of hearing in
their “perfect” baby can be very challenging for workers in the field. Although lack of speech, poor vocabulary, articulation, and grammatical errors in children (even with mild hearing losses) are well known to experienced workers, initial reluctance of parents to accept hearing aids (and sometimes even the diagnosis itself) is not uncommon. There are cultural biases about hearing aids, and indeed other rehabilitative devices, in some ethnic community groups, where the admission of disability in a young child might bring shame on a family (Mok C, Ibrahim R, Royal Institute for Deaf and Blind Children, North Rocks, New South Wales, Australia, personal communication). Previous experiences and lack of access to unbiased information might also affect a family’s willingness to consider sign language as a communication medium for their child, and to use it in public.

Even in developed countries there might be negative perceptions of deaf people that temporarily overwhelm the parents of the deaf infant. Why is this so? It is interesting to speculate why a small device placed in or behind the ear is so much more unacceptable than a larger more obvious device placed in front of the eyes (spectacles). We might also ask why the community is often less sympathetic to the deaf child than to the blind child with a guide dog or white cane.

Screening programmes are about saving health dollars and preventing disability and suffering.6,7 Hand in hand with providing optimum detection, optimum hearing aids (or implants), and optimum teaching of spoken and/or sign language for the deaf child, we need to be promoting the image of the deaf achiever. Hearing children and adults need to be impressed by the many successful deaf people who can be role models for so many. These deaf achievers offer the most effective persuasion for the financial and humanitarian gains made by universal screening.

*Patricia Mutton, Kenneth Peacock
Deafness Centre, Children’s Hospital at Westmead, Westmead, NSW 2145, Australia
patrim@chw.edu.au

We declare that we have no conflict of interest.


Role of the global civil society

The concept of health promotion as a public-health policy is of surprisingly recent vintage. Emerging in the mid-1970s, it was enthusiastically embraced at a 1978 conference sponsored by WHO and UNICEF at Alma Ata, Kazakhstan. There, almost all of the 134 nations represented reaffirmed health as a fundamental human right, and endorsed an ambitious new Declaration of “Health for all by 2000”. The Alma Ata conference is noteworthy for its recognition that health is a societal responsibility, but the great difficulty of providing primary health care to the extremely poor soon became apparent.1 Although world gross domestic product more than tripled between 1978 ($US 9·1 trillion, $2129 a head) and 2000 (31·5 trillion, $5198 a head), at the turn of the millennium there were still more than a billion people living on less than a dollar a day, with limited or no access to health care, clean water, or sanitation. Gross national income in then low-income countries, which account for 40% of the world’s population, was $430 a year, compared with $26 310 in high-income countries.2 This appalling disparity illuminates the failure to live up to the Alma Ata consensus, and led to a new United Nations declaration in 2000, which included the Millennium Development Goals—the first of which is to halve extreme poverty and hunger by 2015.3 Progress towards reducing poverty is on target,3 but success will be remarkably variable in different world regions. China and India are experiencing rapid economic growth, but the number of people living in abject poverty in the poorest African countries is increasing rapidly. Present rates of progress for health and education are well behind those required to reach the 2015 targets.4 One can only hope that recent G8 deci-
sions about debt relief and aid to poor African countries will improve the situation, but, as pointed out by Benjamin Loevinsohn and April Harding in today’s *Lancet*, money is not enough.

Although ill-health compounds poverty, it is not the root cause of poverty. Poverty, in turn, increases the prevalence of illness while decreasing access to healthcare. Thus, both poverty and disease need to be addressed simultaneously with political, economic, and social measures. Unfortunately, in low-income countries, which also carry the highest disease burdens, governments spend, on average, 1·1% of gross domestic product—less than $5 a head an annum—on health, and little of this goes to the poor. Even so, available funds could reduce mortality and morbidity rates if focused on prevalent health problems that can be effectively controlled by simple inexpensive measures applied to a high proportion of the population. Herein lies another problem: the enormous deficiencies about the quantity and quality of institutions, and of health-service providers in low-income countries, as well as their uneven distribution. Health will not be improved unless the necessary infrastructure for effective service delivery to the poorest populations can be developed. Service providers might include central or regional governments, universities, corporations, international organisations, private practitioners, or even community volunteers. Each can be effective, but success is highly dependent on local circumstances, and the knowledge and accountability of the providers. The potential for abuse is significantly increased when politics and profit enter into the equation.

One bright spot on the horizon is the enormous growth of what has been called global civil society, which is mostly comprised of non-governmental organisations (NGOs). Although estimates vary, there were no more than a hundred or so international NGOs in 1900. There are now tens of thousands, more than 2000 of which function as consultants to the United Nations (figure). NGOs—national or international—cover an enormous range of human activities, and are usually established by persons motivated to create, by advocacy or by the provision of services, needed “public goods” not normally generated by the for-profit marketplace. Many are dedicated to the empowerment of the underprivileged and vulnerable members of society. Their operating strategy relies heavily on cooperation, and NGOs often work together in networks, which greatly increases their chances of success.

Loevinsohn and Harding evaluated ten examples of providing of primary health and nutrition services in low-income countries via non-state contractors and concluded that this approach, which cost between $3 and $6 a head a year, could result in rapid, and in many cases, sustained improvements. Their findings relate only to contracts with NGOs and might not apply to other private sector elements. Engaging commercial entities or private practitioners directly would probably complicate the contractual process, while market-type competition will not necessarily improve the health benefit per dollar expended. Independent quality assurance is essential because the health-care consumer will rarely be able to judge or influence the quality of services. Further, for external for-profit providers, the incentive to transfer the full range of necessary skills and knowledge to the local population—an essential element of sustainability—would be small.

Because of their limited resources, the implementation of effective measures to control poverty, illiteracy, and disease in most developing countries will require external assistance. The record of governments in this regard is, in general, dismal, but the rapid and accelerating growth of a global civil society that Kofi Annan has referred to as the
world’s new superpower is likely to play an increasing role in ensuring that developing countries will eventually realise the potential of their greatest asset—their people.

Ian Magrath
International Network for Cancer Treatment and Research, INCTR at Institut Pasteur, B1180 Brussels, Belgium imagrath@inctr.be

I declare that I have no conflict of interest.


Does a NEAT difference in energy expenditure lead to obesity?

Disentangling the neuroendocrine systems that regulate energy homoeostasis and adiposity has been a long-standing challenge, especially with obesity as an increasingly important public-health problem. Energy homoeostasis has become a busy research topic; barely a week goes by without the spotlight falling on some new potential regulator of appetite or bodyweight. There is even a new journal devoted to metabolic research. Despite the growing understanding of the fundamental biology of energy balance, critical pathways have yet to be identified.

James Levine and colleagues recently showed that lean and obese individuals are different in the energy dedicated to postural habits. If obese individuals adopted the NEAT-type of lean individuals, they could potentially expend on average 350 kcal a day more energy in posture allocation. NEAT (non-exercise activity thermogenesis) encompasses, besides the calorie expenditure of activities of daily living, the energy costs of all non-volitional muscle activity, such as muscle tone, posture maintenance, and fidgeting. NEAT got off to a promising start when it was initially shown to confer resistance to weight gain during overfeeding. Changes in NEAT account for differences in fat storage between individuals, suggesting that as human beings overeat, NEAT dissipates excess energy to preserve leanness and that failure to activate NEAT may lead to weight gain. In thermogenesis, NEAT is the best predictor of interindividual differences in fat gain during overeating. Because thermogenesis only accounts for about 10% of total energy expenditure, the potential for a significant contribution in human beings is small but not negligible (figure 1, A).

The interesting finding of decreased NEAT in obese individuals, however, reveals the complexity of comparing the metabolic activity of individuals who differ greatly in body size, body composition, and other characteristics, which themselves affect energy expenditure. A larger size implies more fat-free mass with a greater metabolic activity devoted to cardiac output, body support, and other physiological processes influenced by body size. Similarly, obese individuals have more adipose tissue that is metabolically less active. If NEAT is calculated on a per-whole-body mass basis (kcal per day per kg bodyweight, as expressed by Levine and colleagues), obese individuals have 21% less energy consumption. However, if NEAT is expressed on a per-fat-free mass basis (kcal per day per kg fat-free mass, figure 1, B), there is no significant difference between lean and obese individuals (+5% in the obese group).

What are the pathophysiological implications of NEAT, and might these lead to a new conceptual strategy for weight control? Little is known about how NEAT is regulated. Does the potential contribution of NEAT during fluxes in energy balance operate through central mediators, hormonal modulators, peripheral signals, or even a mixture of those? In this context, orexin A, a neuromediator associated with arousal, increases NEAT in the short term. Plausible candidates for hormonal modulators of NEAT are thyroid...
hormones and leptin. To date, there is no direct evidence for a peripheral signal of NEAT. Every new finding generates new questions. There is no escaping the laws of thermodynamics. For example, are there obese individuals with increased NEAT who are overweight simply because their energy intake exceeds their energy expenditure? On the contrary, it seems possible to stay lean even without an increased NEAT. Environmental influences probably affect the amount of NEAT expended by individuals. More fascinating, however, will be to decipher how much NEAT is biologically regulated by disentangling the underlying mechanisms involved (figure 1, C). We can speculate about a specific contribution from sympathetic activation and endocrine changes in response to energy imbalance.

Total energy expenditure reflects the sum of three major components: basal metabolic rate, thermogenesis, and physical activity (figure 1, A). The basal metabolic rate is around 75% of total energy expenditure in sedentary people and is the energy required to maintain basic physiological functions, such as respiration, circulation, cellular homoeostasis, and tissue regeneration. About 80% of the interindividual variance in basal metabolic rate can be accounted for by age, fat-free mass, and sex. The remaining variance can be attributed to differences in activity in the sympathetic nervous system and skeletal muscle metabolism. In human beings, the freedom to do exercise and everyday physical activities accounts for further interindividual variability in this energy fraction. Physical work can be divided into weight-independent and weight-dependent activities. Consequently, the energy cost of weight-bearing exercises is highly related to body mass and is, therefore, much higher in obese individuals.

In modern civilised societies, mechanisation has replaced most manual labour. Therefore leisure-time physical activity is gaining a dominant role in determining energy expenditure. Although genetics certainly plays a role in determining a person’s energy homeostasis, the obesity epidemic is less the result of individual genetic propensities than the consequence of the shift to a sedentary lifestyle. Analysis of secular trends indicates that the primary causes of obesity lie in environmental or behavioural changes, because the escalating rates of obesity occur in a relatively constant gene pool and, hence, against a constant metabolic background.

Gema Frühbeck
Department of Endocrinology & Metabolic Research Laboratory, Clínica Universitaria de Navarra, University of Navarra, 31008-Pamplona, Spain
gfruhbeck@unav.es
I declare that I have no conflict of interest.

Figure: Energy expenditure and NEAT
A=major components of daily total energy expenditure in 70 kg man. B=NEAT differences between lean and obese individuals depending on way data are expressed. C=potential factors influencing NEAT.
Health research: what’s culture got to do with it?

In attempting to interpret racial and ethnic variation in statistical data, otherwise rigorous health researchers seem mysteriously drawn to a complex and nebulous concept: culture. “Culture” has recently emerged in health literature as a common default explanation, especially in research on certain groups, such as Mexican Americans, who are popularly thought to be particularly culture-driven.

For example, Bridget Grant and co-workers1 recently examined data from a national survey of psychiatric disorders in foreign-born and US-born Mexican Americans and non-Hispanic whites. In an innovative approach, their analysis considers immigration status, producing interesting and unexpected findings. They report that immigrants, both Mexican and non-Hispanic white, have lower rates of psychiatric disorders than do non-immigrants, and that US-born Mexican Americans have lower rates than US-born non-Hispanic whites. To interpret these correlations, they follow what has become a popular trend in health research: they suggest that traditional Mexican culture somehow protects people of Mexican descent from ill-health. Specifically, they propose that because the traditional Mexican family is close-knit and supportive, it may provide a buffer from mental illness. Without any measure or observation of family life or familial structure, they conclude that their findings point to “what appear to be the protective effects of culture and the deleterious effects of acculturation”. This is but one example of a common practice in current health research: to routinely attribute positive, negative, or neutral health-outcomes in ethnic minorities to traditional culture acting as a source of dysfunction or as a therapeutic panacea.2

Despite its popularity, such freewheeling applications of culture cannot be presumed to be innocuous nor justifiable. We examined the concept of “culture,” and more specifically “acculturation,” in current literature on Hispanic health.3 We found that data interpretation in these articles commonly invokes widely held cultural stereotypes about Hispanics to explain health status. These studies almost never include indicators of the specific cultural traits in question, but instead assume that by knowing someone’s ethnic identity or national origin, their beliefs and behaviours can reliably be inferred. But is this a reasonable assumption? What do we know, for example, about the family life of people who happened to choose “Mexican” or “non-Hispanic white” on a survey form? These groups are highly heterogeneous, and cultural beliefs and behaviours do not track well with ethnicity.4–6

Attributing variable health-outcomes to unexamined cultural differences requires several major leaps of faith. We must first assume the cultural trait is significantly different between the subgroups; next that it co-varies with the presence or absence of the disorder being studied; and finally that the trait can affect the outcome in question. Although perhaps plausible, accepting such an explanation brings us far afield from the study itself. It would seem more parsimonious to seek insight about the observed correlations among variables included in the study, thus minimising the number of assumptions made.

When the income and educational status of the groups studied are radically different, as is often the case when comparing ethnic minorities to non-Hispanic whites, one might begin by considering more carefully the effect of class status.7 For example, rather than asking what there is about being Mexican that keeps people healthy, one might consider what there is about the better-educated higher-income non-Hispanic whites that might explain their greater morbidity. Or could it be simply the effect of middle-class culture? Are people in this group more likely to have experience with psychotherapy and thus be more conversant with the psychometric instruments used, more hyper-vigilant about their mental states, and more willing to discuss them with strangers? Although such an exercise might seem plausible and interesting, the filament of reality slips gradually further...
Rethinking lipid mediators

One of the characteristics of the modern western diet is the imbalance of the two essential fatty acids, the omega-6 and the omega-3 fatty acids (figure). The most important omega-6 fatty acid is arachidonic acid (20:4n-6); the important omega-3 fatty acids are eicosapentaenoic acid (20:5n-3) and docosahexaenoic acid (22:6n-3). Today's western diet has a ratio of omega-6 to omega-3 fatty acids of 15:1, when data suggest that human beings evolved, and lived for most of their existence, on a ratio close to 1:1. Changes in eating habits and in agriculture within the past 100–200 years are the main cause of this changed ratio.1 This shift might be important because these fatty acids are the precursors of the eicosanoid lipid mediators: the prostaglandins (PGs) and leukotrienes, lipoxins, epoxyprostanoids and isoprostanes.

Most molecular studies of the fatty-acid-based mediator system have focused on arachidonic acid and elucidated several pathways of action. The cyclo-oxygenases catalyse the first step in PG synthesis, the generation of PGH₂ from arachidonic acid. These enzymes are the molecular target of non-steroidal anti-inflammatory drugs and thus highly relevant clinically.2 There are two different isoforms, with cyclo-oxygenase 1 expressed constitutively in many tissues and cyclo-oxygenase 2 expressed under physiological conditions in some tissues, but particularly in acute stress situations such as inflammation with resulting high levels of PGs.3

PGG₂ is then first metabolised to PGH₁, which in turn is metabolised to PGD₂, PGF₂, PGE₂, and PGI₂, and to thromboxane A₁ (TXA₁) by specific synthases depending on the cell and tissue. The resulting mediators bind to specific G-protein-coupled prostanoid receptors. These mediators and receptors are essential for the transduction of various signals, ranging from inflammatory responses to nociception, renal function, haemodynamics, and blood clotting.4 The same PGs can have distinct effects in different cell and tissue types.5 In particular, PGE₂, the only PG so far for which four receptor subtypes have been described, can have hugely varying effects depending on which of its receptors it is bound to.6 Furthermore, the isoprostanes and epoxyprostanoids, generated by oxidation, might act as ligands at prostanoid receptors.7

Another series of lipid mediators generated from arachidonic acid are the leukotrienes, which are formed by lipoxygenases in many cells and tissues. Leukotriene B₄ (LTB₄), acting through its specific G-protein-coupled receptor,
mediates chemotaxis. LTC₄, LTD₄, and LTE₄, via their G-protein-coupled receptors, mediate bronchial smooth-muscle constriction, mucus production, and submucosal oedema. Leukotrienes are thus highly potent proinflammatory mediators. An exciting twist in this mediator field is the discovery of lipoxins, mediators arising from cell-cell interaction and the sequential transformation by different lipooxygenases. Leucocyte 5-lipoxygenase generates LTA₄ from arachidonic acid, which is then transformed to the lipoxin LXA₄ in platelets by the oxidase activity of their 12-lipoxygenase. Lipoxins have potent anti-inflammatory and inflammation-resolving properties, including the inhibition of inflammatory cytokine formation and immune cell proliferation and migration. These lipoxin pathways also offer a new explanation for the anti-inflammatory action of aspirin: the acetylation of cyclo-oxygenase 2 enables it to act as a lipooxygenase, forming the lipoxin precursor 15-hydroxyeicosatetraenoic acid from arachidonic acid, which is then transformed by leucocyte 5-lipoxygenase to 15-epi-LXA₄ or 15-epi-LXB₄, the so-called aspirin-triggered lipoxins. These aspirin-triggered lipoxins seem to be more potent anti-inflammatory compounds than their conventional counterpart LXA₄.

But if human beings largely evolved on a 1:1 ratio of omega-3 to omega-6 fatty acids, the lipid-mediator system might be dependent on the presence of omega-3 and omega-6 fatty acids in roughly equal amounts. The beneficial aspects of supplementation with omega-3 fatty acid, particularly for its antithrombotic, cardioprotective, and anti-inflammatory effects, have been widely studied and publicised. Most of the effects caused by such supplementation have been attributed to: the inhibition of the metabolism of arachidonic acid to proinflammatory mediators; by competition for the same enzymes, the formation of the less bioactive 3-series PGs and thromboxanes and the 5-series leukotrienes from eicosapentaenoic acid; and to direct effects of long-chain polyunsaturated fatty acids with cell proteins.

Recently, several studies have looked at metabolites derived from omega-3 fatty acids as bioactive mediators in their own right. By using lipidomics (chromatography combined with mass spectroscopy), Yang et al. found that addition of physiological concentrations of eicosapentaenoic acid to a lung cancer cell line rapidly changed the major PG produced from PGE₂ to PGE₃, with a significant inhibitory effect on proliferation. Furthermore, in a cell-free system, cyclo-oxygenase 2 rapidly metabolised eicosapentaenoic acid, leading to the formation of PGE₃ instead of PGE₂. PGE₃ has also been attributed with less inflammatory and mitogenic potential than PGE₂.

Charles Serhan used lipidomics to study the role and significance of metabolites of docosahexaenoic acid, the 22-carbon omega-3 fatty acid, and eicosapentaenoic acid, the 20-carbon compound, in inflammation. Two studies showed generation of potent anti-inflammatory mediators metabolised from docosahexaenoic acid by acetylated cyclo-oxygenase 2 and 5-lipoxygenase, in analogy to the aspirin-triggered lipoxins formed from arachidonic acid. The omega-3-derived mediators have been implicated in the resolution of inflammation and called resolvins. Docosahexaenoic acid metabolites comprise several resolvins (RvD1 to RvD6) and the docosatrienes. The effect of these mediators could be confirmed in vitro and in vivo. The resolvin formed from eicosapentaenoic acid is RvE1. Recently, Makoto Arita and colleagues (in Serhan’s group) identified a specific receptor and mechanism of action for RvE1 and showed RvE1 generation and anti-inflammatory and antimigratory
action in vivo.\textsuperscript{16-19} RvE1 inhibited activation of nuclear factor κB by tumour necrosis factor α. Further experiments identified the orphan receptor ChemR23 as a specific G-protein-coupled receptor for RvE1, and found transcription of ChemR23 in cardiovascular, gastrointestinal, renal, brain, and myeloid tissue. In siRNA knockdown experiments, downregulation of ChemR23 disrupted RvE1-induced inhibition of interleukin-12 generation in dendritic cells. ChemR23 shares homology with the receptor identified for the arachidonic-acid-derived and aspirin-triggered lipoxins, but is molecularly distinct from this receptor. Arita and colleagues have thus identified an omega-3 metabolite-specific receptor with potent anti-inflammatory function: RvE1 is the first omega-3 metabolite for which a specific receptor pathway of action, distinct from that for arachidonic-acid-derived mediators, has been identified.

Omega-6 arachidonic acid is thus not the only precursor of interest to generate bioactive lipid mediators; the omega-3 fatty acids, eicosapentaenoic acid and -docosahexaenoic acid, are also important precursors. The field should focus now on the relative contributions of omega-3 and omega-6 fatty acids, and on modulation rather than total inhibition of mediator formation by inhibitors of cyclo-oxygenases. Lipidomics will be valuable for the molecular identification of signalling molecules. Combining these approaches with molecular biology methods, such as siRNA, should identify more protein targets of lipid-mediator pathways.

The complexity and the importance of the lipid-mediator system in physiology and pathology highlight the potential effect of diet and nutrition on human health, because diet is the source of the precursors, and determines, to a great extent, the concentration or availability of each mediator.

To fully understand the relative contributions of omega-3 and omega-6 fatty acids in lipid-mediator cascades, several steps are now necessary: more mediators derived from omega-3 fatty acids need to be identified and characterised; the synthesis pathways for these mediators and their regulation need to be analysed; the molecular targets of the mediators need to be understood; and the effect of non-steroidal anti-inflammatory drugs and others on these pathways need to be analysed.

*Karsten H Weylandt, Jing X Kang
Department of Gastroenterology and Hepatology, Charité University Hospital, Campus Virchow, 13352 Berlin, Germany (KHW); and Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts, USA (JXK)
karsten.weylandt@charite.de

KHW declares that he has no conflict of interest. JXK is the inventor and co-applicant (with Massachusetts General Hospital) of a US patent application (US 10/468318; PCT/US02/07649) on fat-1 transgenic mice. If granted, Massachusetts General Hospital will own the patent. He had to make this application as a condition of his employment, has received no income from the invention, and, should he do so, any such income will not be used for personal gain.

12 Racke MK, Leaf A. Evidence that free polyunsaturated fatty acids modify Na+ channels by directly binding to the channel protein. Proc Natl Acad Sci USA 1996; 93: 3542–56
15 Serhan CN. Novel omega-3-derived local mediators in anti-inflammation and resolution. Pharmacol Ther 2005; 105: 7–21
19 Aita M, Yoshida M, Hong S, et al. Resolvin E1, an endogenous lipid mediator derived from omega-3 eicosapentaenoic acid, protects against 2,4,6-trinitrobenzene sulfonic acid-induced colitis. Proc Natl Acad Sci USA 2005; 102: 7671–76
While current EU president Tony Blair, prime minister of the UK, calls for increased aid to African countries, the fate of a major EU grant for Uganda’s war-torn northern region is highlighting how procedural requirements can stand in the way of donors’ good intentions.

The project at issue is the Improving Sexual and Reproductive Health in Uganda (ISRH) initiative, launched in 2000 as a 5-year, €8 million programme to serve northern Uganda. Devastated by the ongoing war waged by the Lord’s Resistance Army rebel group, the region has some of the highest HIV prevalence rates in the country.

By many measures, the project has been a great success. Despite the area’s chronic volatility, demand for services—including HIV testing and counselling, family planning, and care for HIV-infected individuals—far exceeded initial projections. One partner, the AIDS Information Center (AIC), anticipated administering 60,000 HIV tests per year; the actual figure in 2004 exceeded 130,000. 6 months after establishing an HIV/AIDS care centre in Gulu, a major northern town in early 2005, another partner, The AIDS Service Organization (TASO), was serving seven times more clients than anticipated.

However, there have also been problems, particularly with release of funds. Although the project was scheduled to start in January, 2001, grantees did not receive funds until July, 2001. Since then, there have been further delays, which according to Maxwell Abok, Financial Manager for IRSH, are partly a result of the many links in the chain of command—including the Ugandan Ministry of Finance, the Ugandan office of the EU delegation, and the Brussels headquarters.

As of July 2005, the EU had released funds for just 35–40 months of the 60-month project, despite the fact that partners operated their services continuously. To maintain operations during the unfunded months, the implementing agencies scaled back activities, and borrowed from other donors or their general budgets. After 16 AIC employees working on the IRSH project went unpaid for several months, AIC asked another donor to permanently cover their salaries, says AIC executive director Charles Hitimana.

As IRSH enters its final months, some of the grantees fear they will lose many thousands of euros in unrecovered operating expenses, since current EU regulations do not allow for retrospective reimbursements. To date, TASO and AIC have not received any funds at all for 2005. Marie Stopes Uganda did not receive its funds until April, 2005. €3 million—roughly 40% of the budget—remains unspent, Abok says.

In addition, some of the partners have been asked to refund money to the EU. While the EU says that a portion of the funds at issue are “unjustified” expenses, Geoffrey Byamugisha, IRSH’s independent auditor at Ernst & Young, says that most of the money was spent to meet the overwhelming demand for services.

TASO, for example, has been asked to repay more than €50,000 spent on stocking and running its Gulu centre. AIC faced questions about transport outlays when it opted for air travel rather than ground transport due to safety concerns.

EU regulations require that grant recipients obtain advance approval for over-spending on specific budget lines, says Byamugisha. By their own admission, the organisations did not always comply, maintaining that the approval process was slow, and a rapid response was needed. “We agree we didn’t follow procedures”, says Hitimana of AIC. But he nonetheless argues that the EU should be flexible and recognise that “we didn’t misuse the money”.

Other IRSH partners agree that their experience is an object lesson in how strict adherence to the fine print can undermine the ambitious plans for aid to Africa laid out by Blair and other world leaders.

“They’re talking about aid efficiency, and they’re talking about corruption, but what could be worse than a work programme that starts in January, and doesn’t get funded until July?” asks Mark Blackett, former Marie Stopes Uganda programme director, who oversaw much of its participation in IRSH.

Emily Bass

NGOs working to improve sexual health in Uganda have been plunged into financial difficulty by delays that held up millions of euros of EU aid money. To make matters worse, the EU now wants some of the money back, citing wilful violations of its strict spending rules. Emily Bass reports.
When José Luis Velasquez, a cheerful 24-year-old health educator with a Catholic Church-sponsored HIV prevention programme, visits schools and youth groups to talk about HIV prevention, he is sometimes asked by the organisers not to talk about sex—even to not mention that HIV can be transmitted sexually. But, when it is time for questions, he says, what young people want to know most about are condoms, contraceptives, abortions, and sex.

Velasquez, a Catholic, is the education coordinator for Proyecto Vida, or Project Life, an HIV/AIDS prevention and treatment programme started by two Maryknoll nuns. The project is based in the town of Coatepeque, in southwestern Guatemala, along a main truck route running north to Mexico.

Because the project falls under supervision of the Guatemalan Catholic Church, Velasquez and his colleagues have to tread carefully when it comes to condoms, contraceptives, and other sex topics considered taboo by the Catholic Church. In his talks, Velasquez always emphasises that abstinence and fidelity are the best ways to avoid infection and that condoms should be a last resort.

Most priests and other religious leaders, he says, have gradually come to accept that it is necessary to discuss condoms and other subjects they consider controversial in HIV/AIDS education.

They allow it, just as long as they’re not involved, he says. “They support us with their silence. They’ll say ‘OK, tell me when, and I’ll leave the room’.”

Though the Church officially condemns contraceptives, abortion, and homosexuality, many Catholics and clergy in Latin America are quietly dissenting from official church policy. In Guatemala and other Latin American countries with majority Catholic populations, the use of birth control is on the rise and national reproductive health and education programmes are spreading. Many of these programmes have the Church hierarchy’s blessing, or at least its tacit acceptance.

Experts say there is also a growing recognition within the Church that HIV/AIDS prevention programmes need to work with at-risk populations, such as homosexuals and sex workers.

Nevertheless, there remains enough Church opposition to hamper efforts to reduce maternal mortality, prevent HIV/AIDS, and improve family health, say activists. “The Church has interfered a lot in what could have been, in terms of prevention of the epidemic”, says Richard Stern, a Costa Rica-based HIV/AIDS activist. He says it is a tragedy that the Church has not done more to help prevent the spread of the virus.

But some Church leaders say it is time to stop waging ideological debates and start preventing HIV/AIDS. “Fighting has done us so much harm that we’re trying to have a greater dialogue”, says Bishop Pablo Vizcaino, head of the Guatemalan Church’s pastoral health programme.

Judging by the high rates of contraceptive use in many staunchly Catholic countries, a large number of Catholics are ignoring the Church’s teachings against birth control. In Brazil and Colombia, for example, both strongholds of the Catholic Church in Latin America, over 75% of women use contraceptives.

Despite widespread rejection of Church policy, activists say there is little chance official Church doctrine will change, especially since the election of conservative Joseph Ratzinger to the Papacy. As head of the Vatican’s Congregation for the Doctrine of the Faith since 1981, Cardinal Ratzinger, now Pope Benedict XVI, maintained a hard line against birth control, abortion, and homosexuality. On Ratzinger’s election, the Latin American Network of Catholics for Choice, which lobbies for the Church’s acceptance of contraception, abortion, and homosexuality, declared that “by electing him, our Cardinals didn’t take into account the millions of parishioners who suffer and the thousands who have died because of his teachings”.

Despite the Vatican’s rigid condemnation of condoms as birth control, its top health official, Cardinal Javier Lozano Barragán, has conceded that a non-HIV infected wife could ask an infected husband to use a condom to prevent the spread of the virus.
It is estimated that some 70 000 people are infected with HIV in Guatemala, one of the highest rates in Central America. Though Church leaders here tow the official line, they often turn a blind eye to the HIV-prevention efforts of some clergy, such as handing out condoms to sex workers and homosexual men. Nevertheless, neither the leaders nor the rebels want to publicise this quiet rebellion.

"Officially, we don't know what they're doing", says Conchita Reyes, head of the Church's network of health programmes in Guatemala, when asked about a specific HIV-prevention programme. Abstinence and fidelity are the Church’s official recommendations for preventing HIV/AIDS and other sexually transmitted diseases, she says.

When asked about birth control, Reyes said that although the Guatemalan Church’s health programme only recommends natural methods approved by the Vatican, it provides information on all types of contraceptives. "We make it known that people have the right to decide, as long as they are well informed", says Reyes. Those individuals who want to use artificial birth control will do so regardless of what the Church recommends, she adds.

In fact, most experts in Latin America say religious beliefs are not usually people's first concern when it comes to making decisions about family planning. "Very few people don't use contraceptives because of God", says Dalila de la Cruz, education coordinator at APROFAM, the biggest provider of reproductive health care and education in Guatemala.

De la Cruz says the reasons most women don't use birth control methods are that they don't know about them, have no access to them, or are concerned about possible harmful side-effects. According to the 2002 National Maternal Infant Health Survey, only 43% of Guatemalan women and men use contraceptives, up from 31.5% in 1995.

Rolando Figueroa, a gynaecologist and coordinator of the health programme for Catholic Relief Services in Latin America and the Caribbean, says Church opposition to reproductive health services was a result of Church leaders’ views that reproductive health meant contraceptives, sexuality, and abortion. But some Catholic organisations are trying to change that stance by moving the debate away from taboo topics, Figueroa says. For example, some church programmes that will not discuss contraceptives have begun to recommend spacing between pregnancies using natural methods. "It’s a step", he says.

Alejandro Silva, director of Guatemala’s National Reproductive Health Program, says that since the Catholic Church helped formulate the Social Development Law in 2001, which created the National Reproductive Health Program, there have been few major conflicts with the Church on public reproductive health policies, including family planning and sexual education.

But Silva says issues like emergency contraception and homosexuality are still too touchy to be included in public-health initiatives, for fear of a backlash that could hold up other reproductive health programmes.

He added that there are more ominous barriers to good reproductive health in Guatemala than conflicts over religious beliefs. These include lack of education, under funding of the health system, and even biases among health workers.

According to Silva, some doctors will not prescribe contraceptives to teenagers, while others recommend methods of contraception based on their personal religious beliefs, not science.

Some activists say the Catholic Church's official stance against condoms and homosexuality is still hurting AIDS-prevention efforts. Nevertheless, the spread of Catholic-run programmes that are addressing HIV/AIDS, like Project Life, may indicate a change.

"When we first started, the Church was very slow in responding because it’s taboo to talk about sex", says Dee Smith, a Maryknoll nun and co-founder of Project Life. "But now that so many people have died in the parishes, they're becoming much more helpful."

Smith, along with Project Life co-founder Marlene Condon, believes the energy wasted on the debate over condoms and the Church could be better used to fight the epidemic. "Making everything controversial does not help the fight against AIDS", says Smith.

Jill Replogle
Swiss hospital investigates heart transplant “experiment”

The death of a Swiss woman following a transplant operation in which she was given a mismatched heart has forced a Zurich hospital to suspend all transplant services—and launch an investigation into newspaper claims that the operation was a “medical experiment”. Bojan Pancevski reports.

Heart transplants have been temporarily banned at a top Zurich hospital while investigations continue into an operation in which a patient was transplanted a heart of a different blood group, and died within days of the operation.

Rosemarie Voser, who was 57 at the time of her death in April last year, was transplanted a heart from a patient with blood group A, despite the fact that she had the incompatible O blood group, at the Zurich University Hospital after spending a year on an organ-donation waiting list. She died 3 days after the operation.

Soon after, reports in Swiss media appeared citing unnamed sources from the hospital claiming that Marko Turina, 67, a professor and one of the country’s top heart specialists, had approved the operation as a medical experiment.

Two chief surgeons have been temporarily suspended and the state prosecutor has started an investigation into the case following the allegations. When contacted by The Lancet the hospital refused to comment on the case; Swiss state prosecutors also refused to give out details of their investigation.

Chief prosecutor Ulrich Weder said: “This is a complex case and the investigation is likely to go on for several months.”

Investigative reporters at the Neue Zuercher Zeitung, one of Switzerland’s most respected newspapers, quoted sources from Turina’s team, which did the operation, speaking on condition of anonymity, saying that Turina, driven by professional ambition, had ordered the operation to go ahead with the aim of “breaking medical taboos”.

The newspaper’s sources claimed the chief surgeon, named only as Dr Andreas K for legal reasons, suggested to the senior physician Dr Oliver R, again not fully named for legal reasons, that the A blood group heart be transplanted to Ms Voser.

Oliver R then called in Turina, who “after thinking for several seconds” decided to go ahead with the operation. According to the sources, at least four people who participated in the operation were aware of the blood group mismatch.

When Professor Turina could not get the implanted heart to start beating, “it was clear for all present that something was not right”, an anonymous doctor was quoted as saying.

Ms Voser was then connected to a respirator and artificial heart machine and put into an induced coma. She died 3 days later despite efforts to track down another heart for her.

All the surgeons involved in the operation have denied the allegations. Professor Turina has however admitted he received a call from his colleague on duty at the clinic at 0400 h on the morning of the operation to notify him that there was a donor heart ready for transplantation, but claimed he misunderstood the blood group of the donor because he was not “wide awake”.

In an interview for Swiss TV he said: “It was a hard blow and I am terribly sorry. It preoccupies me even today. But life goes on. The most important thing is to learn from it.”

He took voluntary retirement shortly after the first media reports on the incident came out.

The incident and subsequent allegations have left the hospital’s transplant department severely shaken, staff have admitted.

Immediately after Ms Voser’s death transplants were suspended briefly while transplantation procedures were changed. Future decisions on whether to go ahead with any operation were to be made after consultation between a surgeon and cardiologist who would then have to keep an entire operational team informed.

But last month transplants were again suspended and hospital officials have said two doctors involved in Ms Voser’s operation were also “relieved of their operational duty”.

Hospital director Christiane Roth, explained: “Some of our employees are having a hard time and there is a great deal of insecurity. This temporary ban should bring back peace and normality to everyday work. It is a chance for the clinic to continue its good work.”

Bojan Pancevski
Book

Drawing the line at the end of life?

At one end of the line lies an act; at the other, an omission. The students’ task is to plot, somewhere on the line, a range of situations. They have been told about James Rachels’ account of the bathtub, two cousins, and the inheritance (N Engl J Med 1975; 292: 78–80). One cousin enters the bathroom where his relative is bathing, and holds the child’s head under the water. With the child dead, the inheritance is his. In a parallel universe, the cousin enters with the same end in mind but discovers that the child has already slipped and fallen unconscious. He refrains from rescuing the child; the child drowns, he inherits.

Impressed with Rachels’ logic, the students appreciate that an act and omission can achieve the same result and might even be seen as morally equivalent, at least where the underlying motivation and intention are identical. But these are not (yet) the issues we want to explore: what we want to establish is when we can apply the labels “act” and “omission”, before moving on to consider whether the conduct in question should be judged good or bad.

We start with the decision not to offer antibiotics to a patient. Some of the students begin to ask why this is happening but they are reminded that the whys and wherefores are not in issue: it may be a good or bad decision, but how do you categorise it? They agree that this is an omission. We talk about lethally injecting the patient. This is unanimously declared to be an act; at one end of the line lies an act; at the other, an omission. The students’ task is to plot, somewhere on the line, a range of situations. They have been told about James Rachels’ account of the bathtub, two cousins, and the inheritance (N Engl J Med 1975; 292: 78–80).

With some convictions quite visibly shaken, it falls to the lawyers and professional bodies to impose a measure of order: the best view, they say, is to place all but the lethal injection in the box marked “omission”. Some such omissions will be good; some will be bad; the lethal injection, however, is always bad. Most of the students, who will be doctors within 6 months, find comfort in this. But a few members of the group, no doubt convinced by Rachels’ thought experiment, complain that the distinction fails to make moral sense. It is this minority who will be most persuaded by Peggy Battin’s Ending Life: Ethics and the Way We Die. Yet, wherever one stands, there is more to be learnt from this book than from almost any of the countless others that seek to explore the issues that surround death and dying.

“. . .there is more to be learnt from this book than from almost any of the countless others that seek to explore the issues that surround death and dying”.

There are many dimensions to these issues and many ways of engaging with them. Amidst the usual philosophical argumentation and historical studies, Battin has placed works of “creative non-fiction” and fiction. That she has chosen to do so is both bold and startlingly obvious, for few philosophical treatises, no matter how eloquently argued, could hope to have the impact of a fictional story like that of “Robeck”. Battin’s tale tells of the final weeks in the lives of an elderly academic and his wife, and is deservedly placed at the heart of this collection. Like the female protagonist, Battin’s words—both in this story and in the more “academic” selections—have a grace and power all too rarely found in such collections.

These are, therefore, canny choices for inclusion in such a book, which comes a decade on from her previous collection, the seminal A Least Worst Death. Battin no doubt knows and surely intends (or does she merely foresee?) the impact they will have upon the reader, but she nowhere excludes or diminishes the contributions made by her opponents. In fact, that she appreciates such diversity and the fine nuances of the debates she tracks, fits well with Battin’s central theme, which is to allow the individual to make sense of these issues for himself or herself. That most popular of principles—respect for autonomy—forms the central strand running through these papers, albeit bolstered on occasion by appeals to mercy or humanity.

These convictions lead Battin to some familiar conclusions: if, for example, the choice is between an unwanted life of suffering and dependency on the one hand, and physician-assisted suicide on the other, the latter option can be right. However, unlike some contributors to these debates, she recognises also that one’s options are rarely so starkly dichotomous. Negotiation, compromise, simple discussion; there are many overlapping alternatives that lie between the two poles, and Battin does well in exploring such courses, both in abstract and, on occasion, in relation to clinical scenarios, as in the case of “Scott Ames”.

Furthermore, Battin sensibly acknowledges that her conclusions will not fit all cultures and contexts. Support for her arguments might, indeed, come from a North American patient—the “independent, confrontational, self-analyzing, do-it-yourself, authority-resisting patient”. Elsewhere such support may be rare or absent. We will, she writes, get a better appreciation of how to tackle death and dying,


See Viewpoint page 682
once we understand “the way we do it, the way they do it”. How then do (and should) we, in the UK, do it?

On first inspection we are moving in the direction of Battin’s Americans. Autonomy has taken hold in the minds of the specialists and the public alike, and we too no longer die in the home, but instead anticipate that our final moments will be spent in a care setting, invaded by machines and drugs. The House of Lords, the House of Keys, and Scottish Parliamentarians, among many others, have therefore begun to look again at arguments for relaxing the law, spurred on by reports that autonomous citizens have decided they want the “right to die”; indeed, the right to be helped to die. Although one must reserve a degree of scepticism about the accuracy and neutrality of the opinion polls, if this is what some people want, why not allow them to have it?

The answer (such as it is) might lie in a subtle transatlantic difference that Battin detects, not in her philosophical analysis, but in her fictional account of Mrs Robeck’s visit to England. “In the country with which she is familiar, fields are divided by a single rail fence; these are replaced when they fall, but here in the land of her ancestors the barriers between one field and the next show centuries of growth, decay, continued growth”. This metaphor hints at two aspects of the British approach to death and dying. Firstly, our ethical boundaries may have taken a beating but they can still be glimpsed: the line depicted at the outset of this essay, for example, continues to command support, not least among health professionals. Secondly, we might not be as atomised as the North Americans Battin depicts. And atomisation, Battin concedes, risks overlooking “various interpersonal and situational pressures”, particularly if assistance in dying were to be legalised.

Perhaps, then, we are simply not ready to accept Battin’s arguments, as powerful as they are. We must, however, continue to explore these issues and follow Battin in refusing to frame them in “all or nothing” terms. Battin is an expert guide in this exploration of the way we die and her insightful, original, and paradoxically life-affirming collection cannot be commended highly enough.

Richard Huxtable
R.Huxtable@bristol.ac.uk

In brief

Book  Cicely Saunders’ spirit

When I was visiting Malawi earlier this year, to assess a new project for assistance, I visited Mary, a young woman with three young children. Mary was dying, and although she was growing weaker she was free of pain. Her pain was controlled using the analgesic ladder, promoted throughout the palliative care world, and initially started by Dame Cicely Saunders. Without the modern methods of palliative care, researched and taught by Saunders, Mary’s suffering would have been much worse.

Saunders’ model of palliative care, which began in the UK at St Christopher’s Hospice, Sydenham, London, has now reached more than 100 countries worldwide. The struggles and joys of this remarkable woman are captured in David Clark’s engaging selection of her correspondence. Through these letters we can absorb the true spirit of love and caring as Saunders writes, seeking advice and discussing her many concerns, which are common to those of us who work in palliative care. Her words of wisdom give us a glimpse of the breadth of the areas she sought advice in as she pioneered the hospice model, which was to change the face of death and dying for so many.

Clark brings together the letters written by Saunders over her lifetime that were typed by her secretary or left as copies in the archives. We should remember, however, that much advice was given confidentially to colleagues and Saunders also sought advice verbally from many of those she worked alongside. For example, in her letter to Josefina Magno, founder of the International Hospice Institute, in 1990, she says “I really must give credit to Barbara McNulty and Mary Baines who started the first ever Hospice home service from St Christopher’s”. But apart from this reference, there is no further mention of these two pioneers in the letters, because Saunders would have discussed issues with them in person. It would be important to harvest experiences from those still alive who can remember advice, incidents, and Saunders’ many quips that so often hit the nail on the head. Since the letters are arranged in chronological order by date, it would have been useful to have a thematic index where selected items of her words of wisdom could be accessed more easily.

I would recommend this book to anyone involved in hospice work, be they health professionals or affected families, as a reference to the hospice spirit and how we need to go the extra mile for the comfort of a patient at this special time of life. Clark has become the preserver of the legacy of one of the greatest women of the past century. This is our inheritance.

Anne Merriman
merriman@utlonline.co.ug
Fidgeting

The etymology of fidgeting remains obscure, but the way its meaning has changed is symbolic of wider changes in the practice of medicine during the past 200 years. In its earliest 17th-century uses, fidgeting was seen as something one did to others: it was a way of worrying or needling people into a state of discomfort. Two centuries later, fidgeting had become a way that we display or betray our own inner worries to the wider world. In his 1894 Harveian Society Lectures on The Common Neuroses, James Goodhart complained that his contemporaries had succumbed to an outbreak of fidgets and saw this as evidence of the declining strength of English character. Just over a decade later, Sigmund Freud could interpret his patient Dora’s persistent fidgeting with her handbag as a sign of her obsessive sexual interest in her family and friends. Fidgeting was no longer an irritating practice, rather it was a sign of a pathological inner state.

Advances in pharmacology and biophysics from the 1930s undid the psychological significance of fidgeting. The rise of amphetamine treatments for juvenile problems led to the rejection of the old psychoanalytic model of the “emotionally disturbed” child. In its place appeared “hyperkinetic syndrome” identified by American paediatrician Maurice Lauffer in 1957, in which fidgeting became the defining characteristic of the illness in its own right. And, of course, the emergence of this syndrome was bound up with the appearance of new drug treatments, such as Ritalin (methylphenidate) in 1955. As this new market emerged the fidget moved from being a simple behavioural sign to become a product. Today fidgets even appear as rubber toys or trinkets for those with attention deficit disorder and attention deficit hyperactivity disorder as practical aids to stimulate attention.

Fidget has moved full circle—from a way of engaging attention to become a sign of the failure of attention, and now reappearing as a tool for managing our attention. And, of course, its therapeutic potential is not limited to mental health. In 1999, James Levene and his colleagues at the Mayo Clinic, announced that fidgeting was an effective strategy in the pursuit of weight loss. An irritating habit has now become a new technique in the pursuit of psychological and bodily perfection.

Rhodri Hayward
Wellcome Trust Centre for the History of Medicine at UCL
rhodri.hayward@ucl.ac.uk

Mayana Zatz
Professor of Genetics and Director of the Human Genome Research Center at the University of São Paulo, Brazil. She is committed to helping those with genetic disorders, especially progressive muscular dystrophies, and is President of the Brazilian Muscular Dystrophy Association and head of a research group in neuromuscular diseases.

What has been the greatest achievement of your career?
To be able to contribute to two teams: as a scientist at the University of São Paulo’s Human Genome Research Center and as a patients’ representative as the President of the Brazilian Muscular Dystrophy Association.

What do you think is the most over-hyped field of science or medicine at the moment?
Stem cells. I have been personally involved in the political battle to allow embryonic stem-cell research in Brazil. However, we now have to tell patients that there is still a long way to go from research to treatment.

And the most neglected?
The lack of humanity and personal interest in the relation between physicians and patients. Once I read a sentence that I never forgot: “Patients do not care as much about what you know as long as they know how much you care.”

What is the best piece of advice you have received, and from whom?
From my father: if you want something, fight for it.

How do you relax?
Jogging every morning. By the end of a run, a problem that seemed as big as an elephant has became the size of a small ant.

What is your greatest regret?
Not having had more children—I have a son and a daughter. They are the greatest result of recombinant DNA we may ever experience.

Do you believe there is an afterlife?
No. When my son was 5 years old, I told him I was going to take a nap because I felt sleepy. He replied that when I died I would sleep forever, so why sleep now? I think we should always keep this in mind and try to do as much as possible while we are alive.

What is your worst habit?
I’m always in a hurry.

What makes a good research mentor?
A good mentor is one who is provocative and gives freedom for research and development of new ideas, thus allowing “natural selection” to find the best scientific minds.

With which historical figure do you most identify?
Madame Marie Curie.
Dame Cicely Mary Strode Saunders


In 1947, while working as a lady almoner at Archway Hospital, London, UK, Cicely Saunders took care of a dying patient, David Tasma, who was a survivor of the Warsaw ghetto. She and Tasma began discussing how the care of those at the end of their lives might be improved, according to David Clark, a historian at Lancaster University, UK, who worked with Saunders to document the history of hospice care. Eventually, Saunders, who had originally trained as a nurse, went on to medical school after a thoracic surgeon she was working for told her “go and read medicine, it’s the doctors who desert the dying, and there’s so much more to be learned about pain”, she said in a 2003 interview.

After graduating in 1958, Saunders spent several years doing research and laying the groundwork for modern hospice care. When she began her research, “physicians thought that morphine was inevitably addictive and that it was not useful by mouth”, Nigel Sykes, now medical director of St Christopher’s Hospice, told The Lancet. “The patient had to ‘earn their morphine’. That meant that a lot of people were suffering a lot of pain. What she showed was that if you titrated it just against the pain response that it worked by mouth, and did not cause addiction, and also did not obtund people. That produced a huge change over time in the approach to pain management.”

“Over the years St Christopher’s has taught many about changing things in London. She wanted to change things for the whole country and for the whole world.”

“She certainly moved physicians forward in understanding that death is not a failure”, said Val Halamanadas, president of the US National Association for Home Care and Hospice. Within 20 years of St Christopher’s founding, palliative care was recognised as its own specialty in the UK. Saunders had no illusions that a single hospice in London would be able to serve every patient who needed its services, and she knew that end-of-life care should take place “as much as possible at home, and when that was impossible, in a controlled setting”, Halamanadas said. St Christopher’s established a home care unit in 1969 that became a model for others. She would remain as medical director of the hospice until 1985.

Saunders no longer did routine daily clinical work when Sykes arrived at St Christopher’s as a registrar 20 years ago, but she used to take turns in the weekend rota. “When you came back on Monday morning it could be a bit disconcerting”, Sykes said. “Your patients would say to you, I met this wonderful doctor. She came and spent time with me.” Saunders showed that health providers could discuss dying with patients. “One was aware right from that point that although she was no longer fully engaged in clinical work, she had not lost her touch”, Sykes said. Having trained as a nurse, almoner, and doctor, “she was a complete multiprofessional team in one person”, he said.

Religion was extremely important to Saunders, who made a daily practice of spiritual reading, contemplation, and prayer. “I think she found in that something very, very important that maintained her through a lot of difficult times both personally and professionally”, Clark said. In 1980, Queen Elizabeth II bestowed upon Saunders the title of Dame, one of many honours she was awarded throughout her life. Saunders is survived by two brothers, John Saunders and Christopher Saunders. Her husband of 15 years, the Polish artist Marian Bohusz-Sysko, died in 1995.

Ivan Oransky ivan-oransky@erols.com
Cerebral perfusion deficits in dysbaric illness

In 1989, The Lancet published a short paper by G H Adkisson and colleagues describing cerebral perfusion deficits in 23 divers with decompression sickness. Shortly after publication, we received correspondence pointing out concerns relating to methodological aspects of the study and to the lack of reporting of ethics approval and patients’ consent. In a reply, published with that letter, the corresponding author argued that ethics approval was not required because the “investigation was based on observations during patient management”. He wrote that patients did give their consent.

For several years, concerns about the validity of the study findings were raised, and these concerns eventually led to an investigation by a Preliminary Proceedings Committee of the UK’s General Medical Council (GMC). This investigation was completed in September, 2002. The committee found no grounds on which to proceed.

Following the conclusions of this committee, Peter Wilmshurst asked The Lancet again to pursue the matter. As a member of the Committee on Publication Ethics and as a signatory to its code of conduct for editors, we sought advice about how best to take these concerns forward. As a result of these deliberations, we are now bringing the conclusions of the GMC committee to readers’ attention. Peter Wilmshurst and all contactable authors who wished to comment on the GMC’s summary publish their reactions alongside this judgment.

The following is a copy of a letter to Peter Wilmshurst from the General Medical Council:

Dear Dr Wilmshurst,

I am writing to inform you that on 30 August 2002 the Preliminary Proceedings Committee considered allegations against Dr Revell, Dr Sykes, Dr Macleod, Dr Torok and Dr Pearson stemming from the publication of the article published in The Lancet in 1989 entitled “Cerebral Perfusion Deficits in Dysbaric Illness”. As you may be aware, this Committee’s role is to determine whether any case which has been referred to it ought to be referred for public inquiry by the Professional Conduct Committee. It is for the Committee to decide if, in its opinion, a question is raised whether a doctor has committed serious professional misconduct which may consequently raise issues about the doctor’s registration and fitness to continue to practise. In so deciding, it may consider (using due caution) whether the case against a doctor has any real prospect of being established.

Having considered all the circumstances of the case, the Committee did not consider that allegations against any of the doctors ought to be referred for such further inquiry. It was its view that there was no real prospect of establishing a finding of serious professional misconduct against any of the doctors concerned.

In reaching that decision, the Committee noted that the allegations against the doctors were in connection with a paper which had been published in The Lancet in 1989. The subject of the Lancet paper was dysbaric illness. The paper described the effects of HMPAO brain scanning on 28 divers with neurological decompression sickness and one patient with non-neurological decompression sickness.

It also noted that the papers contained a copy of an article by you which had been published in the British Medical Journal on 17 August 2002. The Legal Assessor advised the Committee that it should not give any weight to the article, but merely note it and the article, and the potential prejudice to the respondent doctors by its publication at this time.

The Committee noted that two expert reports had been obtained in this case, one on the statistical aspects and one on the paper itself and HMPAO imaging. The Committee noted its obligation to judge each set of allegations against the standards in place in 1989, and to deal with each doctor individually.

In relation to Dr Revell, the Committee noted that the allegations against him were in connection with his response to a complaint about the paper. It was alleged that the contents of his letter of February 1994 were false and misleading because he gave the impression that the study had Ethics Committee approval and that the Royal Navy was covered by the appropriate ARSAC certificates when this has since been discovered to be untrue.

The Committee noted that Dr Revell’s letter had been based on information supplied by Dr Sykes, whom he had consulted as he was involved in the research and because he was an expert in the field of diving medicine. The Committee considered that given the seriousness of the allegations made, he should have done more to satisfy himself that the allegations were without foundation. In particular Dr Revell could have asked for an independent opinion on the allegations made. However, given that his response had been written in good faith, based on technical information from a senior clinician in the relevant field it did not consider that a question of serious professional misconduct arose. In reaching this decision it also took account of the standards of the time in relation to what was expected


The Lancet, 32 Jamestown Road, London NW1 7BY, UK

*Sabine Kleinert, Richard Horton sabine.kleinert@lancet.com

The Lancet, 32 Jamestown Road, London NW1 7BY, UK

Correspondence

E-mail submissions to correspondence@lancet.com
of a doctor in investigating complaints. A much more rigorous approach would be expected today.

Nevertheless, the Committee has asked me to convey its advice to Dr Revell that he has a responsibility to investigate complaints thoroughly and to verify the accuracy of information to ensure that it is correct before it is passed on to a third party.

In relation to the allegations against Dr Sykes, Dr Macleod, Dr Torok and Dr Pearson, the Committee noted that the allegations against each doctor were identical. It was alleged that a current Administration of Radioactive Substances Advisory Committee (ARSAC) certificate was required for the HMPAO brain scanning but no such certificate was issued to the Royal Navy at that time, and that patients who took part in the study were told it was research although there was no Ethics Committee approval for it; alternatively, if it was not research, the patients were misled.

The Committee noted the documentation on the subject of the ARSAC certificate, which indicated that although a certificate had been required in order to administer therapeutic radioactive products, the method of granting such certification had changed in the period 1986–89. The other main issue was whether the paper was true research or whether it was a retrospective study reporting on the outcome of clinical investigations. Leading on from this was the matter of informed consent, whether the patients thought they were taking part in research or routine investigations.

In relation to Dr Sykes, the Committee noted that in the letter of explanation submitted on his behalf, it was admitted that he now understood that the ARSAC certificate was required for HMPAO scanning, but that was not his understanding in 1989 when he thought HMPAO was licensed for brain scanning and its use in patients following recompression treatment was entirely appropriate. The Committee considered that the timings of changes in the certification requirements were not clear, and that there would be no real prospect of establishing culpability in the failure to ensure an ARSAC certificate was in place.

The Committee noted that Dr Sykes asserted that the Lancet paper was a retrospective review of case studies. Therefore it was submitted that Ethics Committee approval was not required, and only informed consent in the form usual in 1988 was required from the patients. The Committee considered there may have been lack of clarity in the language used to patients about the purpose of treatment and investigation, and whilst this was regrettable, given the standards of the time this did not raise a question of serious professional misconduct and that therefore the matter should not be referred for public inquiry by the Professional Conduct Committee.

In relation to Dr Macleod, the Committee noted that he was required to have an ARSAC certificate in respect of the HMPAO brain scanning. It also noted the letter of explanation on his behalf which stated that it was his understanding that if a new agent became licensed during the period of a five year certificate, he could start using it and only need have it added to his ARSAC certification at the next issue of the five year certificate. It noted the confusion that surrounded the issue of ARSAC certification but considered that as a consultant in nuclear medicine he had a responsibility to ensure he was properly authorised for the work he undertook. However, the Committee did not consider, in the circumstances, that Dr Macleod’s conduct in this respect raised a question of serious professional misconduct.

The patients in the study were referred to Dr Macleod for diagnosis by perfusion scanning. The Committee accepted that he would not have known whether Ethics Committee approval was necessary or had been granted. The Committee considered there may have been lack of clarity in the language used to patients about the purpose of treatment, and whilst this was regrettable, given the standards of the time this did not raise a question of serious professional misconduct.

Nevertheless, the Committee has decided to issue a warning and advice in relation to this matter. It has asked me to stress to both Dr Sykes and Dr Macleod the importance of ensuring that the appropriate documentation is in place to allow procedures to be carried out on their patients. In relation to the consent issues, the Committee directed that Dr Sykes and Dr Macleod should be advised of the need to be very clear about the purposes of treatment when seeking consent, and to inform patients if any study arising from treatment is contemplated.

As regards Dr Torok, the Committee noted that he was a civilian medical practitioner who worked at the Institute of Naval Medicine at the time the Lancet paper was published. His involvement was to carry out the clinical assessment of divers. In the letter of explanation submitted on his behalf, it was stated that he would have had no reason to question the technical qualifications of those undertaking the scans, or the existence of appropriate regulatory consents. The Committee accepted this point. The Committee also noted that while it is now usual for papers with multiple authors to show how responsibility was split and the level of involvement in research, this was not the case in 1989 when normal practice was to name all authors equally.

In relation to Dr Pearson the Committee noted that he had been an occupational physician in 1989, and had no clinical responsibilities at the Royal Naval Hospital. It noted that in his letter of explanation he stated that he was only involved with one of the patients detailed in the Lancet paper. It also noted his denial that he had any responsibility to check that the appropriate ARSAC certification had been issued. The Committee accepted this point.
The Committee determined that both Dr Torok and Dr Pearson had minor roles in the preparation of the paper, and that their actions could not raise a question of serious professional misconduct. It determined that no further action should be taken against either Dr Torok or Dr Pearson in relation to this case.

Thank you for bringing this matter to our attention.

Yours sincerely

Assistant Registrar

Conduct Case Presentation Section, General Medical Council, 178 Great Portland Street, London W1W 5JE, UK

For contact: Andrew Wood (awood@gmc-uk.org)

The UK government recognises the British Sub-aqua Club as the governing body for amateur scuba diving. I am a member of the club’s medical committee. I was the committee’s chairman when G H Adkisson and colleagues reported research on British amateur divers.1 I was immediately concerned that there was no statement about ethics committee approval for injection of the radioactive isotope and about accuracy of data reporting.2

The authors claimed that cerebral perfusion deficits were always detected by technetium-99 hexamethylpropyleneamine oxime (HMPAO) scans when individuals had neurological dysbaric (decompression) illness, but those with non-neurological dysbaric illness had a normal scan. T O Nunan and I challenged this claim.3 We were also concerned that, by then contemporary standards, the research required ethics committee approval. Adkisson responded that ethics approval was not required because the scans were for patient management.3 I believe that such use required published evidence of utility in that disease. I believe that the authors should explain how the scans assisted patient management. Adkisson stated that ethics committee approval was not required because this was a retrospective review of cases. I would like him to explain how “a retrospective review” can be squared with the statement in the paper that “Patients were selected for inclusion if they met the following criteria . . . ”. I would like Sykes and Revell to explain how Revell came to believe that ethics committee approval had been granted when Sykes believed that it was not required.

In 1992, The Lancet rejected my attempts to get this matter reopened. Imogen Evans wrote “In response to your first point about the ethical considerations of Adkisson, we are talking about something that happened in 1989 and therefore we do not think it is useful for The Lancet to return to that specific aspect.” In 1995, The Lancet announced a willingness to withdraw aegis from an article if the journal would not have published it had the full facts surrounding the research been known at the time of acceptance.4 I approached Richard Horton about this paper. At his request, I produced an article with supporting documents outlining my case. In 1997, Horton wrote to the Royal Navy requesting that the authors respond. Instead, in July, 1997, Surgeon Rear Admiral Paine, Revell’s successor, asked the GMC to investigate whether I had disparaged the authors. Paine informed the GMC that my “accusations were taken very seriously indeed and the matter was thoroughly investigated [presumably by Revell]; no evidence was found of anything that could reasonably be construed as fraud or unethical conduct”.

Horton was surprised by the events. The solicitor appointed by the Medical Defence Union wrote to me “I have now had the opportunity of reading the papers you have forwarded to the MDU and felt great indignation on your behalf when doing so. On the one hand the General Medical Council are advising members that they have a duty to report unethical behaviour but...”
on the other hand seem to have some reluctance to take on government establishment.” In March, 1998, the GMC exonerated me. The GMC also said “The members consider that Dr Horton’s proposed means of establishing whether The Lancet’s aegis should be withdrawn from the research in question is the proper and most effective way of investigating your concerns, and should be pursued.”

I asked the GMC to consider the behaviour of the authors of the paper and the senior doctors in the Royal Navy. The GMC’s eventual decisions are shown in the accompanying letter from the GMC dated Sept 24, 2002. The speed with which the GMC investigated me for disfigurement contrasts with the slow pace used to investigate the authors of the paper. The Royal Navy was uncooperative, but in October, 2000, the GMC exercised new powers to compel the Royal Navy to provide relevant documents. I understand that the relevant clinical records were incomplete and could not be interpreted. I would like the Royal Navy to explain why they did not take more trouble to preserve the data when queries were repeatedly raised. The GMC responded to me on Sept 24, 2002, more than 4 years after my complaint. I believe that the unresolved issues relating to this research require explanation.

Peter Wilmhurst
peter.wilmshurst@rsh.nhs.uk
Royal Shrewsbury Hospital, Shrewsbury SY3 8XQ, UK


I am one of the six authors of the paper which appeared in The Lancet in July 1989 and which has been the subject of considerable criticism by Peter Wilmshurst who has, I understand, called for the aegis of the paper to be removed and which culminated in him complaining about the paper to the GMC in 2000. Those criticisms extended beyond concerns regarding the application of ARSAC certification and the ethical nature of the conduct of the investigation and included the far more serious allegation of research fraud.

As part of the GMC’s investigations, which were conducted with the full co-operation of the Naval authorities, the paper was submitted for independent statistical and peer review. The eminent statistician concluded that: “there is no sound statistical basis for a claim of scientific fraud with regard to the 1989 Lancet results”. These allegations by Wilmshurst were therefore rejected by the GMC at the screening stage and were not submitted to the Preliminary Proceedings Committee as part of the charges made against the authors. Furthermore, the independent Consultant Neurologist who provided an expert opinion to the GMC on the papers concluded: “Given the previous concerns at the time in respect of subclinical damage in neurologically damaged divers and the need to try and quantify such insults, in my view it was entirely reasonable to use HMPAO imaging as part of the assessment of patients with dysbaric illness . . . “.

The shortcomings of the paper, which were limited to the failure to include a control group, were no greater than that recognised by the authors at the time. A prospective, controlled study was established the following year and the results published. This immediately led to the responsible abandonment of the use of HMPAO scans for diagnostic purposes in this context by the Royal Navy. I would submit that it is far from a unique situation where subsequent research does not support an initial presumed correlation.

It was also alleged in the complaint to the GMC that ethics committee approval was required for the work done using HMPAO screening by the Undersea Medicine Division of the Institute of Naval Medicine. The original Lancet article arose from a retrospective review of case studies and therefore no ethics committee approval was required—a view defended by the first author at the time. The Preliminary Proceedings Committee of the GMC accepted that position and determined that the study had in fact been conducted in accord with the standards of the day and therefore only informed consent was required, which I believe was obtained properly. Readers should also be aware that the work of the Undersea Medical Division of the Institute of Naval Medicine was closely monitored by the Medical Research Council’s Underwater Physiology Subcommittee and that the early findings of HMPAO had also been widely discussed (and recorded) with a wide range of authorities from academic, scientific, regulatory, and other backgrounds and had given rise to no concerns either ethically or scientifically.

I acknowledge with the benefit of hindsight that ARSAC certification was required for HMPAO screening, albeit that the responsibility for ensuring its presence fell outside my department,
but the GMC recognised that in 1989 the changes to the certification process were not clear and as a consequence determined that this matter did not need to be investigated further. It is quite possible that other units using HMPAO may also have been affected by the changes in the certification process.

The GMC, as regulatory body, has dismissed the claims of scientific fraud and ethical impropriety against the authors registered with it. It has advised me of best practice and the standards it requires from any future involvement in similar activities, which I fully accept. The case was not referred to the Professional Conduct Committee of the GMC and there was no hearing. There is therefore no case to answer. It is now hoped that the matter will be consigned to a footnote in underwater medicine and that it should be properly allowed to rest 16 years after the paper was published.

J J W Sykes
jimsyke@dscd.mod.uk
Defence Post Graduate Medical Deenery, ICT Centre, Birmingham Research Park, Vincent Drive, Birmingham B15 25Q, UK


Your decision to publish the letter from the GMC to Peter Wilmshurst is, at best, perverse. I do not consider that any of your current readers have any interest whatsoever in issues arising from an article published by The Lancet some 16 years ago. Suffice to say, it was submitted in good faith and has been more than adequately defended. It is also pertinent to add that the GMC, after consulting expert opinion, did not support any allegations about the scientific merits of the article itself.

Ramsay Pearson
ramsdale6@aol.com

Combination anti-hypertensives in WHO essential medicine list

In his Comment (June 25, p 2160), Valery Feigin rightly emphasises the urgent need to implement effective preventing strategies for stroke in developing countries, with particular emphasis on blood-pressure control. In the related original study done over 2 years in Chile by Pablo Lavados and colleagues (p 2206), the mean age of the 380 patients who had a first-ever stroke was 68·5 years (SD 14·6) for women and 61·2 years (16·6) for men, the proportion with intracerebral haemorrhage was 23%, and the proportion with prestroke hypertension 60%.

No population-based stroke study has been done in Nepal. However, in our study of 40 new stroke patients consecutively admitted over 4 months from October, 1995, to January, 1996, to the general medicine unit of the central government hospital in Nepal, 25 (62%) were younger than 60 years and 17 (41%) had intracerebral haemorrhage. Of those with intracerebral haemorrhage, 10 (59%) had hypertension with irregular or no treatment.1 Similarly, among 100 consecutive new stroke patients admitted between November, 1999, and February, 2000, to a Nepalese teaching hospital, 34 were aged between 40 and 59 years and 57 had haemorrhagic stroke.1 Because haemorrhagic strokes are more severe than ischaemic strokes and require hospital admission, the proportion of intracerebral haemorrhage is higher in these hospital-based reports than in population reports. However, the high prevalence of uncontrolled hypertension could contribute.2,3,4

Proper control of hypertension is an important possible means of preventing stroke. But almost half of patients with hypertension require more than one medicine to properly control blood pressure. Many people do not like taking many tablets. Thus, hypertension is an important condition for which combination medicines need to be made available in the community. Such combination products are mostly not available in Nepal.

3 years ago on June 15, 2002, I wrote a letter to the Department of Drug Administration about the need to make available combination antihypertensive medicines, and subsequently discussed the issue many times. The editorial team of the Journal of the Nepal Medical Association, including myself, also wrote an editorial entitled “Need for combination anti-hypertensive medicines in the market” in October, 2002. But the situation remains the same.

The major hindrance frequently expressed by the officials concerned is non-inclusion of combination antihypertensive medicines in the WHO Model List of Essential Medicine. Combination medicines are rightly listed separately for diseases such as tuberculosis, but not for hypertension. Different combinations of two to three antihypertensive drugs are available in many countries, including in the industrialised ones.2 If such combinations were only included in the Model List of Essential Medicines, it would be easier to make them available in developing countries like Nepal and thus prevent many premature disabilities and deaths. Without a simultaneous effort to make effective preventive strategies available, there is little point in doing scientific studies, whether hospital-based or population-based.

I declare that I have no conflict of interest.

Madhur Bhattarai
mdb@ntc.net.np
Bir Hospital, Post Box 2245, Kathmandu, Nepal

Adjuvant trastuzumab for HER2-positive breast cancer

In your June 18 Editorial (p 2064), 1 you bemoan a lack of accurate information to help informed patients make decisions affecting their treatment. In the example used, a woman with early breast cancer is reported to be selling her home to pay for adjuvant trastuzumab (a monoclonal antibody specific for the growth-factor receptor HER2). She is presumably basing this decision on the striking results of trials presented in a special session at the recent American Society of Clinical Oncology annual meeting.2 In your Editorial, you imply that this patient has been misinformed by inaccurate reporting of these trial results.

In fact, despite only short follow-up to date, these European and North American trials showed hazard ratios for disease-free survival of 0.54 (95% CI 0.43–0.67, p<0.0001) and 0.48 (0.39–0.60, p<0.0001), respectively, in favour of trastuzumab. Disease-free survival is widely accepted as a major end-point and a surrogate for overall survival when assessing adjuvant trials in early breast cancer, because most metastatic disease is not curable.4 Furthermore, the North American data already show a significant overall survival improvement with a relative risk reduction of 33% (p=0.015). It is unusual to see any such effect on survival so early in an adjuvant breast cancer trial.

Cardiac toxicity is the only significant adverse effect of trastuzumab, but in these trials was almost always reversible, and trastuzumab-related cardiac deaths were very rare (none has so far been reported in the experimental group of the European trial). As a result of the remarkable efficacy shown in these recent trials, trastuzumab is rapidly becoming the standard of care after surgery for high-risk HER2-positive breast cancer. The data have been widely disseminated in forms readily accessible to motivated patients,5 who will understandably ask their doctors about them. The opinions of our patients, informed or otherwise, demand our respect, especially when they are correct.

PE has received grants from Roche for the conduct of clinical trials and has been a member of advisory boards for the company.

*James Spicer, Mark Harries, Paul Ellis james.spicer@gstt.nhs.uk
Department of Medical Oncology, Guy’s Hospital, London SE1 9RT, UK

2 Remond E, Perez EA, Bryant J, et al. Doxorubicin and cyclophosphamide followed by paclitaxel with or without trastuzumab as adjuvant therapy for patients with HER-2 positive operable breast cancer: combined analysis of NSABP B-31/NCCTG-N9831. http://www.asco.org/ac/1,1003,_12-000251-00-00-0034-00-00-005816-00-21-001,00.asp (accessed July 20, 2005).
3 Piccart-Gebhart M. First results of the HERA trial, a randomised three-arm multicentre comparison of Herceptin in women with HER-2 positive primary breast cancer who have completed adjuvant chemotherapy. http://www.asco.org/ac/1,1003,—12-000251-00-18-0034-00-19-005816-00-21-001,00.asp (accessed July 20, 2005).

Controlling the three “P”s in Africa

I agree with the tenor of your June 18 Editorial (p 2064). 1 Africa’s Gold Coast gained independence from Britain in 1957. In less than 25 years, the fledgling Ghana had five coups d’état plus one revolution, before which soldiers invaded Korle Bu Hospital compound and beat us up. I had to flee my house to sleep on the ward for several days in July, 1977. Doctors left Ghana in droves, inflation rocketed, poverty was rife, and the health of the nation was non-existent.

Developed countries successfully control what I call the three “P”s: Politics, Poverty, and Population.2 In developing countries, these get so out of hand that national health suffers. The more out of control the three “P”s are, the less healthy the nation is. Lasting help is required to deflate our ballooning three “P”s, and for this we say to the G8, International Monetary Fund, and World Bank: (1) “Thank you” for pressing democratic government on us, and for cancelling our huge debts. (2) Please stop dictating the price of our raw materials. (3) Remove tariffs on our produce and allow us to sell to Europe. (4) You subsidise your farmers, and dump rice and sugar on our markets. Why? (5) You allow your banks to accept billions from our corrupt dictators, and then keep all the money when they die. Why? (6) Why do you always advise devaluation of our currency?

What we really need is a paradigm shift in our approach to all Africa’s health problems,3 acknowledging also that the stellar qualities of our doctors make them a valuable export, especially to the USA.

I declare that I am on Ghana Government Pension working abroad and supporting Ghanaian projects in clinical epidemiology and African anthropogenetics.

Felix I D Konotey-Ahulu
felix@konotey-ahulu.com
Cromwell Hospital, London SW5 0TU, UK

2 Konotey-Ahulu FID. The three “P”s in health care delivery in developing countries. The Healing Hand (Journal of the Edinburgh Medical Missionary Society) 1985; Summer/Autumn: 8–15.

Department of Error

Bhandari N, Bahia R, Maximum S, et al. Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illnesses and growth: a cluster randomised controlled trial. Lancet 2003; 361: 1418–23. In table 2 of this article (April 26, 2003) the number for non-breastfed babies in the control group at 3 months should be 12 (3%).
Chemotherapy versus allogeneic transplantation for very-high-risk childhood acute lymphoblastic leukaemia in first complete remission: comparison by genetic randomisation in an international prospective study

Adriana Balduzzi, Maria Grazia Valsecchi, Cornelio Uderzo, Paola De Lorenzo, Thomas Klingebiel, Christina Peters, Jan Stary, Maria S Felice, Edina Magyarosy, Valentino Conter, Alfred Reiter, Chiara Messina, Helmut Gadner, Martin Schrappe

Summary

Background The dismal prognosis of very-high-risk childhood acute lymphoblastic leukaemia could be improved by allogeneic haemopoietic cell transplantation. We compared this strategy with intensified chemotherapy protocols, with the aim to improve the outcome of children with very-high-risk acute lymphoblastic leukaemia in first complete remission.

Methods A cooperative prospective study was set up in seven countries. Very-high-risk acute lymphoblastic leukaemia in first complete remission was defined by the presence of at least one of the following criteria: (1) failure to achieve complete remission after the first four-drug induction phase; (2) t(9;22) or t(4;11) clonal abnormalities; and (3) poor response to prednisone associated with T immunophenotype, white-blood-cell count of $100 \times 10^9$/L or greater, or both. Children were allocated treatment by genetic chance, according to the availability of a compatible related donor, and assigned chemotherapy or haemopoietic-cell transplantation. The primary outcome was disease-free survival and analysis was by intention to treat.

Findings Between April, 1995, and December, 2000, 357 children entered the study, of whom 280 were assigned chemotherapy and 77 related-donor haemopoietic-cell transplantation. 5-year disease-free survival was 40·6% (SE 3·1) in children allocated chemotherapy and 56·7% (5·7) in those assigned transplantation (hazard ratio 0·67 [95% CI 0·49–1·09]; p=0·12).

Interpretation Children with very-high-risk acute lymphoblastic leukaemia benefit from related-donor haemopoietic-cell transplantation compared with chemotherapy. The gap between the two strategies increases as the risk profile of the patient worsens.

Introduction

High-risk childhood acute lymphoblastic leukaemia, which is defined by biological characteristics at onset—eg, high white-blood-cell count, T immunophenotype, or t(9;22) and t(4;11) clonal abnormalities—or by resistance to treatment (poor prednisone response, induction failure), has a poor prognosis.1–3 During the past decade, differences in definitions of high-risk criteria and in frontline chemotherapy protocols have produced various results, with a cure rate for the disorder of 30–60%.4–9 Findings of retrospective studies have suggested that allogeneic haemopoietic-cell transplantation from HLA-identical siblings improves the dismal prognosis of high-risk childhood acute lymphoblastic leukaemia compared with further intensified chemotherapy protocols.10–17 As far as we are aware, no randomised trials have compared chemotherapy with transplantation in patients with acute lymphoblastic leukaemia in first complete remission; therefore, most results reported so far, being retrospective, are affected by two major biases. First, selection of patients to be transplanted is not only dictated by availability of a donor but also affected by the risk profile of the patient, as perceived by clinicians during the course of the disease. Second, the waiting time to transplantation means the transplanted group does not include individuals who do not survive in first complete remission long enough to undergo the procedure.18

In 1995, a cooperative prospective study was set up in collaboration with the International Berlin-Frankfurt-Münster Study Group and the Pediatric Working Party of the European Blood and Marrow Transplantation Group to compare allogeneic haemopoietic-cell transplantation from a compatible related donor with chemotherapy, with the aim to improve the outcome of children with acute lymphoblastic leukaemia in first complete remission.19

Participants and methods

Participants Between April, 1995, and December, 2000, all consecutive patients younger than 18 years with newly diagnosed acute lymphoblastic leukaemia (based on morphologic, cytochemical, and immunophenotypic criteria, excluding mature B-cell disease) were assessed for eligibility by five study groups from seven different countries: Italy, Germany, Switzerland, Austria, Czech Republic, Hungary, and Argentina.
Hungary, and Argentina. Study groups and their members are listed in the webappendix.

We included children with very-high-risk disease, which was defined by the presence of at least one of the following criteria, listed in order of importance: (1) failure to achieve complete remission after four-drug induction treatment, henceforth referred to as induction failure; (2) t(9;22) or t(4;11) clonal abnormality, defined by conventional cytogenetic or molecular analysis; and (3) poor response to prednisone, defined by persistence of at least 1·0 × 10⁹ blast cells per L in peripheral blood after 7 days of prednisone monotherapy and one injection of intrathecal methotrexate, associated with T immunophenotype, white-blood-cell count of 100 × 10⁹/L or greater, or both.

To enter the study, every child had to meet the definition of very-high-risk disease, had to achieve first complete remission, and his or her family had to be HLA-typed for a compatible related donor search. Those who had induction failure were included in the study as long as they achieved complete remission by the end of consolidation treatment. Children who failed to achieve complete remission after induction failure and who did not ultimately achieve complete remission by the end of consolidation treatment were not included in the study.

Complete remission was defined by fewer than 5% blast cells in the bone-marrow aspirate, with normal cellularity and trilineage haemopoiesis, associated with no more than five cells per μL and no blasts in the cerebrospinal fluid. A compatible related donor was defined as either a genotypically identical sibling or phenotypically compatible family member with no more than one antigenic or allelic mismatch at either class I A, B, or C loci (typed by low-resolution molecular techniques) or the class II DRB1 HLA locus (typed by high-resolution molecular techniques).

The institutional review board of every participating centre approved the protocol before enrolment of children into the study. Informed written consent approved by local ethics committees was obtained from parents or guardians for every child before participation; parents or guardians had the right to refuse to enter the study.

**Procedures**

Eligible children underwent similar frontline chemotherapy protocols based on the Berlin-Frankfurt-Münster approach for paediatric acute lymphoblastic leukaemia. Induction consisted of the same drugs at the same doses and timing, except for a 1-week delay at the beginning of asparaginase administration in the Italian schema. Subsequent treatment included six chemotherapy BFM blocks followed by one reinduction course for all but the Italian group, which included three chemotherapy blocks followed by two reinduction courses.

Children were assigned to one of two treatment arms depending on whether an HLA-compatible related donor was available. Those without a suitable donor continued the chemotherapy protocol to which they had been started on at diagnosis, whereas for the other children, allogeneic haemopoietic-cell transplantation was planned within a recommended time frame of at least 2 months after achievement of remission and within 5 months after diagnosis. For children allocated haemopoietic-cell transplantation, treatment for graft-versus-host disease consisted of ciclosporin 3 mg/kg per day, and the conditioning regimen entailed fractionated total-body irradiation (2 Gy twice a day on days –6, –5, and –4) and etoposide (60 mg/kg intravenously on day –3) for children older than 2 years, and busulfan (5 mg/kg per day by mouth on days –8, –7, –6, and –5), etoposide (40 mg/kg intravenously on day –4), and cyclophosphamide (60 mg/kg per day intravenously on days –3 and –2) for those younger than 2 years. Supportive therapy, including total parenteral nutrition, treatment for infectious diseases, and management of graft-versus-host disease, was done according to every institution’s policy.

Transplantation from alternative donors—either unrelated or more than one antigen-mismatched related—was not considered as a treatment option, although such procedures were sporadically undertaken as a choice of the treating doctor.

**Statistical analysis**

Very-high-risk acute lymphoblastic leukaemia was expected to occur in 8% of children in the participating study groups, and expected overall recruitment was estimated to be around 90 per year. 4-year recruitment was thus forecast to yield an approximate total sample size of 300, with which one could expect 80% power to detect a difference in the 4-year disease-free survival of at least 17% or 18%, given that 30% or 25% of children, respectively, would have a suitable related donor. The power calculation was done assuming a 4-year disease-free survival of 35% in children with very high-risk disease conventionally treated with chemotherapy, and a one-sided test with 5% type I error.

The primary endpoint was disease-free survival, and overall survival was the secondary endpoint. In both arms, disease-free survival was defined as the time from achievement of first complete remission to last follow-up or any of the following events, whichever occurred first: relapse (at any site), death while in complete remission (of any cause), or second malignant disease. Death from any cause was the sole event in determining overall survival. The main analyses were done according to the intention-to-treat principle, with treatment allocation based only on donor availability and thus on genetic randomisation.

Disease-free survival and survival curves were produced according to the Kaplan-Meier method and their SEs were estimated with the Greenwood formula. The Cox model was used in two ways: (1) to assess treatment effect, after stratifying by study group according to the principles of...
prospective meta-analysis;10 and (2) to adjust treatment effect according to prognostic features, namely those defining the very-high-risk criteria.22 Furthermore, the time-dependence of the treatment effect was measured by including a term for the interaction between time and treatment in the model. The likelihood ratio test assessed heterogeneity of treatment effect among study groups. A one-sided Wald test was used for the intention-to-treat comparison of disease-free survival only; all other tests were two-sided. The probabilities of relapse and of death in first complete remission were calculated with the cumulative incidence estimator, thus allowing for competing risks, and they were compared according to Gray.23 Additional analyses according to treatment given were also undertaken as secondary analyses; adjustment for waiting time to transplantation was done both in the Cox model (with a time-dependent treatment indicator) and in the estimation of the survival curves.22

Every year, data were pooled by a trial data centre, checked for consistency, and processed for reporting, interim, and final analyses. Follow-up was updated in December, 2002, for every study group. Analyses were done with SAS version 8.1 (SAS Institute, Cary, NC, USA) and R statistical software.

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 the report for publication.

Results
Between April, 1995, and December, 2000, 389 consecutive children (246 boys, age 0–17 years) were assessed for eligibility; overall, they accounted for 8% of the registered acute lymphoblastic leukaemia population in the five study groups. Figure 1 shows the flow of participation in the study. 18 children were not assessable or did not fulfil the criteria for very-high-risk disease, and HLA-type was not known for 14, so a total of 357 entered the study (156 from Germany and Switzerland; 23 Austria; 122 Italy; 22 Czech Republic; 16 Hungary; and 18 Argentina). 97 children had no siblings. Of the 260 who had one or more, 76 had a compatible sibling; for one child, another compatible family member was found. Overall, 280 children were allocated chemotherapy and 77 were assigned related-donor haemopoietic-cell transplantation.

Distribution of sex, age, white-blood-cell count at diagnosis, and T immunophenotype were well balanced between groups (table 1). Among the 56% (201) for whom cytogenetic data were known, about 18% (35) presented with t(4;11) in both treatment groups, whereas 48% (75) allocated chemotherapy and 19% (8) assigned transplantation presented with t(9;22). Of patients in whom induction failure status was known, 22% (58) of children assigned chemotherapy and 34% (25) of those allocated haemopoietic-cell transplantation had induction failure.

With respect to the criteria for very-high-risk disease, 23% (n=83) of children had induction failure, 21% (74) and 9% (32) had t(9;22) and t(4;11) clonal abnormalities, respectively, and 47% (168) had a poor response to prednisone, associated with T-immunophenotype, high white-blood-cell count, or both (table 1). Some presented with a combination of these very-high-risk features, moreover, not all features were known for every child, so the overall prevalence of biological features could differ from the eligibility criteria. For instance, children who had induction failure, a poor response to prednisone, and t(9;22) would be counted only in the worst eligibility criterion, induction failure. 5-year disease-free survival of children sharing these very-high-risk features was 44.2% (SE 2.7), and overall survival was 51.4% (SE 2.8), with a median follow-up of 5 years (IQR 4–6).

Figure 1 shows the outcome according to donor availability (intention-to-treat analysis) in terms of numbers and types of events; in both treatment groups, relapse was the most common cause of treatment

<table>
<thead>
<tr>
<th>Eligibility criteria*</th>
<th>Induction failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>t(9;22)</td>
<td>58 (21%)</td>
</tr>
<tr>
<td>Other abnormalities</td>
<td>19 (7%)</td>
</tr>
<tr>
<td>Absent</td>
<td>25 (9%)</td>
</tr>
<tr>
<td>T immunophenotype</td>
<td>12 (4%)</td>
</tr>
<tr>
<td>WBC</td>
<td>34 (12%)</td>
</tr>
<tr>
<td>No</td>
<td>12 (4%)</td>
</tr>
<tr>
<td>Good</td>
<td>21 (7%)</td>
</tr>
<tr>
<td>Eligibility criteria*</td>
<td></td>
</tr>
<tr>
<td>T immunophenotype</td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Mediately</td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Eligibility criteria*</td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Mediately</td>
<td></td>
</tr>
</tbody>
</table>

Data are numbers of children (%), unless otherwise indicated. PPR=poor response to prednisone. WBC=white-blood-cell count ≥109/L. T=T-immunophenotype. *Children are only counted in the first listed criterion that made them eligible to the study—ie, those with several very-high-risk features are counted in the first appropriate category, irrespective of other coexisting features.

Table 1: Children’s characteristics and eligibility criteria

See http://www.r-project.org

www.thelancet.com Vol 366 August 20, 2005 637
failure. Of 280 children assigned chemotherapy, 28 (10%) died at a median of 8 months (IQR 6–11) after first complete remission, and 134 (48%) relapsed at a median of 12 months (7–24). Relapse sites were: bone marrow in 106; bone marrow combined with other sites in 17; CNS in seven; CNS, testis, and eye in one; eye in one; and unknown in two. Of 77 children assigned haemopoietic-cell transplantation, seven (9%) died a median of 7 months (6–7) after first complete remission, and 26 (34%) relapsed at a median of 10 months (7–14). Sites of relapse were: bone marrow in 21; CNS in two; bone marrow and testis in one; pericardium in one; and unknown in one.

Availability of a compatible related donor was associated with significantly better outcome (figure 2). 5-year disease-free survival in the chemotherapy group was 40·6% (SE 3·1) compared with 56·7% (5·7) in the haemopoietic-cell transplantation group (hazard ratio 0·67 [95% CI 0·46–0·99], p=0·02). When this analysis was adjusted by eligibility criteria (to account for imbalances in children’s characteristics in the two treatment groups) results were similar (0·65 [0·44–0·96], p=0·01). The advantage of related-donor haemopoietic-cell transplantation over chemotherapy became more apparent with every successive year of follow-up, suggesting greater protection against late relapses in children who survived the early toxic effects of treatment (table 2). No significant heterogeneity of treatment effect was detected among study groups (p=0·12).

Cumulative incidences of relapse differed between the chemotherapy and haemopoietic-cell transplantation groups (49·3% [SE 3·1] vs 34·2% [5·4], p=0·06), whereas those of death in remission were similar (10·1% [1·8] vs 9·1% [3·3], p=0·85; figure 3). The 5-year survival estimate in children assigned chemotherapy was 50·1% (SE 3·1), which was lower than that in those allocated transplantation (56·4% [5·9], p=0·12 [stratified by study group]; figure 4).

The advantage of allocated related-donor transplantation over chemotherapy for disease-free survival was consistent across subsets of children with very-high-risk disease. The
A induction failure subset reported a 5-year disease-free survival of 26·5% (SE 5·9) in the chemotherapy group versus 56·0% (9·9) in the haemopoietic-cell transplantation arm (p=0·03). Children who were enrolled only because of a poor response to prednisone, associated with T immunophenotype, white-blood-cell count greater than $10^9/L$, or both, reported a 5-year disease-free survival of 54·3% (SE 4·5) in the chemotherapy group versus 62·4% (8·0) in the transplantation arm (p=0·32). Notably in this subset, children who had both T immunophenotype and hyperleukocytosis had a 5-year disease-free survival of 48·0% (SE 5·7) if allocated chemotherapy versus 55·9% (11·7) if assigned haemopoietic-cell transplantation.

Each clonal translocation was not assessed as a subset, in view of the few children involved. Rather, subsets of all children presenting with a relevant translocation, irrespective of other features, are described here. For those presenting with t(9;22), 23 of 75 assigned chemotherapy were alive in continuous complete remission (disease-free survival 25·6% [SE 6·0]) compared with four of eight who were allocated haemopoietic-cell transplantation. Of the 75 children assigned chemotherapy, 11 also had a poor response to prednisone and none was alive in complete remission. For 27 children presenting with t(4;11), ten were alive in continuous complete remission (disease-free survival 36·4% [SE 9·4]) compared with three of eight children assigned transplantation. For all children presenting with T immunophenotype, irrespective of other very high-risk features, 5-year disease-free survival was 47·9% (SE 4·5) in children assigned chemotherapy compared with 55·9% (11·7) if assigned haemopoietic-cell transplantation.

Overall, 65 (18%) of 357 children deviated from their assigned treatment; the lower part of figure 1 shows treatment actually received and subsequent outcome. Causes for deviations were known in only some cases, but were mostly because of the choice of the treating doctor, rarely parental decision. Of the 280 children allocated chemotherapy, 43 (15%)—who were enrolled for induction failure (n=20), clonal translocation (11), or poor response to prednisone with T immunophenotype (12)—deviated from the protocol, and at a median of 6 months (IQR 3–8) after first complete remission they underwent transplantation from an alternative unrelated donor who was either compatible (29) or partly incompatible (12), or from an unspecified donor (2). Of the 77 children allocated related-donor transplantation, 55 actually underwent transplantation according to protocol, at a median of 4 months (3–5) after first complete remission, whereas the remaining 22 (29%)—enrolled for induction failure (n=8), clonal translocation (3), or poor response to prednisone with T immunophenotype (6) or hyperleukocytosis (5)—deviated from their treatment allocation and continued on chemotherapy.

<table>
<thead>
<tr>
<th>Number at risk</th>
<th>280</th>
<th>190</th>
<th>146</th>
<th>112</th>
<th>85</th>
<th>54</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years from first complete remission</td>
<td>100</td>
<td>80</td>
<td>60</td>
<td>40</td>
<td>20</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Relapsed HCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>280</td>
<td>190</td>
<td>146</td>
<td>112</td>
<td>85</td>
<td>54</td>
<td>31</td>
</tr>
<tr>
<td>Related donor</td>
<td>77</td>
<td>65</td>
<td>53</td>
<td>47</td>
<td>34</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>Deaths, p=0·85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relapses, p=0·06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: Estimates of cumulative incidence of relapse and death, by treatment assigned
HCT=haemopoietic-cell transplantation.

<table>
<thead>
<tr>
<th>Number at risk</th>
<th>280</th>
<th>218</th>
<th>169</th>
<th>135</th>
<th>103</th>
<th>70</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years from first complete remission</td>
<td>100</td>
<td>80</td>
<td>60</td>
<td>40</td>
<td>20</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Related-donor HCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>280</td>
<td>218</td>
<td>169</td>
<td>135</td>
<td>103</td>
<td>70</td>
<td>40</td>
</tr>
<tr>
<td>Related donor</td>
<td>77</td>
<td>65</td>
<td>53</td>
<td>47</td>
<td>34</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>5-year survival (%), p=0·12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Estimates of survival, by treatment assigned
HCT=haemopoietic-cell transplantation. One death after relapse occurred after 6 years.
Of the 259 children who received chemotherapy, 109 remained in complete remission, 134 relapsed at a median of 12 months (IQR 7–24), and 16 (6%) died at a median of 6 months (2–8). Of the 55 who actually underwent haemopoietic-cell transplantation from compatible related donors, 35 remained in complete remission, 15 relapsed at a median of 7 months after transplantation (5–12), and five (9%) died at a median of 2 months (2–4). Of the 43 children who received a transplant from an alternative donor, 18 were in complete remission, 11 relapsed at a median of 5 months (3–9) after transplantation, and 14 (33%) died at a median of 3 months (2–5). The 5-year disease-free survival for children who actually received chemotherapy was 45·0% (SE 3·3) compared with 62·7% (6·6) for those who underwent haemopoietic-cell transplantation from a compatible related donor (p=0·08) and 34·3% (7·8) for those who underwent transplantation from an alternative donor (p=0·06), after adjustment for waiting time to transplant (figure 5).

Discussion
This international study has shown that availability of a compatible related donor for children with very-high-risk acute lymphoblastic leukaemia in first complete remission improves disease-free survival.

International collaboration was crucial to the success of the study. Enrolment of as many as 80 children per year allowed the planned sample size to be reached, which would not have been possible within one study group because very-high-risk disease is rare. Children at very high risk for treatment failure were identified as those who reported a 4-year event-free survival less than 35% after intensive chemotherapy protocols, based on the Berlin-Frankfurt-Münster approach.3,4 The participating study groups used common eligibility criteria, a common conditioning regimen for haemopoietic-cell transplantation, and front-line chemotherapy protocols which, although not identical, were very similar in the treatment schedule. This innovative strategy applied—in the rigorous frame of a prospective study—the principles of meta-analysis.19

Early family HLA-typing allowed for children to be assigned either chemotherapy or haemopoietic-cell transplantation, according to the availability of a compatible related donor. HLA-typing was known in 96% of eligible children, consistently across study groups, which resulted in a remarkably large set of assessable patients. Adherence to treatment allocation by genetic chance was carefully tracked; deviations from the assigned treatment group accounted for 18% of assessable patients, which compares favourably with prospective randomised studies of transplantation.24–27

Previous reports of patients with acute lymphoblastic leukaemia have compared treatments (chemotherapy and haemopoietic-cell transplantation) retrospectively.16–17 Major drawbacks of these studies have been absence of homogeneous high-risk features and selection bias that affected treatment allocation.18 The novelty of the present study was the prospective comparison. This study design relied both on an a-priori restrictive definition of very-high-risk acute lymphoblastic leukaemia (8% of the overall population), common to all participating study groups, and on genetic randomisation, ie, treatment allocation solely decided by availability of a suitable donor, which allowed an unbiased analysis by intention to treat.19,22

Standard randomisation was not feasible in this framework because haemopoietic-cell transplantation from a compatible related donor is widely regarded as the most promising treatment in children with very-high-risk acute lymphoblastic leukaemia, although no evidence of this benefit has been provided in a clinical trial.

The analysis by intention to treat showed a benefit of 16% in 5-year disease-free survival for children who had a compatible related donor compared with those who did not. This advantage was not shown in the survival rate. The gap between survival and disease-free survival was about 10% for children allocated chemotherapy, indicating that few could be rescued in second complete remission when relapse occurred after intensive chemotherapy. No gap was noted between survival and disease-free survival in those allocated haemopoietic-cell transplantation, which accords with the clinical observation that there is little or no room to rescue children with very-high-risk disease who relapse after transplant. The disease-free survival curves for the two treatment groups were similar and dropped strikingly up to 18 months, when they began to diverge. This finding accords with the clinical observation that treatment failure in chemotherapy patients, although higher, occurs later than in...
those receiving a transplant, for whom treatment-related complications increases the early death rate.

Divergence in disease-free survival was also noted in 43 children allocated chemotherapy but who deviated from the assigned treatment and underwent haemopoietic-cell transplantation from an alternative donor. These children were highly selected, since about half had presented with induction failure. Similarly, 22 children who were allocated related-donor haemopoietic-cell transplantation but continued with chemotherapy had a worse outcome compared with the others in the transplantation arm. This subset probably includes a mixture of children characterised by either the worst or the best expected prognosis. In fact, those who were either too compromised or relapsed too early to undergo haemopoietic-cell transplantation (four of 22 relapsed within 6 months after first complete remission) could be counted in the worst prognostic group, whereas those who were not judged to be eligible for transplantation by the treating doctor could be positively selected. Noticeably, 11 of 22 children only had a poor response to prednisone, which could be judged the mildest prognostic feature for very-high-risk disease.

Although the adjustment by waiting time to transplantation allowed for control of some selection bias, outcome by assigned treatment could not be directly compared with outcome by treatment actually received. Nevertheless, both analyses point in the same direction—i.e., an advantage of haemopoietic-cell transplantation from a related donor over chemotherapy for children with very-high-risk acute lymphoblastic leukaemia. This advantage is especially remarkable since chemotherapy yielded better results than the 35% disease-free survival expected on the basis of historical data available when this study was planned, highlighting an improvement over time of frontline protocols.

Ancillary analyses, focusing on subsets defined by the characteristics of very-high-risk disease, showed that children who were eligible for the study because of induction failure (irrespective of which other features they had) could benefit more from related-donor availability than could others. Conversely, those in the haemopoietic-cell transplantation arm who were at very high risk only because of a poor response to prednisone achieved the best outcome with yet less benefit over chemotherapy. The biological subgroups of children with t(9;22) and t(4;11) clonal abnormalities reported an intermediate outcome, with some apparent benefit from transplantation.

To the best of our knowledge, of studies comparing chemotherapy with haemopoietic-cell transplantation in patients with acute lymphoblastic leukaemia in first complete remission, the only report of an intention-to-treat analysis was by Wheeler for the Medical Research Council. In that study, the advantage of transplantation over chemotherapy, which was seen when assessed by treatment received, was not confirmed when assessed by treatment assigned; nevertheless, the study was retrospective, no stringent criteria for very-high-risk disease were used (13% of the population), and HLA typing was not known for 16% of patients. Another report from the Nordic Society for Pediatric Hematology and Oncology compared all patients receiving a transplant in first complete remission, irrespective of high-risk criteria, with matched chemotherapy controls. In that study, a significant advantage of haemopoietic-cell transplantation was reported. Our previous retrospective report, within the Associazione Italiana di Ematologia ed Oncologia Pediatria, showed a similar result, but was not significant.

The issue of whether patients with very-high-risk acute lymphoblastic leukaemia in first complete remission should undergo haemopoietic-cell transplantation from an alternative donor remains controversial and cannot be addressed by the present study, since this treatment option was not taken into account in the study design. The outcome of unrelated transplantation has greatly improved in recent years, mostly because of high-resolution HLA-typing and improved management of graft-versus-host disease.28,29 Even when assuming that final outcome after related and unrelated donor transplantation could be more similar, the effectiveness of any successful allogeneic transplantation has to be counterbalanced by late effects associated with the conditioning regimen and multiple immunological complications.30,31 Possibly, determination of minimal residual disease could allow better selection of those patients who need the immunological effect of an allograft to be cured, despite not having a compatible related donor.

In conclusion, this international prospective study, based on treatment allocation by genetic chance, provided evidence that children with very-high-risk acute lymphoblastic leukaemia benefit from related donor haemopoietic-cell transplantation compared with chemotherapy and showed that the gap between the two strategies increases as the risk profile of the patient worsens.

Contributors
A Balduzzi planned the study and wrote the report. M G Valsecchi was the study statistician and designed the study, did analyses, coordinated the trial data centre, and wrote the report. C Uderzo was the clinical coordinator of the study and wrote the report. P De Lorenzo was in charge of data pooling, study reporting, data checking and analyses, and wrote the report. T Klingebiel was the transplant coordinator for children in Germany and reviewed the report. C Peters was the transplant coordinator within the international Berlin-Frankfurt-Muenster study group and reviewed the report. J Stary coordinated the study in Czech Republic and reviewed the report. M S Felice coordinated the study in Argentina and reviewed the report. E Magyarosy coordinated the study in Hungary and reviewed the report. V Coster was the Italian chemotherapy coordinator, contributed to study planning, and reviewed the report. A Reiter reviewed the report. C Messina was the transplant coordinator for children in Italy and reviewed the report. H Gadner planned the study, coordinated the study in Austria, and reviewed the report. M Schrappe was the ALL Committee chairman within the international Berlin-Frankfurt-Muenster study group, planned the study, coordinated the study in Germany, and reviewed the report.
Conflict of interest statement
We declare that we have no conflict of interest.

Acknowledgements
We thank the “Comitato Maria Letizia Verga per lo studio e la cura della leucemia del bambino” for its continuous support; Joanna Upton for linguistic consultancy; and Andreina di Palma for secretarial assistance. This work was partly supported by grants from the Associazione Italiana Ricerca sul Cancro (AIRC; to MGV).

References
Direct decompressive surgical resection in the treatment of spinal cord compression caused by metastatic cancer: a randomised trial


Summary

Background The standard treatment for spinal cord compression caused by metastatic cancer is corticosteroids and radiotherapy. The role of surgery has not been established. We assessed the efficacy of direct decompressive surgery.

Methods In this randomised, multi-institutional, non-blinded trial, we randomly assigned patients with spinal cord compression caused by metastatic cancer to either surgery followed by radiotherapy (n=50) or radiotherapy alone (n=51). Radiotherapy for both treatment groups was given in ten 3 Gy fractions. The primary endpoint was the ability to walk. Secondary endpoints were urinary continence, muscle strength and functional status, the need for corticosteroids and opioid analgesics, and survival time. All analyses were by intention to treat.

Findings After an interim analysis the study was stopped because the criterion of a predetermined early stopping rule was met. Thus, 123 patients were assessed for eligibility before the study closed and 101 were randomised. Significantly more patients in the surgery group (42/50, 84%) than in the radiotherapy group (29/51, 57%) were able to walk after treatment (odds ratio 6·2 [95% CI 2·0–19·8] p=0·001). Patients treated with surgery also retained the ability to walk. Secondary endpoints were urinary continence, muscle strength and functional status, the need for corticosteroids and opioid analgesics, and survival time. All analyses were by intention to treat.

Interpretation Direct decompressive surgery plus postoperative radiotherapy is superior to treatment with radiotherapy alone for patients with spinal cord compression caused by metastatic cancer.

Introduction

Metastatic epidural spinal cord compression (MESCC) is a debilitating and common complication of cancer, occurring in 5–14% of cancer patients. More than 20 000 new cases are reported every year in the USA.1,2 Acute onset of MESCC needs immediate treatment.3,4 Standard treatment for MESCC consists of corticosteroids and radiotherapy,5,6 with which only about 50% of patients are able to walk and few non-ambulatory patients ever walk again.7,8 The role of surgery in the management of MESCC has not been established. Before radiation became available, surgery (in the form of simple laminectomy) was the only treatment. With the introduction of radiotherapy, results with laminectomy plus radiation did not seem to differ from results with radiation alone. Surgical treatment was largely abandoned when several retrospective studies9–16 and a small randomised trial17 did not show any benefit for laminectomy alone or in combination with radiotherapy. However, laminectomy might not be the best operation for MESCC. Most spinal metastases causing MESCC are located in the vertebral body, anterior to the spinal cord.18 Laminectomy involves the removal of posterior elements of the spinal column and does not remove tumour, and thus often does not result in immediate decompression. Furthermore, the procedure can cause destabilisation of the spine because often only the posterior elements are intact and removal of these elements causes instability.

In the 1980s, another type of surgical procedure was developed for the treatment of MESCC. The tumour was removed and immediate circumferential decompression was achieved, usually through an anterior approach. When needed, reconstruction of the spine intraoperatively was possible to provide immediate stabilisation. Several uncontrolled surgical series19–24 and a meta-analysis25 reported that direct decompressive surgery, with or without postoperative radiotherapy, was superior to radiation alone. However, these non-randomised studies were subject to patient selection bias, heterogeneous tumour types, unclear inclusion criteria, and imprecise endpoints. Consequently, these studies have not established direct decompressive surgery as an effective treatment. To determine the value of surgery in the management of MESCC, we undertook a randomised trial comparing the efficacy of direct decompressive surgery plus postoperative radiotherapy with that of radiotherapy alone.

Methods

Patients

Patients at least 18 years old with a tissue-proven diagnosis of cancer (not of CNS or spinal column origin) and MRI evidence of MESCC were eligible for the study.


Published online July 21, 2005
DOI:10.1016/S0140-6736(05)66954-1

See Comment page 609

Department of Surgery (Neurosurgery) (R A Patchell MD, P A Tibbs MD, B Young MD), Department of Neurology (R A Patchell), Department of Statistics (R J Kryscio PhD), and Department of Radiation Medicine (M Mohiuddin MD), University of Kentucky Medical Center, Lexington, KY, USA; Department of Radiation Oncology, University of Maryland Medical School, Baltimore, MD, USA (W F Regine MD); Duke Institute on Care at the End of Life, Duke University Medical Center, Durham, NC, USA (R Payne MD); and Department of Neurosurgery, Rhode Island Hospital, Providence, RI, USA (S Saris MD)

Correspondence to: Dr R A Patchell, Division of Neurosurgery, University of Kentucky Medical Center, 800 Rose Street, Lexington, Kentucky 40536, USA

rpatchell@aul.com
MESCC was defined radiographically as a true displacement of the spinal cord (by an epidural mass) from its normal position in the spinal canal. Patients also had to have at least one neurological sign or symptom (including pain) and not have been totally paraplegic for longer than 48 h before study entry. The MESCC had to be restricted to a single area, which could include several contiguous spinal or vertebral segments. Patients with a mass that compressed only the cauda equina or spinal roots were excluded. Those with multiple discrete compressive lesions were also excluded (unless they had one area of compression and multiple non-compressive lesions). Patients with certain radiosensitive tumours (lymphomas, leukaemia, multiple myeloma, and germ-cell tumours) were excluded, as were patients with pre-existing or concomitant neurological problems not related directly to their MESCC (eg, brain metastases). Additionally, patients with previous MESCC and those who had received spinal radiation such that they were unable to receive the study dose were excluded. Patients also had to have a general medical status good enough to be acceptable surgical candidates and an expected survival of at least 3 months.

The study was approved by the institutional review boards of the University of Kentucky and other participating institutions, and written informed consent was obtained from all patients before study entry.

**Procedures**

The study was a randomised, multi-institutional, non-blinded trial with two treatment groups. Before randomisation, all patients had imaging of the entire spinal cord. The imaging technique consisted of MRI with whole spine sagittal T1 and T2 imaging and axial T1 imaging. Additional MRI techniques were used as clinically appropriate. There was a central review of all MRI scans for confirmation of MESCC.

When diagnosed, all patients were given 100 mg dexamethasone immediately, then 24 mg every 6 h until the start of radiotherapy or surgery. Corticosteroids were then reduced and continued until completion of radiotherapy. Patients with severe diabetes or other relative contraindications to high-dose corticosteroids were treated with reduced doses when appropriate.

Before randomisation, patients were stratified according to treating institution, tumour type, ambulatory status, and relative stability of the spine. Spinal stability was ascertained according to Cybulski’s guidelines. Patients with pathological spine fractures or evidence of bone in the spinal canal were also judged to have spinal instability. Randomisation within strata by permutated blocks was done separately at each institution with a computerised technique, which ensured immediate randomisation at study entry. The study was undertaken by the Bluegrass Neuro-Oncology Consortium with seven participating institutions (University of Kentucky [n=70 patients], MD Anderson [n=14], Brown University [n=12], University of Alabama-Birmingham [n=2], University of Michigan [n=1], University of Pittsburgh [n=1], University of South Florida [n=1]).

For patients randomised to the radiation group, radiotherapy was started within 24 h after randomisation. The total dose was 30 Gy given in ten fractions (3·0 Gy×10 fractions). Treatments were delivered to a port that encompassed one vertebral body above and below the visible lesion. There was a central review of radiotherapy treatment plans to monitor protocol compliance. Patients allocated to surgery were operated on within 24 h after randomisation. The protocol did not specify operative techniques or fixation devices. However, the aim of surgery was to provide immediate direct circumferential decompression of the spinal cord. The operation was tailored for each patient depending on the level of the spine involved and the patient’s circumstances. In general, for anteriorly-located tumours the approach in the cervical spine was anterior, and in the thoracic and lumbar spine, depending on the tumour location, the approach was through a transversecotomopy or anterior approach. For laterally-located tumours, a lateral approach was used, and for posteriorly-located tumours, a laminectomy was done and any other posterior elements involved were removed. Stabilisation of tumours in all locations was performed if spinal instability was present; cement (methyl methacrylate), metallic rods, bone grafting, or other fixation devices were used. Within 1 month of treatment Philip Tibbs reviewed operative reports and William Regine reviewed plans for post-surgery radiotherapy to monitor protocol compliance. Patients were given radiotherapy, as in the radiation group, within 14 days after surgery. Steroids were given on the same schedule for both groups.

Patients had neurological assessments before treatment, weekly during radiotherapy, and within 1 day after completion of treatment. Patients then had regular study follow-up assessments every 4 weeks until the end of the trial or death. Patients were also reassessed at any time they had symptoms suggestive of neurological progression.

The primary endpoint of the study was the ability to walk after treatment. A patient was deemed ambulatory if he or she could take at least two steps with each foot unassisted (4 steps total), even if a cane or walker was needed. We assessed ambulatory status in two ways, and both methods were prespecified. The combined ambulatory rate was the percentage of patients who maintained or regained the ability to walk immediately after completion of radiotherapy and quantified the initial success rate of treatment. Ambulatory time after treatment was a measure of long-term success. Secondary endpoints were urinary continence, changes in Frankel functional scale scores and American Spinal Injury Association (ASIA) motor scores, and use of...
corticosteroids and opioid analgesics. Corticosteroid use was assessed by calculating and comparing mean daily dexamethasone equivalent doses. Pain relief was assessed by calculating and comparing mean daily morphine equivalent doses. Survival time after treatment was also recorded. All time dependent endpoints were measured from the day of randomisation until death or last follow up.

**Statistical analysis**

Results from previous uncontrolled studies have suggested that the expected combined post-treatment ambulatory rate in patients treated with radiation alone is about 45%, and uncontrolled surgical series, in which direct decompressive surgery was used, reported post-treatment ambulatory rates of about 75%. These studies suggest that the advantage from surgery is an additional 30% increase in the post-treatment ambulatory rate compared with radiation alone. To determine sample size for this study, we used a more conservative expected difference in post-ambulatory rate between study groups of 20%. With this assumption and with 100 patients in each treatment group, the chance of achieving overall statistically significant results at the p<0.05 level, using a two-sided test, was 82%. The study design also included provision for an interim analysis to be done at the halfway point (after 100 patients were entered into the trial) according to the O’Brien-Fleming rule. An intention-to-treat analysis was used throughout. Multivariate analyses were based on a Cox regression model. The covariates used were treatment group, age, sex, primary tumour type, spinal level involved, predominant position of metastasis in vertebra, stability of spine, Frankel and ASIA scores at study entry, length of time motor symptoms associated with cord compression were present before treatment, and length of time between diagnosis of the primary tumour and development of cord compression. All these analyses were prespecified.

**Role of the funding source**

The sponsor had no role in study design, data collection, data analysis, data interpretation, or 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**

We compared combined ambulatory rates after treatment between the two groups using a Cochran-Mantel-Haenszel statistic based on ambulatory status. This comparison yielded a p value of 0.001, which fell below the predetermined significance level for early termination of the trial according to the O’Brien-Fleming rule (p<0.0054). Because of proven superiority of surgical treatment, the data safety and monitoring committee deemed the trial should be stopped early.

Between Sept 1, 1992, and Dec 31, 2002, 123 patients were assessed for eligibility and, of these, 101 were entered into the trial before it closed (figure 1). Protocol violations occurred with five patients. In the surgery group, three patients did not receive postoperative radiotherapy. Between Sept 1, 1992, and Dec 31, 2002, 123 patients were assessed for eligibility and, of these, 101 were entered into the trial before it closed (figure 1). Protocol violations occurred with five patients. In the surgery group, three patients did not receive postoperative radiotherapy.

**Table 1: Baseline characteristics of study patients**
radiotherapy and a fourth patient stopped radiotherapy before receiving the complete course. In the radiation group, one patient was treated with surgery as well as postoperative radiotherapy. Table 1 shows baseline characteristics of patients entered in the study. Overall median follow-up times were 102 days (IQR 0–1940) in the surgery group and 93 days (0–1117 days) in the radiation group (p=0·10).

The combined post-treatment ambulatory rate in the surgery group was 84% (42/50) and 57% (29/51) in the radiation group. Ambulatory rates were compared between the two groups using a Cochran-Mantel-Haenszel statistic after stratifying by pretreatment ambulatory status. This analysis yielded a p value of 0·001 with an odds ratio of 6·2 (95% CI 2·0–19·8). Patients in the surgery group retained the ability to walk for significantly longer than did those in the radiation group (median 122 days vs 13 days, p=0·003; figure 2). Multivariate analysis showed surgery (p=0·0017) and pretreatment Frankel score (p=0·0008) to be associated with longer ambulatory time.

In the subgroup of patients who could walk at study entry, 94% (32/34) in the surgery group continued to walk after treatment compared with 74% (26/35) in the radiation group (p=0·024). Patients in the surgical group were able to walk for a median of 153 days compared with 54 days in the radiation group (odds ratio 1·82 [95% CI 1·08–3·12] p=0·024; figure 3). Multivariate analysis showed surgery (p=0·0048), Frankel score (p=0·016), and breast primary tumour (p=0·029) to be associated with longer ambulatory times.

32 patients (16 in each group) entered the study unable to walk; of these, ten patients (62%) in the surgery group regained the ability to walk compared with three patients (19%) in the radiation group (p=0·012). Additionally, non-ambulatory patients treated with surgery walked for a median of 59 days compared with a median of 0 days for patients in the radiation group (p=0·04).

Surgical treatment resulted in significant differences in maintenance of continence, muscle strength (ASIA

**Table 2: Secondary endpoints**

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Surgery group (n=51) median</th>
<th>Radiation group (n=51) median</th>
<th>Relative risk*</th>
<th>95% CI*</th>
<th>P*</th>
<th>Significant predictors**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance of continence</td>
<td>17 days</td>
<td>156 days</td>
<td>0·47</td>
<td>0·25–0·87</td>
<td>0·016</td>
<td>Surgery RR=0·51 (0·29–0·90) Baseline Frankel Score RR=0·56 (0·3–0·73)</td>
</tr>
<tr>
<td>Maintenance of ASIA score</td>
<td>72 days</td>
<td>566 days</td>
<td>0·28</td>
<td>0·13–0·61</td>
<td>0·001</td>
<td>Surgery RR=0·30 (0·14–0·62) Baseline Frankel Score RR=0·65 (0·46–0·91) Stable Spine RR=0·43 (0·22–0·83) Cervical Spinal Level RR=0·49 (0·26–0·90)</td>
</tr>
<tr>
<td>Maintenance of Frankel score</td>
<td>72 days</td>
<td>566 days</td>
<td>0·24</td>
<td>0·11–0·54</td>
<td>0·0006</td>
<td>Surgery RR=0·28 (0·12–0·54) Stable Spine RR=0·39 (0·20–0·75) Cervical Spinal Level RR=0·53 (0·74–0·98) Baseline Frankel Score RR=0·62 (0·44–0·88)</td>
</tr>
<tr>
<td>Survival time</td>
<td>100 days</td>
<td>126 days</td>
<td>0·60</td>
<td>0·38–0·96</td>
<td>0·033</td>
<td>Surgery RR=0·60 (0·40–0·89) Breast Primary Tumour RR=0·29 (0·13–0·62) Lower Thoracic Spinal Level RR=0·65 (0·43–0·98)</td>
</tr>
</tbody>
</table>

*Based on a Cox model with all covariates included. **Based on a Cox model with only significant predictors included (stepwise selection).
scores), functional ability (Frankel scores), and increased survival time (table 2). The surgical group also had a substantial reduction in use of corticosteroids and opioid analgesics. In the surgery group, the median mean daily dexamethasone equivalent dose was 1·6 mg (IQR 0·1–44·0) compared with 4·2 mg (0·0–50·0) in the radiation group (p=0·0093). In the surgery group, the median mean daily morphine equivalent dose was 0·4 mg (0·0–60·0) compared with 4·8 mg (0·0–200·0) in the radiation group (p=0·002).

The 30-day mortality rates were 6% in the surgery group and 14% in the radiation group (p=0·32). 30-day morbidity rates were calculated by deterioration in ASIA and Frankel scores. At 30 days, surgery group patients maintained or improved their pretreatment ASIA muscle strength scores at a significantly (p=0·0064) higher rate than did patients in the radiation group (86% vs 60%). Also, at day 30 after treatment, the percentage of patients with Frankel scores at or above study entry level was significantly (p=0·0008) higher in the surgery group than in the radiation group (91% vs 61%). Surgery did not result in prolonged hospitalisation: the median hospital stay was 10 days in both the surgery group (IQR 2–51 days) and the radiation group (0–41 days; p=0·86). Extended hospital stays (greater than 20 days) occurred in seven patients in the surgery group and 11 in the radiation group.

Ten patients in the radiation group (20%) had a substantial decline in motor strength during radiotherapy and crossed over to receive surgery. The primary tumour histologies of these crossover patients were: lung (four), gastrointestinal (two), prostate (one), other genitourinary (two), sarcoma (one). At the time of surgery, none of these patients could walk. Three (30%) regained the ability to walk. Of the crossover patients, four (40%) had surgical complications consisting of three wound infections and one failure of fixation that needed additional surgery.

Discussion
This prospective randomised trial shows that patients with MESCC treated with direct decompressive surgery plus postoperative radiotherapy retain the ability to walk for longer and regain the ability more often than do patients treated with radiotherapy alone. Surgery allows most patients to remain ambulatory for the remainder of their lives, whereas patients treated with radiation alone spend a substantial proportion of their remaining time paraplegic. Surgical treatment also results in increased survival time. The better survival time in the surgical group was probably because a greater proportion of patients in this group were ambulatory and remained so for longer than those in the radiation group. Therefore, patients in the surgery group were less susceptible to infections, blood clots, and other problems that result in the death of paraplegic patients. Surgical treatment also reduces the need for corticosteroids and opioid pain relief.

The cause of damage to the spinal cord from compression is complex and multifactorial, although two mechanisms predominate. Direct compression results in oedema, venous congestion, and demyelination. If the compression is of short duration, the effects are reversible; remyelination and recovery of function is possible. However, with prolonged compression, secondary vascular injury occurs with infarction of the spinal cord. After this type of injury, no meaningful recovery is possible. Surgical decompression is immediate, whereas radiotherapy takes several days to have an effect. Surgery was probably able to provide relief from compression before irreversible vascular injury occurred in a substantial number of patients in our study. Thus a higher percentage of patients were able to recover function in this treatment group, which explains the number of patients who regained the ability to walk after treatment and the initial success of the treatment. The fact that surgery preserved the ability to walk much longer than did radiation is because of the ability of surgery to remove tumour. In patients treated with radiation alone, tumour was left behind and regrowth with secondary compression was more likely.

With any surgical procedure, operative mortality and morbidity have to be weighed against any possible benefit from surgery. Surprisingly, surgery did not result in an increase in length of hospital stay. 30-day mortality rates did not differ significantly between the two groups, and 30-day morbidity was substantially worse in the radiation group. Therefore, there was no excess mortality or morbidity due to surgery.

A possible limitation of the study was patient selection bias. Any study that has exclusion criteria selects a subset of the total number of patients with a disease for study. Our study was designed to reflect the way patients with MESCC were being treated routinely in community and academic medical centres. The patient population studied consisted of those patients for whom surgery would be regarded as a realistic treatment option. Patients with very radiosensitive tumours, multiple areas of spinal cord compression, or total paraplegia for longer than 48 h were excluded. Therefore, the results of this trial cannot be used to justify surgery in all patients with MESCC and apply only to patients comparable to those included in our study. Even in this group of patients, reasonable clinical judgment should be used in the selection of patients for surgery.

Our trial shows that surgery is an effective treatment for MESCC, but should surgery be the initial treatment for all patients similar to those in the study who have operable lesions? An argument could be made that ambulatory patients should be treated with radiation first, and surgery reserved for those patients who progress. This approach would reduce the number of surgeries done and might be as effective. However, the results of our trial do not lend support to the use of
radiation alone as first-line treatment. In the subgroup of patients who were ambulatory at the start of therapy, initial treatment with surgery was significantly better at preserving the ability to walk. Further evidence comes from the radiation patients who crossed over to surgery. These patients were treated initially with radiation and then operated on when they failed radiation and lost the ability to walk. In these patients, only 30% regained the ability to walk. This result compares unfavourably with the 62% post-treatment ambulatory rate of the patients who were originally not able to walk and received surgery as their first treatment. Clearly, first-line treatment with surgery was superior. For these reasons, the best treatment for spinal cord compression caused by metastatic cancer is surgery as initial treatment followed by radiotherapy.

Contributors
R A Patchell had the idea for the study and designed and wrote the protocol in collaboration with colleagues from various disciplines (P A Tibbs and B Young for neurosurgery, W F Regine and M Mohiuddin for radiation oncology, and R J Kryscio for statistics). R Payne and S Karis were principal investigators at their institutions and made substantial contributions of patients to the study.

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

Acknowledgments
This study was supported by grants from the National Cancer Institute (ROI CA55256) and the National Institute for Neurological Disorders and Stroke (K24 NS502180). We thank the following individuals and institutions for contributing patients to the study: E J Dropcho, University of Indiana, Indianapolis, IN, USA; H S Greenberg, University of Michigan, Ann Arbor, MI, USA; B Zacharia, University of Pittsburgh, Pittsburgh, PA, USA; and R Siegal, University of Pittsburgh, South Florida, Tampa, FL, USA.

References
NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses

Montserrat García-Closas, Núria Malats, Debra Silverman, Mustafa Dosemeci, Manolis Kogevinas, David W Hein, Adonina Tardón, Consol Serra, Alfredo Carrato, Reina García-Closas, Josep Lloreta, Gemma Castaño-Vinyals, Meredith Yeager, Robert Welch, Stephen Chanock, Nilanjan Chatterjee, Sholom Wacholder, Claudine Samanic, Montserrat Torà, Francisco Fernández, Francisco X Real, Nathaniel Rothman

Summary

Background Many reported associations between common genetic polymorphisms and complex diseases have not been confirmed in subsequent studies. An exception could be the association between NAT2 slow acetylation, GSTM1 null genotype, and bladder-cancer risk. However, current evidence is based on meta-analyses of relatively small studies (range 23–374 cases) with some evidence of publication bias and study heterogeneity. Associations between polymorphisms in other NAT and GST genes and bladder-cancer risk have been inconsistent.

Methods We investigated polymorphisms in NAT2, GSTM1, NAT1, GSTT1, GSTM3, and GSTP1 in 1150 patients with transitional-cell carcinoma of the urinary bladder and 1149 controls in Spain; all the participants were white. We also carried out meta-analyses of NAT2, GSTM1, and bladder cancer that included more than twice as many cases as in previous reports.

Findings In our study, the odds ratios for bladder cancer for individuals with deletion of one or two copies of the GSTM1 gene were 1.2 (95% CI 0.8–1.7) and 1.9 (1.4–2.7) respectively (p for trend <0.0001). Compared with NAT2 rapid or intermediate acetylators, NAT2 slow acetylators had an increased overall risk of bladder cancer (1.4 [1.2–1.7]) that was stronger for cigarette smokers than for never smokers (p for interaction 0.008). No significant associations were found with the other polymorphisms. Meta-analyses showed that the overall association for NAT2 was robust (p<0.0001), and case-only meta-analyses provided support for an interaction between NAT2 and smoking (p for interaction 0.009). The overall association for GSTM1 was also robust (p<0.0001) and was not modified by smoking status (p=0.86).

Interpretation The GSTM1 null genotype increases the overall risk of bladder cancer, and the NAT2 slow-acetylator genotype increases risk particularly among cigarette smokers. These findings provide compelling evidence for the role of common polymorphisms in the aetiology of cancer.

Relevance to practice Although the relative risks are modest, these polymorphisms could account for up to 31% of bladder cancers because of their high prevalence.

Introduction

The inability to replicate results on many associations between common genetic polymorphisms and complex diseases has raised scepticism in this area of research. One of the few exceptions could be the association between the risk of bladder cancer and polymorphisms in two carcinogen-detoxification genes—NAT2 and GSTM1. However, evidence for an association relies on analyses of pooled data and meta-analyses of relatively small studies (range 23–374 patients, average about 100 per study), and concern has been raised about publication bias and heterogeneity of results. Tobacco smoking is an important cause of bladder cancer, and previous analyses have suggested that the relative risk from smoking is stronger for NAT2 slow acetylators than for rapid or intermediate acetylators. This interaction is biologically plausible, since aromatic amines, which are thought to be the most important class of bladder carcinogens in tobacco smoke, are detoxified by NAT2. However, epidemiological evidence for this interaction is even weaker than for the overall genotype association. Associations between bladder-cancer risk and polymorphisms in other carcinogen-detoxification genes such as NAT1 and other glutathione-S-transferases have been less frequently explored, with inconsistent results across studies.

We report results on the associations of polymorphisms in NAT and GST genes with bladder-cancer risk and their interaction with cigarette smoking among participants in the Spanish Bladder Cancer Study. This study was designed to have adequate statistical power for rigorous evaluation of the proposed associations between genetic variation in NAT2 and GSTM1 and bladder-cancer risk. We also report meta-analyses of NAT2, GSTM1, smoking, and bladder cancer that include more than twice as many patients as in previous reports.

www.thelancet.com Vol 366 August 20, 2005 649


See Comment page 610

Division of Cancer Epidemiology and Genetics (M García-Closas MD, D Silverman ScD, M Dosemeci PhD, N Chatterjee PhD, S Wacholder PhD, C Samanic MS, N Rothman MD) and Core Genotype Facility at the Advanced Technology Center (M Yeager PhD, R Welch MBA, S Chanock MD), National Cancer Institute, Department of Health and Human Services, Bethesda, MD, USA; Institut Municipal d’Investigació Médica, Barcelona, Spain (N Malats MD, Prof M Kogevinas MD, J Lloreta MD, G Castaño-Vinyals BSc, M Torà MD, F Fernández BSc, Prof FX Real MD); Department of Pharmacology and Toxicology and James Graham Brown Cancer Center, University of Louisville School of Medicine, KY, USA (Prof W Hein PhD); Universidad de Oviedo, Oviedo, Spain (A Tardón MD); Consorcio Hospitalari Parc Taulí, Sabadell, Spain (C Serra MD); Universitat Pompeu Fabra, Barcelona, Spain (J Lloreta MD); and Department of Pathology, Hospital del Mar-IMAS, Barcelona, Spain (J Lloreta)

Correspondence to: Dr Montserrat García-Closas, Hormonal and Reproductive Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 6120 Executive Blvd, Room 7076, MSC 7234, Rockville, MD 20852–7234, USA montse@nih.gov
To limit heterogeneity, 16 cases with neoplasias of 11 controls were excluded because of low amounts of buccal-cell sample for DNA extraction. Seven cases and 1173 (92%) controls provided a blood or bladder sample for DNA extraction. In total, 1219 cases and 1271 controls interviewed, 1188 (97%) of whom were white and 84% took part in the study and were interviewed. Of the 1219 cases and 1271 controls participating hospitals with diagnoses thought to be potential risk factors for bladder cancer for cases and controls was collected by means of computer-assisted personal interviews during the hospital admission. 84% of eligible cases and 88% of eligible controls agreed to take part in the study and were interviewed. Of the 1219 cases and 1271 controls interviewed, 1188 (97%) cases and 1173 (92%) controls provided a blood or buccal-cell sample for DNA extraction. Seven cases and 11 controls were excluded because of low amounts of DNA. To limit heterogeneity, 16 cases with neoplasias of non-transitional histology and six non-white individuals (five cases, one control) were excluded from the analyses. 15 individuals (seven cases, eight controls) with missing information on smoking status and seven (three cases, four controls) with DNA quality-control difficulties were also excluded from the analyses. Thus, the final study population available for analysis was 1150 cases and 1149 controls, all of whom were white.

Participants were classified as never smokers if they had smoked fewer than 100 cigarettes in their lifetime and ever smokers otherwise. Every smoker was further classified as regular smokers if they had smoked at least one cigarette per day for 6 months or longer and occasional smokers otherwise. We defined current smokers as those regular smokers who had smoked within a year of the reference date; individuals who had smoked regularly but who had stopped smoking more than 1 year before the reference date were defined as former smokers. Most (81%) smokers of known tobacco type reported smoking black tobacco. In addition, the risks of bladder cancer in relation to the risk for never smokers were similarly raised among smokers of black tobacco alone, smokers of black and blond tobacco, and smokers of unknown tobacco type (data not shown). These subgroups were therefore combined as known or likely black-tobacco smokers. We obtained informed consent from potential participants in accordance with the National Cancer Institute and local institutional review boards.

Methods

Study population

The Spanish Bladder Cancer Study is a hospital-based case-control study based in 18 hospitals in five areas in Spain (Asturias, Barcelona metropolitan area, Vallsës/ Bages, Alicante, and Tenerife). Eligible “cases” were aged 21–80 years and had newly diagnosed, histologically confirmed carcinoma of the urinary bladder in 1998–2001. Diagnostic slides from each patient were reviewed by a panel of expert pathologists to confirm the diagnosis and to ensure uniformity of classification criteria, based on the 1998 system of WHO and the International Society of Urological Pathology. Controls were selected from patients admitted to participating hospitals with diagnoses thought to be unrelated to the exposures of interest, such as tobacco use. The distribution of reasons for hospital admission was: 37% hernias, 11% other abdominal surgery, 23% fractures, 7% other orthopaedic problems, 12% hydrocoele, 4% circulatory disorders, 2% dermatological disorders, 1% ophthalmological disorders, and 3% other diseases. Controls were individually matched to the cases for age at interview within 5-year categories, sex, ethnic origin, and region. Information on known or potential risk factors for bladder cancer for cases and controls was collected by means of computer-assisted personal interviews during the hospital admission. 84% of eligible cases and 88% of eligible controls agreed to take part in the study and were interviewed. Of the 1219 cases and 1271 controls interviewed, 1188 (97%) cases and 1173 (92%) controls provided a blood or buccal-cell sample for DNA extraction. Seven cases and 11 controls were excluded because of low amounts of DNA.

Table 1: Characteristics of study population

<table>
<thead>
<tr>
<th></th>
<th>Cases (n=1150)</th>
<th>Controls (n=1149)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demography</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age (SD), years</td>
<td>66 (10)</td>
<td>65 (10)</td>
</tr>
<tr>
<td>Female</td>
<td>146 (13%)</td>
<td>147 (13%)</td>
</tr>
<tr>
<td>Male</td>
<td>1004 (87%)</td>
<td>1002 (87%)</td>
</tr>
<tr>
<td>Educational attainment*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than primary</td>
<td>525 (46%)</td>
<td>539 (47%)</td>
</tr>
<tr>
<td>Primary and less than high school</td>
<td>452 (39%)</td>
<td>437 (38%)</td>
</tr>
<tr>
<td>At least high school</td>
<td>156 (14%)</td>
<td>154 (13%)</td>
</tr>
<tr>
<td>Other</td>
<td>14 (1%)</td>
<td>14 (1%)</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>159 (14%)</td>
<td>138 (29%)</td>
</tr>
<tr>
<td>Occasional</td>
<td>50 (4%)</td>
<td>88 (8%)</td>
</tr>
<tr>
<td>Regular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>474 (41%)</td>
<td>458 (40%)</td>
</tr>
<tr>
<td>Current</td>
<td>467 (41%)</td>
<td>265 (23%)</td>
</tr>
<tr>
<td>Type of tobacco smoked†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blond tobacco only</td>
<td>92 (10%)</td>
<td>114 (16%)</td>
</tr>
<tr>
<td>Black tobacco only</td>
<td>383 (41%)</td>
<td>281 (39%)</td>
</tr>
<tr>
<td>Both types of tobacco</td>
<td>284 (30%)</td>
<td>194 (27%)</td>
</tr>
<tr>
<td>Unknown tobacco type</td>
<td>182 (15%)</td>
<td>132 (18%)</td>
</tr>
</tbody>
</table>

Unless otherwise stated, data are number of participants (%). *Information on education missing for three cases and five controls. †Defined only for regular smokers; information on type of tobacco missing for two controls.

Procedures

DNA for genotype assays was extracted from leucocytes with the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA) or MASSArray (Sequenom, San Diego, CA, USA), MGB Eclipse (Epoch Biosciences, Bothel, WA, USA), or TaqMan (Applied Biosystems, Foster City, CA, USA). Genotyping was done at the Core Genotyping Facility of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, with the TaqMan (Applied Biosystems, Foster City, CA, USA). Genotype assays were done for NAT1 (Ex1-88A>T rs1057126, Ex1-81A>C rs15561, V149I rs4987076, R187Q rs4986782, R187* rs5030839, R33*, D251V, R64W), NAT2 (K268R rs1208, G286E rs1799931, R187Q rs1801279, Y94Y rs1041983, I114T rs1801280, L161I rs1799929, R197Q rs1799930), GSTM1 deletion (SNP500Cancer ID:GSTM1-02), GSTT1 deletion (SNP500Cancer ID:GSTT1-02), GSTP1 (I105V rs947894, A114V), and GSTM3 (V224I rs7483, IVS7–30G>T rs1573234). All genotypes studied were in Hardy-Weinberg equilibrium among the control population. Duplicate quality-control samples showed 100% agreement for all but four assays (range 98·2% to 99·6%).
Information from the NAT1 and NAT2 single-nucleotide polymorphisms analysed in this study was used to assign the most likely NAT1 and NAT2 alleles previously identified in human populations. Individuals homozygous for NAT2 rapid-acetylator alleles (NAT2*4, NAT2*11A, NAT2*12A, NAT2*12B, NAT2*12C, NAT2*13) were classified as rapid-acetylator phenotype; individuals homozygous for slow-acetylator alleles were classified as slow-acetylator phenotype, and heterozygous individuals (one rapid and one slow NAT2 allele) were classified as intermediate-acetylator phenotype. Individuals with missing information for four rare NAT1 single-nucleotide polymorphisms (R187*, R33*, D251V, and R64W with more than 99% homozygous wild-type individuals) were assumed to be homozygous for NAT1*4. On the basis of previous studies, the NAT1*10 allele was deemed to be the “at risk” allele. GSTM1 genotypes were defined as null (–/–) if a deletion was present in both copies of the gene and present if one (+/+ or none (+/–) of the copies had a deletion. The two GSTP1 (I105V and A114V) and GSTM3 (V224I and IVS7 -30G>T) polymorphisms investigated were in strong linkage disequilibrium (D'=1.0, R²=0.10 and D'=1.0, R²=0.68, respectively). Participants were classified according to the presence of three GSTP1 variants that have been found to encode functionally differing GSTP1 proteins: GSTP1*A (105 Ile; 114 Ala), GSTP1*B (105 Val; 114 Ala), and GSTP1*C (105 Val; 114 Val).  

### Statistical analysis

Odds ratios, as measure of relative risk, and 95% CI were estimated from logistic regression models, with adjustment for sex, age at interview, region, and smoking status (never, occasional, former, or current). These unconditional models provided estimates similar to those from conditional logistic regression models for individually matched pairs. Interactions between genotypes and smoking habits were also investigated by the semiparametric maximum likelihood estimator method (SPMLE) to allow estimation of parameters under the assumption of genotype–smoking and genotype–sex independence in the source population. This assumption is supported by strong evidence from previous studies for independence of NAT2 and GSTM1 genotypes from cigarette smoking and sex in the control populations. Tests for multiplicative interaction were used to assess whether the genotype odds ratios within categories of smoking habits differed significantly from each other, or whether smoking odds ratios within genotype categories differed significantly from each other. When no multiplicative interactions were present, we also tested for additive interactions, because departures from the additive model can exist in the absence of multiplicative interactions and they might have biological implications under certain biological models. The synergy index was used as a measure of additive interaction and its CI was calculated by use of previously published formulae.  

We updated previous meta-analyses on NAT2, GSTM1, and bladder cancer and used similar selection criteria for studies—ie, case-control studies in the general population. Relevant studies published up to February, 2005, were identified in a MEDLINE search. For studies of NAT2 and GSTM1 included in previously published meta-analyses, we used data from

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Cases</th>
<th>Controls</th>
<th>Odds ratio (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NAT2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid</td>
<td>55</td>
<td>66</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td>351</td>
<td>427</td>
<td>1.0 (0.71–1.5)</td>
<td>0.97</td>
</tr>
<tr>
<td>Slow</td>
<td>728</td>
<td>637</td>
<td>1.0 (0.92–1.1)</td>
<td>0.10</td>
</tr>
<tr>
<td>Slow vs rapid/intermediate</td>
<td></td>
<td></td>
<td>1.0 (1.21–1.7)</td>
<td>0.0002</td>
</tr>
<tr>
<td><strong>GSTM1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+/+</td>
<td>70</td>
<td>107</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>+/–</td>
<td>352</td>
<td>454</td>
<td>1.0 (0.81–1.3)</td>
<td>0.38</td>
</tr>
<tr>
<td>–/–</td>
<td>716</td>
<td>571</td>
<td>1.0 (1.42–2.7)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Null vs present</td>
<td></td>
<td></td>
<td>1.0 (1.42–2.0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>NAT2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAT2<em>4/NAT2</em>14</td>
<td>585</td>
<td>574</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>NAT2<em>10/NAT2</em>14</td>
<td>327</td>
<td>326</td>
<td>1.0 (0.81–1.2)</td>
<td>0.62</td>
</tr>
<tr>
<td>NAT2<em>12/NAT2</em>10</td>
<td>53</td>
<td>42</td>
<td>1.0 (0.81–1.2)</td>
<td>0.48</td>
</tr>
<tr>
<td><strong>GSTM1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTP1*I105V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re/Reval</td>
<td>486</td>
<td>488</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Re/Val</td>
<td>525</td>
<td>531</td>
<td>1.0 (0.81–1.2)</td>
<td>0.93</td>
</tr>
<tr>
<td>Val/Val</td>
<td>330</td>
<td>119</td>
<td>1.0 (0.89–1.2)</td>
<td>0.35</td>
</tr>
<tr>
<td>GSTP1*I114V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala/Ala</td>
<td>966</td>
<td>917</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Ala/Val</td>
<td>113</td>
<td>85</td>
<td>1.0 (1.01–1.8)</td>
<td>0.07</td>
</tr>
<tr>
<td>Val/Val</td>
<td>4</td>
<td>5</td>
<td>0.9 (0.23–4.1)</td>
<td>0.85</td>
</tr>
<tr>
<td>GSTP1*I114V/A114V combination†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+/+</td>
<td>456</td>
<td>441</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>+/–</td>
<td>409</td>
<td>402</td>
<td>1.0 (0.81–1.2)</td>
<td>0.92</td>
</tr>
<tr>
<td>–/–</td>
<td>95</td>
<td>69</td>
<td>1.0 (1.01–1.8)</td>
<td>0.09</td>
</tr>
<tr>
<td>Null/other variant</td>
<td>113</td>
<td>90</td>
<td>1.0 (0.9–1.8)</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>GSTM3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTM3 V224I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val/Val</td>
<td>565</td>
<td>588</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Val/Reval</td>
<td>472</td>
<td>451</td>
<td>1.0 (0.99–1.3)</td>
<td>0.30</td>
</tr>
<tr>
<td>Val/Re</td>
<td>92</td>
<td>88</td>
<td>1.0 (0.77–1.7)</td>
<td>0.89</td>
</tr>
<tr>
<td>GSTM3*V224I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val/Val</td>
<td>439</td>
<td>464</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Val/Reval</td>
<td>529</td>
<td>504</td>
<td>1.0 (0.99–1.3)</td>
<td>0.19</td>
</tr>
<tr>
<td>Val/Re</td>
<td>160</td>
<td>154</td>
<td>1.0 (0.99–1.4)</td>
<td>0.64</td>
</tr>
</tbody>
</table>

### Table 2: Odds ratios for the associations of polymorphisms in NAT and GST genes and bladder-cancer risk
Table 3: NAT2 slow-acetylation genotype association by smoking characteristic

<table>
<thead>
<tr>
<th>Smoking characteristic</th>
<th>NAT2 rapid/intermediate</th>
<th>NAT2 slow</th>
<th>Odds ratio (95% CI) for joint NAT2 slow genotype association by smoking characteristic</th>
<th>Odds ratio (95% CI) for joint NAT2 slow genotype association</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>66</td>
<td>131</td>
<td>0.9 (0.6–1.3)</td>
<td>10</td>
</tr>
<tr>
<td>Ever</td>
<td>340</td>
<td>362</td>
<td>1.6 (1.3–1.9)</td>
<td>0.9 (0.6–1.3)</td>
</tr>
<tr>
<td>Occasional</td>
<td>15</td>
<td>37</td>
<td>1.4 (0.6–2.9)</td>
<td>2.9 (2.0–4.2)</td>
</tr>
<tr>
<td>Former</td>
<td>161</td>
<td>212</td>
<td>1.7 (1.3–2.2)</td>
<td>1.2 (0.6–2.4)</td>
</tr>
<tr>
<td>Current</td>
<td>163</td>
<td>113</td>
<td>1.4 (1.1–2.4)</td>
<td>2.4 (1.6–3.7)</td>
</tr>
<tr>
<td><strong>Type of tobacco</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>66</td>
<td>131</td>
<td>0.9 (0.6–1.3)</td>
<td>10</td>
</tr>
<tr>
<td>Black</td>
<td>284</td>
<td>272</td>
<td>1.6 (1.3–2.0)</td>
<td>0.9 (0.6–1.3)</td>
</tr>
<tr>
<td>Blond</td>
<td>40</td>
<td>52</td>
<td>1.2 (0.7–2.1)</td>
<td>2.5 (1.4–4.3)</td>
</tr>
<tr>
<td><strong>Smoking intensity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>66</td>
<td>131</td>
<td>0.9 (0.6–1.3)</td>
<td>10</td>
</tr>
<tr>
<td>&lt;10</td>
<td>26</td>
<td>55</td>
<td>1.7 (0.9–3.2)</td>
<td>0.9 (0.5–1.8)</td>
</tr>
<tr>
<td>10–19</td>
<td>67</td>
<td>57</td>
<td>1.2 (0.7–1.9)</td>
<td>1.6 (0.9–3.0)</td>
</tr>
<tr>
<td>20–29</td>
<td>143</td>
<td>108</td>
<td>1.4 (1.0–2.0)</td>
<td>2.3 (1.3–4.1)</td>
</tr>
<tr>
<td>30–39</td>
<td>31</td>
<td>27</td>
<td>1.8 (0.9–3.5)</td>
<td>1.4 (0.6–3.0)</td>
</tr>
<tr>
<td>≥40</td>
<td>54</td>
<td>73</td>
<td>1.7 (1.1–2.8)</td>
<td>1.0 (0.5–2.0)</td>
</tr>
</tbody>
</table>

*For differences between the odds ratio for NAT2 slow-acetylation genotype within strata defined by smoking characteristics compared with never smokers. This test is equivalent to testing whether the observed joint odds ratio for NAT2 slow-acetylation genotype and smoking characteristics differs from the product of the odds ratio for NAT2 slow-acetylation genotype among never smokers and the smoking characteristic among NAT2 rapid/intermediate genotype. Odds ratios are from conventional logistic regression models adjusted for sex, age, and region. The p for interaction for former vs current smokers is 0.44 and for blond vs black tobacco is 0.33. Odds ratios are from conventional logistic regression models adjusted for sex, age, and smoking cessation (former/current). Black is for known or likely black tobacco smokers. §Cigarettes per day. Odds ratios are from conventional logistic regression models adjusted for sex, age, and region. ‡The p for interaction for former vs current smokers is 0.58, and 0.90 respectively for total intensity.

those papers rather than the data from the original reports, with a few exceptions: for Taylor (1998) in Marcus and colleagues’ meta-analysis,41 and for Lin (1994) and Bell (1993) in Engel and colleagues’ meta-analysis42 we used the original report to distinguish between black and white individuals; for Horai (1989) and Karakaya (1986) in Marcus and colleagues’ paper,11 we recalculated odds ratios and 95% CI to obtain exact estimates. For studies not included in previous meta-analyses that did not present crude odds ratios and 95% CI, we calculated them from published data.

Random-effects summary measures were calculated by weighting of each study result by a factor of within-study and between-study variance.43 Homogeneity of study results was assessed by the Q statistic, and publication bias was assessed by Begg’s44 and Egger’s tests.45 A case-only design49 was used in meta-analyses to assess the presence of a multiplicative interaction between NAT2 and GSTM1 genotypes and smoking status (ever/never) because that approach meant we could include some studies without information on the cross-classification of genotype and smoking status among controls, it removed possible biases resulting from the inclusion of hospital controls with diseases related to tobacco use, and it is a powerful design to test for multiplicative interactions under the assumption of independence of NAT2 and GSTM1 from smoking status in the population. Statistical analyses were done with STATA (version 8.2, special edition).

**Role of the funding source**

The study sponsors had no role in the design of the study; in the collection, analysis, or interpretation of the data; or in the writing of the report. The corresponding

Figure 1: Association between smoking intensity (average number of cigarettes per day in categories of 10 cigarettes) and bladder-cancer risk compared with never smokers, stratified by NAT2 acetylation genotype

Odds ratios are from conventional logistic regression models adjusted for sex, age, region, smoking duration (<20 years, 20–29 years, 30–39 years, 40–49 years, ≥50 years), and smoking cessation (current/former smokers). Error bars represent 95% CI. p values for interaction are shown in table 3.
author had full access to all the data in the study and had final responsibility for the decision to submit the paper for publication.

**Results**

The study population was white, predominantly male, and a high proportion were smokers, mostly of black tobacco (table 1). In this population, NAT2 slow-acetylator and GSTM1 null (–/–) genotypes significantly increased the risk of bladder cancer (table 2). The risk of bladder cancer was 40% higher in NAT2 slow acetylators than in NAT2 rapid or intermediate acetylators (odds ratio 1.4 [95% CI 1.2–1.7]); NAT2 rapid acetylators and intermediate acetylators had similar risks of bladder cancer (table 2). The odds ratios for bladder cancer for individuals with deletion of one or two copies of the GSTM1 gene were 1.2 (0.8–1.7) and 1.9 (1.4–2.7), respectively (trend test p<0.0001). Individuals with the null genotype had a 70% higher risk of bladder cancer than those with one or two copies of the GSTM1 gene (table 2). The associations for NAT2 and GSTM1 genotypes were similar irrespective of tumour grade or stage (table 2), and there was no evidence that these associations differed by age or sex (data not shown).

The joint association for the combined NAT2 slow-acetylator and GSTM1 null genotype, present in 28% of the control population, compared with NAT2 rapid/intermediate-acetylator and GSTM1 present genotype (odds ratio 2.2 [1.7–2.9]) was consistent with a weak multiplicative interaction between these two genetic variants; however, the test for multiplicative interaction was not significant (p=0.15). None of the other genetic polymorphisms investigated was significantly associated with an increased risk of bladder cancer (table 2), and there was no evidence of multiplicative interactions between them (data not shown).
Conventional logistic regression analyses showed a significant multiplicative interaction between NAT2 slow acetylation and cigarette smoking status (ever/never, $p=0.008$; table 3) with an interaction odds ratio of 1.8 (1.2–2.8). The evidence for a multiplicative interaction was somewhat weaker (interaction odds ratio 1.4 [1.0–1.9], $p=0.08$) when we used SPMLE logistic regression, which assumed genotype–smoking and genotype–sex independence conditional on age, in the source population. Estimates for the NAT2 slow-acetylation association with bladder cancer were similar for occasional, current, and former smokers (table 3).

The data suggested that the association of NAT2 slow-acetylation genotype with bladder cancer was stronger for known or likely smokers of black tobacco than for smokers of blond tobacco (table 3). However, this difference was not significant (table 3). The NAT2 and smoking intensity interaction is described by showing the odds ratios for NAT2 slow acetylation genotype by smoking intensity (table 3), for the joint association of NAT2 slow genotype and smoking intensity (table 3), and for smoking intensity by NAT2 acetylation genotype (figure 1). NAT2 slow acetylators were at a higher risk from cigarette smoking than rapid or intermediate acetylators, for all smoking intensities (figure 1). The magnitude of the association between NAT2 slow acetylation and bladder-cancer risk among regular smokers was similar across different smoking intensities (table 3), durations, and pack-years (data not shown). As with the interaction between NAT2 and smoking status, SPMLE odds ratios and $p$ values for interactions with other smoking characteristics were slightly attenuated compared with conventional analyses (data not shown).

Neither conventional nor SPMLE logistic regression showed a significant multiplicative interaction (odds ratio 0.7 [0.4–1.1], $p=0.09$, and 0.8 [0.5–1.1], $p=0.15$, respectively) for the association of GSTM1 null and smoking status (ever/never) on bladder-cancer risk. Thus, the relative risk of bladder cancer for GSTM1 null compared with present genotype does not vary by smoking status. No multiplicative interactions were found for other smoking characteristics such as smoking...
cessation (current vs former smokers), smoking intensity, or duration (data not shown). Since an additive interaction can exist in the absence of a multiplicative interaction, and departures from the additive model might have biological implications under certain assumptions, we then tested for an additive interaction. Both conventional and SPMLE logistic regressions showed significant departures from the additive model (ie, additive interactions) or \( \text{GSTM1} \) null genotype and smoking status, with synergy indices of 1·3 (95% CI 1·0–1·6; \( p=0·04 \)) and 1·4 (1·1–1·7; \( p=0·001 \)), respectively.

We updated a previously published meta-analysis of 22 studies of \( \text{NAT2} \) and bladder cancer to include data from our study and eight additional studies, including a total of 5091 cases and 6501 controls (figure 2). The summary relative risk for \( \text{NAT2} \) slow acetylators compared with rapid/intermediate acetylators was 1·4 (1·2–1·6; \( p<0·0001 \)) with no evidence for publication bias according to Begg’s (\( p=0·94 \)) or Egger’s tests (\( p=0·91 \)). There was some evidence of study heterogeneity (\( Q \) test \( p<0·04 \), which was not present when 15 studies with fewer than 100 cases each were excluded (summary odds ratio 1·4 [1·2–1·5]; \( Q \) test \( p=0·31 \)). Summary estimates for white populations (56% prevalence of \( \text{NAT2} \) slow acetylators in controls) and Asian populations (11% prevalence of \( \text{NAT2} \) slow acetylators in controls) were similar (\( p=0·87 \); figure 2). The summary relative risk for studies of white populations in the USA was lower than that for studies done in Europe, which accounted for most (82%) white cases; however this difference was not significant (\( p=0·18 \); figure 2).

We also updated a case-only meta-analysis of \( \text{NAT2} \) and smoking interaction on bladder-cancer risk to include results from our study and five additional studies published after the meta-analysis (figure 3). This analysis included a total of 4306 cases and showed

### Study Year Country Cases

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Country</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rankes</td>
<td>Unpublished</td>
<td>USA</td>
<td>13</td>
</tr>
<tr>
<td>Bell (blacks)</td>
<td>1993</td>
<td>USA</td>
<td>16</td>
</tr>
<tr>
<td>Heckbert</td>
<td>1992</td>
<td>USA</td>
<td>29</td>
</tr>
<tr>
<td>Daly</td>
<td>1993</td>
<td>UK</td>
<td>53</td>
</tr>
<tr>
<td>Mungan</td>
<td>2000</td>
<td>Netherlands</td>
<td>61</td>
</tr>
<tr>
<td>Salagić</td>
<td>1999</td>
<td>Slovakia</td>
<td>76</td>
</tr>
<tr>
<td>Georgiou</td>
<td>2000</td>
<td>Greece</td>
<td>89</td>
</tr>
<tr>
<td>Lin (whites)</td>
<td>1994</td>
<td>USA</td>
<td>89</td>
</tr>
<tr>
<td>Chen</td>
<td>1994</td>
<td>UK</td>
<td>95</td>
</tr>
<tr>
<td>Zhong</td>
<td>1993</td>
<td>UK</td>
<td>97</td>
</tr>
<tr>
<td>Aktaş</td>
<td>2001</td>
<td>Turkey</td>
<td>103</td>
</tr>
<tr>
<td>Srivastava</td>
<td>2004</td>
<td>India</td>
<td>106</td>
</tr>
<tr>
<td>Moore</td>
<td>2004</td>
<td>Argentina</td>
<td>114</td>
</tr>
<tr>
<td>Torunier</td>
<td>2001</td>
<td>Turkey</td>
<td>121</td>
</tr>
<tr>
<td>Jeong</td>
<td>2003</td>
<td>Korea</td>
<td>126</td>
</tr>
<tr>
<td>Peluso</td>
<td>2000</td>
<td>Italy</td>
<td>130</td>
</tr>
<tr>
<td>Steinhoff</td>
<td>2000</td>
<td>Germany</td>
<td>135</td>
</tr>
<tr>
<td>Schrakenberger</td>
<td>2000</td>
<td>Germany</td>
<td>157</td>
</tr>
<tr>
<td>Kang</td>
<td>1999</td>
<td>Korea</td>
<td>174</td>
</tr>
<tr>
<td>Hung</td>
<td>2004</td>
<td>Italy</td>
<td>201</td>
</tr>
<tr>
<td>Bell (whites)</td>
<td>1993</td>
<td>USA</td>
<td>213</td>
</tr>
<tr>
<td>Kim</td>
<td>2002</td>
<td>Korea</td>
<td>216</td>
</tr>
<tr>
<td>Lee</td>
<td>2002</td>
<td>Korea</td>
<td>232</td>
</tr>
<tr>
<td>Okieh</td>
<td>1996</td>
<td>Denmark</td>
<td>234</td>
</tr>
<tr>
<td>Tsukino</td>
<td>2004</td>
<td>Japan</td>
<td>325</td>
</tr>
<tr>
<td>Kanagas</td>
<td>2005</td>
<td>USA</td>
<td>354</td>
</tr>
<tr>
<td>Brockmoller</td>
<td>1996</td>
<td>Germany</td>
<td>374</td>
</tr>
<tr>
<td>García-Closas</td>
<td>Current</td>
<td>Spain</td>
<td>1139</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cases</th>
<th>Odds ratio (95% CI)</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies (n=28)</td>
<td>5072</td>
<td>1·5 (1·3–1·6)</td>
</tr>
<tr>
<td>Studies of predominantly white populations (n=18)</td>
<td>3539</td>
<td>1·4 (1·2–1·6)</td>
</tr>
<tr>
<td>Europe (n=13)</td>
<td>2841</td>
<td>1·5 (1·2–1·7)</td>
</tr>
<tr>
<td>USA (n=5)</td>
<td>698</td>
<td>1·3 (1·1–1·5)</td>
</tr>
<tr>
<td>Studies of Asian populations (n=6)</td>
<td>1073</td>
<td>1·4 (1·2–1·7)</td>
</tr>
</tbody>
</table>

Figure 4: Meta-analysis of studies of \( \text{GSTM1} \) null genotype and bladder-cancer risk

Number of cases for studies in Engel et al are based on table 1 of that paper.
evidence for an interaction with a summary estimate of 1·2 (1·1–1·5; p=0·009) for all populations combined. There was no evidence of overall study heterogeneity (Q test p=0·84) or publication bias (Begg’s test p=0·40, Egger’s test p=0·13). The point estimate for interaction was higher in white than in Asian populations (1·3 vs 0·9) and in European than in US white populations (1·4 vs 1·0); however, these differences were not significant (p=0·32 and 0·09, respectively; figure 3).

A meta-analysis of 17 studies of GSTM1\(^8\) was also updated to include our study, ten additional studies,\(^{17,21,22,29,30,35,52}\) and an update from a previously published study,\(^9\) yielding a total of 5072 cases and 6466 controls (figure 4). The summary odds ratio for GSTM1 null versus present genotype for all populations combined was 1·5 (1·3–1·6; p<0·0001) with no evidence of study heterogeneity (Q test p=0·10) or publication bias by Begg’s test (p=0·27) or Egger’s test (p=0·57). Summary estimates were similar and significant in white populations (51% of GSTM1 null genotype in controls) and Asian populations (53% of GSTM1 null genotype in controls), as well as in US and European white populations (figure 4).

An updated case-only meta-analysis of studies that investigated the GSTM1-smoking interaction\(^8\) to include our study and seven other studies\(^{17,21,22,29,30,35,52}\) (17 studies of 4043 cases) confirmed the absence of a multiplicative interaction with a summary odds ratio of 1·0 (0·9–1·2; p=0·86; figure 5). The Q test showed no evidence of study heterogeneity (p=0·87), and Begg’s test (p=0·15) and Egger’s test (p=0·03) suggested the presence of publication bias. Summary estimates for the interaction were very similar for all population subgroups (figure 5).

Discussion

This report provides compelling evidence of an increased bladder-cancer risk associated with the GSTM1 null and NAT2 slow-acetylation genotypes. The latter association was particularly important among cigarette smokers. Although the relative risks for polymorphisms in NAT2 and GSTM1 genes are modest, these polymorphisms could account for a large proportion of bladder cancers because they are very common in the population. From our data, we estimate that these polymorphisms cause 31% (95% CI 20–46) of bladder cancers in white populations. In addition, we provide strong evidence against a substantial overall association for polymorphisms in other NAT and GST genes, with the possible exception of small to moderate associations for the NAT1 *10/*10 and GSTP1 114Val/Val genotypes.
The new meta-analysis of studies of NAT2 slow acetylation and bladder-cancer risk showed that this association is robust and similar for white and Asian populations. The lack of significance for the association in Asian populations might be explained by substantially lower statistical power to detect associations in Asian studies owing to a lower prevalence of NAT2 slow acetylators (11% for Asian vs 56% for white populations), along with a smaller number of cases available for the meta-analysis. We also found that NAT2 slow acetylators are especially susceptible to the adverse effects of cigarette smoking on bladder-cancer risk. This gene–environment interaction has strong biological plausibility, because NAT2 slow acetylators have decreased capacity to detoxify aromatic monoamines by N-acetylation, tobacco smoking is a primary source of exposure to aromatic amines in the general population, and aromatic amines are suspected to be the primary bladder carcinogen in tobacco smoke. Our data suggest that NAT2 slow acetylation does not increase bladder-cancer risk among never smokers, although they do not rule out a small increase in risk in this group.

Because the content of aromatic amines is higher in black than in blond tobacco, the effect of NAT2 slow acetylation could conceivably be stronger for smokers of black tobacco. Our data are consistent with this hypothesis, although the differences were not significant. The magnitude of the association between NAT2 slow acetylation and bladder-cancer risk was similar for different smoking intensities in our study population. Our meta-analysis of the interaction between smoking status and NAT2 slow-acetylation genotype suggested a stronger interaction with ever/never smoking in European than in US studies (p=0.09). This difference could result from the lower content of aromatic amines in blond tobacco, which is generally smoked in the USA, than in the black tobacco commonly smoked in parts of Europe. This explanation is consistent with a study of a population in the USA that found an interaction between NAT2 slow-acetylation genotype and smoking only for heavy smokers.13

Distinction of individuals with one and two copies of the GSTM1 gene, an issue that has not been adequately addressed in previous studies of bladder cancer, suggests the presence of a gene-dosage effect with relative risks of 1.2 (0.8–1.7) and 1.9 (1.4–2.7) for individuals with one or no copies of GSTM1, respectively, compared with those with two copies (p for trend <0.0001). Meta-analyses of the association between the deletion of two copies of the GSTM1 gene (null genotype) compared with the presence of one or two copies (present genotype), as calculated from previous studies that could not distinguish between these two groups of individuals, showed that this association is robust (p<0.0001) and similar in magnitude and significant across different population subgroups.

The relative risk for GSTM1 null genotype and bladder cancer was similar for smokers and never smokers in our study population and in meta-analysis within population subgroups. This finding suggests the presence of an additive interaction, which is supported by our data (p=0.04). This observation is compatible with equal protection by GSTM1 activity against tobacco-related and non-tobacco-related bladder cancers. This finding suggests that GSTM1 lowers the risk of bladder cancer through mechanisms that are not specific to the detoxification of polycyclic aromatic hydrocarbons in tobacco smoke. Other mechanisms of action for GSTM1 could be protection from oxidative damage through metabolism of reactive oxygen species.25 Our data did not confirm previously suggested differences in risk for NAT2 slow-acetylation and GSTM1 null genotypes by tumour grade or stage at presentation.26–29 Our findings are consistent with a potential interaction between NAT2 slow-acetylation and GSTM1 null genotypes; however, further evidence is needed to confirm this interaction.

Associations between bladder-cancer risk and polymorphisms in genes encoding the NAT1 enzyme involved in the activation of aromatic amines by O-acetylation, and other GST enzymes that have important roles in the detoxification of polycyclic aromatic hydrocarbons and other carcinogens have been less fully explored. Previous studies have provided inconsistent evidence for an association between bladder-cancer risk and NAT1*10 alone or in combination with NAT2 slow acetylation,14–15,30 GSTT1 null alone or in combination with GSTM1 null genotype,17,20–34 and GSTP1 105 Val/Val genotype.17,21,31,32 The data from our study do not support a substantial association between GSTT1 and GSTM3 genotypes and bladder-cancer risk. We found no significant increases in bladder-cancer risk associated with polymorphisms in NAT1 or GSTP1 genes; however, our estimates did not exclude a small to moderate association for the NAT1*10/*10 genotype compared with the NAT1*4/*4 genotype or for genotypes with the GSTP1 114Val allele compared with the 114Ala/Ala genotype.

Analyses by conventional logistic regression suggested a modification of the association between risk of bladder cancer and NAT2, GSTM1, and NAT1 genotypes by sex. However, the modifications by sex were explained by unexpected differences in the genotype distribution for male and female controls.

Our study had several strengths: high participation rates, large sample size, and high-quality information on exposure and genotype. Specifically, we made an effort to improve the precision in genotype estimation by genotyping the seven single-nucleotide polymorphisms in NAT2 that probably account for virtually all genetic variation in white populations, and we developed assays that successfully distinguished individuals with one or two copies of the GSTM1 and GSTT1 genes. We also used the SPMLE method to increase power and reduce bias in the estimation of interactions, because of the strong
evidence from previous studies for independence of NAT2 and GSTM1 genotypes from cigarette smoking status1,11,36 and sex43 in the general population. To limit selection bias, we carefully selected controls from patients admitted for various diagnoses that were thought to be unrelated to exposures of interest, including tobacco use. Genotype frequencies among the control population were similar to those previously reported. We found no significant overall differences in genotype frequencies across control diagnoses that could have biased our results.

Although this study is the largest to date on the role of genetic polymorphisms and bladder-cancer risk and had adequate statistical power to detect modest genotype associations, the power to detect interactions was limited. Meta-analyses including previous studies improved our ability to make inferences on interactions, when there was an adequate number of previous studies with homogeneous results. A consortium of bladder-cancer studies has been formed to facilitate the pooling of comparable data on environmental and genetic risk factors across studies that will help overcome the limited power of individual studies to investigate complex inter-relations.

Contributors
M García-Closas, N Malats, D Silverman, M Dosemeci, M Kogevinas, F X Real, and N Rothman participated in the study design, enrolment of patients, and gene selection. G Caño-Vinyals, M Torá, F Fernández, C Sasmanic, A Tardón, C Serra, A Carrato, and R García-Closas participated in the study design and enrolment of patients. D W Hein, M Yeager, R Welch, and S Chianock participated in gene selection and genotyping. J Lloreta participated in the pathology review. N Chatterjee and S Wacholder participated in the statistical analyses. M García-Closas did the statistical analyses and drafted the paper with input from all investigators.

Participating study centres in Spain
Institut Municipal d’Investigació Mèdica, Universitat Pompeu Fabra, Barcelona—Coordinating Center (M Kogevinas, N Malats, F X Real, M Sala, G Caño-Vinyals, M Torá, F Fernández, C Sasmanic, A Tardón, C Serra, A Carrato, and R García-Closas) participated in the study design and enrolment of patients. D W Hein, M Yeager, R Welch, and S Chianock participated in gene selection and genotyping. J Lloreta participated in the pathology review. N Chatterjee and S Wacholder participated in the statistical analyses. M García-Closas did the statistical analyses and drafted the paper with input from all investigators.

Acknowledgments
We thank Robert C Saal (Westat, Rockville, MD, USA), Leslie Carroll, and Jane Wang (both IMS, Silver Spring, MD, USA) for their support in study and data management; Maria Sala (Institut Municipal d’Investigació Mèdica, Barcelona, Spain) for her work in data collection; physicians, nurses, interviewers, and study participants for their efforts during fieldwork; Pam Marcus and Larry Engel from the National Cancer Institute for providing datasets for meta-analyses used in their previous publications; and the Genetic Susceptibility to Environmental Carcinogens Study (http://www.gsec.net/) for bringing together the collaborative network of investigators that contributed data used in Engel and colleagues’ study,4 which enabled the update of the meta-analysis on GSTM1, smoking, and bladder cancer. This work was supported by National Cancer Institute Westat contract number N02-CP-11015, FIS/Spain 00/0745 and G03/174, and CA34627.

References
Universal newborn screening for permanent childhood hearing impairment: an 8-year follow-up of a controlled trial

Colin Kennedy, Donna McCann, Michael J Campbell, Lindsay Kimm, Roger Thornton

An 8-year follow-up study of the birth cohort of babies enrolled in the Wessex controlled trial of universal newborn screening (UNS) for permanent childhood hearing impairment (PCHI) was undertaken to establish whether UNS would increase the proportion of all true cases of PCHI in children aged 7–9 years who are referred early. The proportion referred before 6 months of age increased from 11 of 35 (31%) children with true PCHI born during periods without UNS to 23 of 31 (74%) born during periods with UNS (difference 43%, 95% CI 19–60). UNS leads to early referral of PCHI.

Bilateral permanent childhood hearing impairment (PCHI) of 40 decibels (dB) or more hearing level that is congenital, or not known to be acquired, affects 112 per 100 000 population1 and less than half of cases occur in congenital, or not known to be acquired, affects 112 per 100 000 population1 and less than half of cases occur in children at high risk. Although the benefit of universal newborn screening (UNS) has been disputed, there is preliminary evidence that enrolment in an intervention programme before the age of 8 months could reduce the resulting deficit of verbal relative to non-verbal intelligence quotient by up to 19 points.2,4 This finding provides the rationale for UNS for PCHI, recommended by the National Institutes of Health since 1993.

UNS had been introduced at the time of the trial reported here in only two districts in the UK and in 11 hospitals in the USA. Use of the screening programme had increased one hundred-fold by the year 2000,6 but a 2001 systematic review4 identified our programme before the age of 8 months could reduce the resulting deficit of verbal relative to non-verbal intelligence quotient by up to 19 points.2,4 This finding provides the rationale for UNS for PCHI, recommended by the National Institutes of Health since 1993.

UNS had been introduced at the time of the trial reported here in only two districts in the UK and in 11 hospitals in the USA. Use of the screening programme had increased one hundred-fold by the year 2000,6 but a 2001 systematic review4 identified our preliminary report of the present study6 as the only published controlled trial of UNS. There is no published estimate of the effect of UNS on the proportion of all true cases of PCHI >40 dB hearing level referred early, which is a key indicator of performance. The derivation of such an estimate needs knowledge of the total number of true cases and this in turn requires at least 6 years of follow-up of a screened birth cohort so that false negatives with screening and cases of progressive PCHI can be ascertained.1 We have now completed 8 years of follow-up of the birth cohort that was enrolled in the Wessex controlled trial of UNS and we report here on the rise in sensitivity to all true cases of PCHI associated with the introduction of UNS.

All infants born in the four participating hospitals between Oct 1, 1993, and Oct 31, 1996, were included. Details of the ethics committee approval, informed consent, and two-stage UNS programme have been outlined previously.6,7 Two teams of screeners moved back and forth between the two pairs of hospitals every 4–6 months thus creating eight consecutive birth cohorts of infants that were similar, except that alternate cohorts at each hospital had been the target of a programme of UNS. Screening of all infants for hearing impairment at 7–8 months of age with the distraction test continued throughout the trial.

Over a 10-year period from October 1993, we undertook a series of follow-ups and received regular updates from the four local audiology services. We obtained, by review of the case records and outpatient lists, details of the detection and management of all cases of PCHI identified in the 1993–96 trial birth cohort. Information about these cases was also obtained in this way from other paediatric audiologists and audiology scientists, current medical family practitioners, teachers, specialist teachers for the hearing impaired, speech and language therapists, and other professionals involved. Children with an identified postnatal cause (eg, bacterial meningitis) were excluded.

There were 53 781 livebirths in the four participating hospitals in the 3 years of the trial. 25 609 were born during periods of UNS, of whom 21 279 (83%) received screening and 392 (2%) screened positive.6 In September 2003, 66 children with bilateral PCHI >40 dB hearing level were identified from the trial birth cohort (table 1). Seven children with an early diagnosis were reported as having progressed in severity since the time of their first assessment.

<table>
<thead>
<tr>
<th>Risk status</th>
<th>Period with UNS (n=25 609)</th>
<th>Period without UNS (n=28 172)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moderate</td>
<td>Severe</td>
</tr>
<tr>
<td>Low risk</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>High risk</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Ward</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Low risk</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Source of referral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neonatal screen</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>HVDT screen</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Age at referral (months)</td>
<td>6 or more</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 1: Numbers of cases of bilateral PCHI >40 dB hearing level at age 7–9 years in children born during periods with and without universal newborn screening, by risk status, source of referral, and age at referral.
Of 31 cases of PCHI in September 2003, among children born during periods with UNS, a positive screening with UNS was confirmed in 22 (71%); seven had not been screened, including four because consent had not been obtained; and two had screened negative with UNS. No other false negatives from UNS have come to light. Of 35 children with PCHI in September 2003, who were born during periods without UNS, 12 (34%) were referred after failing the distraction test screen (table 1) and six (17%) were false negatives who had passed the distraction test screen.

Referrals before the age of 6 months in the birth cohorts born during periods with and without UNS constituted 74% and 31%, respectively, of all true cases of PCHI in children aged 7–9 years (difference 43%, 95% CI 19–60, p=0·001; table 1). Adjustment for the effect of severity of hearing impairment on age of referral in a logistic regression model slightly strengthened the association between UNS and the proportion of early referrals, increasing the odds ratio from 6·3 to 6·9 (95% CI 2·2–22·0, p=0·001). The percentage of all true cases referred was greater at any given age during the first 3 years of life for children born during periods with UNS than for those born during periods without UNS, but was closely similar thereafter (figure). Ages of children at referral differed significantly for those born in periods with UNS (median 0 months) and without UNS (8 months) when analysed with Breslow’s test (n=66, \( \chi^2 = 17, p < 0·001 \)).

Test sensitivity and specificity were 22/24 (0·92) and 20,960/21,279 (0·98) and the positive and negative likelihood ratios (the odds of disease after screen divided by the odds of disease before screen) were 61 and 0·08, respectively (table 2). One additional case of bilateral hearing impairment was referred before the age of 6 months during periods of UNS for every 1969 (95% CI 1011–12,896) babies in the target population for screening.

Our results suggest that this two-stage method of UNS was effective in increasing the percentage of all cases of bilateral PCHI \( \geq 40 \) dB hearing level in children aged 7–9 years who were referred for hearing assessment before the age of 6 months from 31% to 74%. This finding lends support to our earlier report\(^6\) and strengthens the conclusions through follow-up of the children to an age when almost all true cases, including progressive losses, are likely to have been identified in the UK health system.\(^1\) The number of true cases of PCHI in children who were born during periods of UNS but who did not receive screening because consent was not obtained will probably decrease as UNS becomes established as a clinical service and this will further boost early referral rates.

By contrast with the substantial rise in rates of early referral, data for ages of confirmation of PCHI (data not shown) lend support to our preliminary report\(^6\) of no significant overall increase in the proportion of the study population who were confirmed and managed by the age of 10 months during periods with UNS and indicate that 15 of 31 (48%) cases born during periods with UNS were not managed until after the age of 18 months. However, the advent in the late 90s of management systems capable of translating early referral from UNS into early management has been noted in other reports.\(^5,8\) Therefore, the small increase in rate of early management associated with UNS in our study was probably a problem specific to the early 90s when the screening programme was new.

Eight children with PCHI had screened negative (including two born in periods of UNS) and seven (including two born in periods of UNS) had documented progression in severity after detection in infancy. The sum of these two figures yields 15 of 66, or 23% of cases that might have been progressive on the assumption that the negative screen in infancy had been a true reflection of the hearing status at that time in all eight children. Reliable estimates of the contribution of progressive losses to the burden of PCHI and documentation of the age of confirmation and management both require longitudinal study, which remains important for research and audit of UNS in future birth cohorts.
We estimated the possible contribution of cases missed because field reports were incomplete, as follows: 40 of 66 (61%) of all cases of PCHI identified in this study were confirmed before the age of 18 months (data not shown) and the mean rate of emigration of the general population of 0–15-year-olds out of the four study districts was 1.4% per annum, thus over 95% of true cases in this birth cohort have probably been identified in our study. Furthermore, the similar figures for prevalence of 121 and 124 cases per 100 000 population born during respective periods with UNS and without UNS suggests that any under-ascertainment did not bias the estimate of the size of the effect of UNS.

The US Preventive Services Task Force has stated that better evidence about the effectiveness of UNS is needed and could be obtained via population-based studies that begin with inception cohorts and carefully report outcomes in all possible patients, as well as rates of loss to follow-up. Our report fits this description and is the strongest available evidence of the added benefit of UNS in the early detection of PCHI. Assessment of the effect of early intervention on the speech and language of the children and the costs incurred by their families and by health providers is in progress and will be the subject of future reports. Policymakers should consider these data, as well as those presented here, when contemplating the introduction of UNS.

Contributors
The design and application for funding of this study was developed from an idea by C R Kennedy with help from M J Campbell; A D Thornton, I Kimm, and D C McCann oversaw the conduct of the earlier and later stages of the study, respectively, with help from all the other authors. M J Campbell, assisted by C R Kennedy and D C McCann, undertook statistical analysis. All authors contributed to the preparation of the manuscript.

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

Acknowledgments
We acknowledge the support of The Wellcome Trust and Defeating Deafness (The Hearing Research Trust) and the assistance of the Wessex Universal Newborn Screening Trial Group. We thank all the babies enrolled in the trial and their parents; our research assistants Kristen Paul, Helen Ryder, Helen Davis, Eleanor Lutman, and Shirley Golden; the project speech and language therapist, Sarah Worsfold; the participating audiology, speech, and language educational services for hearing impaired children; the hospitals (Princess Anne, Southampton; St Mary’s, Portsmouth; Royal United, Bath; and Princess Margaret, Swindon) for accommodating the neonatal testing teams; and the MRC Institute of Hearing Research for equipment calibration and other help. The funding source 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.

References

Obesity, cigarette smoking, and telomere length in women

A M Valdes, T Andrew, J P Gardner, M Kimura, E Oelsner, L F Cherkas, A Aviv, T D Spector

Lancet 2005; 366: 662–64

Published online: June 24, 2005
DOI 10.1016/S0140-6736(05)66630-5

Twin Research and Genetic Epidemiology Unit, St Thomas’ Hospital, London, UK
(A M Valdes PhD, T Andrew PhD, E Oelsner BSc, L F Cherkas PhD, Prof T D Spector MD); and Hypertension Research Center, University of Medicine and Dentistry of New Jersey, Newark, New Jersey, USA (J P Gardner PhD, M Kimura MD, Prof A Aviv MD)

Correspondence to: Professor T D Spector
tim.spector@kcl.ac.uk

Obesity and smoking are important risk factors for many age-related diseases. Both are states of heightened oxidative stress, which increases the rate of telomere erosion per replication, and inflammation, which enhances white blood cell turnover. Together, these processes might accelerate telomere erosion with age. We therefore tested the hypothesis that increased body mass and smoking are associated with shortened telomere length in white blood cells. We investigated 1122 white women aged 18–76 years and found that telomere length decreased steadily with age at a mean rate of 27 bp per year. Telomeres of obese women were 240 bp shorter than those of lean women (p=0.026). A dose-dependent relation with smoking was recorded (p=0.017), and each pack-year smoked was equivalent to an additional 5 bp of telomere length lost (13%) compared with the rate in the overall cohort. Our results emphasise the pro-ageing effects of obesity and cigarette smoking.

Telomeres cap the ends of chromosomes and protect them from degradation and end-to-end fusion. Telomeres of cultured somatic cells undergo erosion with each cycle of replication, and oxidative stress enhances this process. Both obesity and cigarette smoking are important risk factors in many age-related diseases, and are associated with increased oxidative stress and inflammation. The latter process is marked by increased white blood cell (WBC) turnover. Telomere attrition (expressed in WBCs)
can serve as a marker of the cumulative oxidative stress and inflammation and, consequently, show the pace of biological ageing. We therefore expected obese individuals and smokers to have shortened telomeres. To investigate this hypothesis we studied WBC telomere length in 1122 healthy white women aged 18–72 years, examining the relations with both smoking and obesity-related phenotypes.

Participants were female twins (45 monozygotic and 516 dizygotic pairs) from the TwinsUK Adult Twin Registry, a group previously developed to study the heritability and genetics of diseases with a higher prevalence among women. These women were recruited from the general population through national media campaigns in the UK, and were similar to age-matched population singletons in terms of disease-related and lifestyle characteristics (www.twinsuk.ac.uk). In our cohort, body-mass index (BMI) was >30 in 119 (11%) women and <20 in 85 (8%). None of the participants had clinical diabetes. 531 (47%) women had never smoked, 369 (33%) were ex-smokers, 203 (18%) were still smoking, and smoking status was unknown for 19 (2%). Smoking history was recorded with a standardised questionnaire. Smoking exposure was measured as pack-years=number of cigarette packs smoked per day×number of years smoking.

A venous blood sample was taken after an overnight fast. We extracted DNA from WBCs, and measured the concentration of leptin in serum with a radioimmunoassay (Linco, St Charles, MO, USA). We measured the mean of the terminal telomere restriction fragment (TRF) lengths, an index of telomere length, with the Southern blot method. Written and oral informed consent was obtained from all participants. The St Thomas’ Hospital Research Ethics Committee approved the study.

Standard linear regression techniques were used to correlate the TRF length with age and the age-adjusted TRF with individual factors. Log-transformed leptin values were used for both the age-adjusted and unadjusted linear regressions. The associations between categorical variables and telomere length, adjusting for age or other covariates, were assessed using analyses of variance. To adjust for non-independence between twins in a pair, bootstrap sets were generated by selecting a random twin from each pair using analysis of variance, and the p value of the mean test statistic from 100 replicates was used to confirm statistical significance. S-Plus 6.0 (Insightful Corp) software was used. No significant difference in the variables studied was noted between monozygotic and dizygotic twins.

Telomere length decreased steadily with age at a mean rate of 27 bp per year (SD 50·2; figure A) and a highly significant negative correlation was detected (table). The proportion of the variance in telomere length accounted for by age was 20·6%. Squared and cubed age terms were also added to the model and had no significant

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>Correlation with TRF</th>
<th>p</th>
<th>Age-adjusted correlation with TRF</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRF (kb)</td>
<td>7·06 (0·67)</td>
<td>-0·455</td>
<td>&lt;0·0001</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Age (years)</td>
<td>47·77 (12·11)</td>
<td>-0·126</td>
<td>&lt;0·0001</td>
<td>-0·077</td>
<td>0·031</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25·05 (4·69)</td>
<td>-0·124</td>
<td>&lt;0·0001</td>
<td>-0·088</td>
<td>0·019</td>
</tr>
<tr>
<td>Serum leptin (ng/mL)</td>
<td>16·26 (12·50)</td>
<td>-0·031</td>
<td>ns</td>
<td>-0·087</td>
<td>0·017</td>
</tr>
<tr>
<td>Smoking status†</td>
<td>n/a</td>
<td>-0·214</td>
<td>&lt;0·0001</td>
<td>-0·110</td>
<td>0·045</td>
</tr>
<tr>
<td>Cigarette pack-years‡</td>
<td>3·56 (3·56)</td>
<td>-0·214</td>
<td>&lt;0·0001</td>
<td>-0·110</td>
<td>0·045</td>
</tr>
</tbody>
</table>

ns=not significant. *Statistical significance of regression coefficient from 100 bootstrap replicates. †Coded as 0=never smoked, 1=ex-smokers, 2=current smokers. ‡Among ex-smokers and current smokers only.

Table 1: Descriptive statistics of study subjects and correlations with telomere terminal restriction fragment (TRF) length before and after adjusting for chronological age.
effect on telomere length (p=0.92 and p=0.98, respectively) suggesting a linear relation between TRF and age.

Additionally, we noted no clear age-related difference in rates of TRF loss; average rate of loss was 27.7 bp per year in women aged 50 years and over and 25.7 bp per year in those younger than 50 years. BMI, leptin concentration in serum, and smoking status were all significantly correlated with age (r=0.12, r=0.13, r=0.10, respectively). Leptin concentration in serum and BMI were strongly positively correlated (r=0.76). However, the correlations between smoking status and BMI (r=0.05) and between leptin concentration in serum and smoking status (r=0.06) were not statistically significant. BMI, leptin concentration in serum, and packs-year of cigarettes smoked were negatively correlated with telomere length. The regression coefficients of these variables remained statistically significant after adjustment for age (table).

In addition to the linear models tested on continuous measures, lean individuals were found to have significantly longer telomeres than women with mid-range BMIs, who, in turn, had longer telomeres than obese individuals (figure, B; p=0.026).

Age-adjusted telomere length was negatively correlated with log-transformed leptin concentration in serum (table) and the mean age-adjusted telomere length showed a progressive decrease through the quartiles of leptin concentration (figure, C).

Individuals who had never smoked had longer age-adjusted telomeres than former smokers and both had longer telomeres than current smokers (figure, D; p=0.02). Moreover, age-adjusted telomere length decreased with the amount of cigarettes smoked (table; figure, D). Each pack-year smoked was equivalent to a loss of an additional 5 bp, or 18% of the average annual per year in those younger than 50 years.

Our findings suggest that obesity and cigarette smoking accelerate human ageing. Our cross-sectional data underscore the considerable variation in telomere length between individuals. Thus large cohorts are needed to capture the effects of inflammation and oxidative stress. However, in view of the hypothesis that telomere length in vivo represents cellular turnover and exposure to oxidative and inflammatory damage, the difference in telomere length between being lean and being obese corresponds to 8-8 years of ageing; smoking (previous or current) corresponds on average to 4-6 years of ageing; and smoking a pack per day for 40 years corresponds to 7-4 years of ageing. Our results emphasise the potential wide-ranging effects of the two most important preventable exposures in developed countries—cigarettes and obesity.

Contributors
A M Valdes participated in the statistical analysis and in the preparation of the manuscript. T Andrew participated in the processing and statistical analysis of the data. E Oelner and L F Cherkas collected and verified the clinical information of the study participants. J Gardner and M Kimura did the telomere assays and participated in the processing of the data. T D Spector and A Aviv designed and coordinated the study and participated in the preparation of the manuscript.

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

Acknowledgments
We thank all the twins participating in TwinsUK. R Swaminathan of Chemical Pathology for the leptin assays, and Gabriela L Surdulescu for the DNA preparation. This work was supported by the Wellcome Trust Functional Genomics Initiative for the TwinsUK program and WT project grant no 07495/t/Z/04/Z, the Healthcare Foundation of New Jersey and NIH grant AG021593. 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.

References
Phaeochromocytomas are rare neuroendocrine tumours with a highly variable clinical presentation but most commonly presenting with episodes of headaches, sweating, palpitations, and hypertension. The serious and potentially lethal cardiovascular complications of these tumours are due to the potent effects of secreted catecholamines. Biochemical testing for phaeochromocytoma is indicated not only in symptomatic patients, but also in patients with adrenal incidentalomas or identified genetic predispositions (eg, multiple endocrine neoplasia type 2, von Hippel-Lindau syndrome, neurofibromatosis type 1, and mutations of the succinate dehydrogenase genes). Imaging techniques such as CT or MRI and functional ligands such as 123I-MIBG are used to localise biochemically proven tumours. After the use of appropriate preoperative treatment to block the effects of secreted catecholamines, laparoscopic tumour removal is the preferred procedure. If removal of phaeochromocytoma is timely, prognosis is excellent. However, prognosis is poor in patients with metastases, which especially occur in patients with large, extra-adrenal tumours.

Phaeochromocytomas are catecholamine-producing neuroendocrine tumours arising from chromaffin cells of the adrenal medulla or extra-adrenal paraganglia. Tumours from extra-adrenal chromaffin tissue are referred to as extra-adrenal phaeochromocytomas or paragangliomas. The term paraganglioma is also used for tumours derived from parasympathetic tissue in the head and neck, most of which do not produce catecholamines. Nearly 80–85% of phaeochromocytomas arise from the adrenal medulla, whereas about 15–20% are from extra-adrenal chromaffin tissue.12 Catecholamine-producing extra-adrenal paragangliomas are usually found in the abdomen.14

In general outpatient clinics, the prevalence of phaeochromocytoma in patients with hypertension is 0.1–0.6%.5,7 Although these tumours are frequently searched for, they are rarely found. The relatively high prevalence of phaeochromocytoma in autopsies studies (about 0.05%) also indicates that many tumours are missed, resulting in premature mortality.5–8

Hereditary phaeochromocytomas occur in multiple endocrine neoplasia type 2, von Hippel-Lindau syndrome, neurofibromatosis type 1, and the familial paragangliomas.11–14 Sporadic forms of phaeochromocytoma are usually diagnosed in individuals aged 40–50 years, whereas hereditary forms are diagnosed earlier, most often before age 40 years.11–17 Phaeochromocytoma is rare in children, but when found it is often extra-adrenal, multifocal, and associated with hereditary syndromes.11,18,19

Clinical presentation of phaeochromocytoma can vary greatly, with similar signs and symptoms produced by many other clinical conditions (panel 1). Phaeochromocytoma is therefore often referred to as the great mimic. Most but not all the clinical signs and symptoms of phaeochromocytoma are due to the direct actions of secreted catecholamines. Hypertension, tachycardia, pallor, headache, and feelings of panic or anxiety, usually dominate the clinical presentation (table 1).15,17,20,21 Metabolic effects include hyperglycaemia, lactic acidosis, and weight loss.22 Less common signs and symptoms are nausea, fever, and flushing. Hypertension is often paroxysmal in nature, in some patients occurring on a background of sustained hypertension, whereas others can have normal blood pressure between paroxysms. Hypertensive episodes can be severe and result in hypertensive emergencies. Blood pressure can also be consistently normal, especially in patients with adrenal incidentalomas, in those screened for an identified familial syndrome, or in those with a very small tumour.23 Because of increasing use of advanced imaging techniques and improved recognition of genetic causes of phaeochromocytoma, where routine screening is becoming mandatory, the numbers of normotensive and asymptomatic patients diagnosed with the disease have steadily risen.21–26 About 5% of all incidentalomas are phaeochromocytomas, with about 25% of all phaeochromocytomas now being discovered incidentally during imaging studies for unrelated disorders.22,26,27

Normal blood pressure or even hypotension is also common in patients with dopamine-producing paragangliomas, in whom diagnosis is often based on the space-occupying complications of tumours.28–30 Presumably as a consequence of their asymptomatic nature, these tumours tend to be large; most present with metastases.

Some patients also present with unexplained orthostatic hypotension that, on a background of hypertension, provides an important diagnostic clue for the presence of a phaeochromocytoma. Occasionally,
patients with predominantly epinephrine-secreting tumours present with hypotension or even shock. Pathophysiological factors contributing to hypotension and shock, include intravascular volume depletion, abrupt cessation of catecholamine secretion due to tumour necrosis, desensitisation of adrenergic receptors, or hypocalcaemia. Shock can also be caused by a complicating cardiovascular emergency, such as myocardial infarction, cardiac arrhythmias, or a dissecting aortic aneurysm. Other cardiovascular complications of phaeochromocytoma include sudden death, heart failure due to toxic cardiomyopathy, hypertensive encephalopathy, a cerebrovascular accident, or neurogenic pulmonary oedema. Since these disorders, when occurring without phaeochromocytoma, are often accompanied by strong increments in plasma catecholamines, the exclusion or confirmation of an eventual underlying phaeochromocytoma in these patients is especially difficult.

Paroxysmal signs and symptoms, a consequence of episodic secretion of catecholamines, provide compelling clues for a phaeochromocytoma. Anaesthesia and tumour manipulation are the most well-known stimuli to elicit a catecholaminergic crisis. Food, micturition (urinary bladder phaeochromocytoma), and various chemical compounds or drugs (eg, glucagon, radiographic contrast substances, tyramine, metoclopromide, and tricyclic antidepressants) might also induce paroxysms. Such spells are usually unpredictable. For most patients they last between several minutes and 1 hour.

Despite improved diagnostic techniques that can bring about an earlier diagnosis of phaeochromocytoma, there still usually remains a delay of 3 years between initial symptoms and a final diagnosis. The most obvious reason for this delay is that in daily clinical patient care, the individual symptoms are quite non-specific—especially headaches, palpitations, and sweating, which are the most frequent. Nevertheless, if all three symptoms present together, the specificity of this combination is reported to be more than 90%.

Hypertension from a phaeochromocytoma during pregnancy can mimic pre-eclampsia, so that the diagnosis is delayed or even missed entirely. Advances in diagnosis and genetics now challenge the traditional rule of 10 for phaeochromocytomas (10% bilateral, 10% extra-adrenal, 10% familial, 10% malignant). Prevalence of bilateral adrenal tumours is higher than 10% in some familial phaeochromocytoma syndromes such as multiple endocrine neoplasia type 2 and von Hippel-Lindau syndrome. Prevalence of extra-adrenal tumours can reach 20%, and up to a quarter or more are hereditary. Finally, although metastases might be rare for adrenal phaeochromocytomas (up to 5%), the prevalence of malignant disease is about 33% for extra-adrenal phaeochromocytomas and even higher in patients with specific mutations such as those causing some forms of familial paragangliomas (eg, SDHB, a gene that encodes the B subunit of mitochondrial succinate dehydrogenase). We further review here the advances in genetics, biochemical diagnosis, and tumour imaging techniques and how these advances affect the spectrum of disease presentations that must now be considered and the available options for disease management and treatment.
Genetics of phaeochromocytoma

Advances in genetics and recognition of the high prevalence of phaeochromocytoma in certain familial syndromes led to mandatory routine screening in patients with identified mutations, even in the absence of typical clinical signs and symptoms. Accumulating evidence also suggests that a hereditary basis for phaeochromocytoma is more frequent than previously thought, accounting for up to 24% of patients with the tumour with no obvious initial evidence of a syndrome or family history.

So far, germline mutations in five genes have been identified to be responsible for familial phaeochromocytomas: the von Hippel-Lindau gene (VHL), which causes von Hippel-Lindau syndrome; the RET gene leading to multiple endocrine neoplasia type 2; the neurofibromatosis type 1 gene (NF1), which is associated with von Recklinghausen’s disease; and the genes encoding the B and D subunits of mitochondrial succinate dehydrogenase (SDHB and SDHD), which are associated with familial paragangliomas and phaeochromocytomas (table 2). Phaeochromocytomas are not always present and usually are not the first clinical manifestation of syndromes due to mutations of VHL, RET, and NF1 genes. Phaeochromocytomas in these three syndromes are usually associated with other benign or malignant neoplasms (panel 2).

Von Hippel-Lindau syndrome

Renal clear-cell carcinomas and cysts, CNS and retinal hemangioblastomas, phaeochromocytomas, pancreatic tumours and cysts, endolymphatic tumours, and epididymal cysts occur in von Hippel-Lindau syndrome, which affects about one in 36 000 livebirths. Presentation of these clinical manifestations can vary, with two broad types (1 and 2) depending on the presence or absence of a family history of phaeochromocytoma. Overall, phaeochromocytoma is present in 10–20% of patients with the syndrome, with a mean age at presentation of 30 years. Phaeochromocytomas in von Hippel-Lindau syndrome produce norepinephrine but not epinephrine. These tumours often have a bilateral adrenal presentation, and occasionally are multifocal with abdominal or thoracic localisations. Malignant disease is rare, at about 5% or lower.

Multiple endocrine neoplasia type 2

In multiple endocrine neoplasia type 2, phaeochromocytoma is the first clinical manifestation in 10–30% of patients, but penetrance is ultimately about 50%. Phaeochromocytomas in these patients produce both epinephrine and norepinephrine, with production of epinephrine occasionally predominating. Most patients (50–80%) develop bilateral adrenal tumours, either simultaneously or at different times. Extra-adrenal localisation or malignant disease are very rare (<5%).

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Exons</th>
<th>Protein</th>
<th>Frequency of germline mutations in apparent sporadic phaeochromocytoma*</th>
<th>Frequency of malignant disease*</th>
</tr>
</thead>
<tbody>
<tr>
<td>VHL 3p25-26</td>
<td>3</td>
<td>pVHL39 and pVHL30</td>
<td>2–11%</td>
<td>5%</td>
</tr>
<tr>
<td>RET 10q11.2</td>
<td>21</td>
<td>Tyrosine-kinase receptor</td>
<td>&lt;5%</td>
<td>3%</td>
</tr>
<tr>
<td>NF1 17q11.2</td>
<td>59</td>
<td>Neurofibromin</td>
<td>Unknown</td>
<td>11%</td>
</tr>
<tr>
<td>SDHB 1p36.13</td>
<td>8</td>
<td>Catalytic iron-sulfur protein</td>
<td>3–10%</td>
<td>50%</td>
</tr>
<tr>
<td>SDHD 11q23</td>
<td>4</td>
<td>Cyb5 (membrane-spanning subunit)</td>
<td>4–7%</td>
<td>&lt;3%</td>
</tr>
</tbody>
</table>

VHL = von Hippel-Lindau syndrome. *Data from references 13, 24, and 42–45.

Table 2: Characteristics of genes associated with familial forms of phaeochromocytoma

Panel 2: Main clinical features of syndromes associated to phaeochromocytoma

**Von Hippel-Lindau syndrome**

Type 1 (no phaeochromocytoma)

- Renal-cell cysts and carcinomas
- Retinal and CNS haemangioblastomas
- Pancreatic neoplasms and cysts
- Endolymphatic sac tumours
- Epididymal cystadenomas

Type 2 (with phaeochromocytoma)

A: Retinal and CNS haemangioblastomas
   - Phaeochromocytomas
   - Endolymphatic sac tumours
   - Epididymal cystadenomas

B: Renal-cell cysts and carcinomas
   - Retinal and CNS haemangioblastomas
   - Pancreatic neoplasms and cysts
   - Phaeochromocytomas
   - Endolymphatic sac tumours
   - Epididymal cystadenomas

C: Phaeochromocytomas only

**Multiple endocrine neoplasia type 2**

A: Medullary thyroid carcinoma
   - Phaeochromocytoma
   - Hyperparathyroidism
   - Cutaneous lichen amyloidosis

FMTC: Familial medullary thyroid carcinoma only

B: Medullary thyroid carcinoma
   - Phaeochromocytoma
   - Multiple neuromas
   - Marfanoid habitus

**Neurofibromatosis type 1**

- Multiple fibromas on skin and mucosae
- “Café au lait” skin spots
- Phaeochromocytomas

**Paraganglioma syndromes**

- Head and neck tumours (carotid-body tumours; vagal, jugular, and tympanic paragangliomas)
- Phaeochromocytomas
- Abdominal or thoracic paragangliomas (or both)
Neurofibromatosis type 1 syndrome
In neurofibromatosis type 1, pheochromocytoma is relatively rare (<5%). Because of this, routine screening for the tumour is not generally recommended. Genetic testing in members of an affected family is possible.47 Phaeochromocytomas in affected individuals usually produce both epinephrine and norepinephrine.

Paraganglioma syndromes
The germline mutations of the succinate dehydrogenase gene family are the most recently identified hereditary causes of pheochromocytoma.43,53 Mutations of each of the SDHB and SDHD genes have been seen in about 3–11% of patients with non-syndromic pheochromocytoma, mainly occurring as paragangliomas in the head, chest, and abdomen (panel 2).43,44 Head and neck paragangliomas are more commonly associated with SDHD than with SDHB mutations.43,53 Penetrance seems slightly higher and multifocal disease seems to be more frequently associated with SDHD than with SDHB mutations, but SDHB mutations are associated with an increased rate of malignant disease (up to 50%).44,45 The chromaffin tumours due to mutations of both genes produce predominantly norepinephrine. Whether other malignant diseases are associated with SDHB mutations such as renal-cell carcinoma and thyroid papillary carcinoma is not currently clear.44 SDHB mutations are maternally imprinted; thus, only carriers who have inherited the mutation from the father develop the disease and those who inherit from the mother are disease-free.44

Genetic testing
Should genetic testing be propagated in all patients with a pheochromocytoma? It has been suggested that all patients with a pheochromocytoma should be considered for genetic testing.44,53 The reasons are twofold. First, the syndromic hereditary forms of pheochromocytoma are associated with other neoplasms, so an early diagnosis of a hereditary syndrome might lead to regular surveillance and eventual treatment and thus improve the prognosis; this could extend to other family members with similar benefits. Second, in patients with proven germline mutations, multiple pheochromocytomas and recurrences are highly probable, so that a more stringent clinical follow-up is recommended throughout life. Presently, the indication for cost-effective genetic testing is recommended to those patients who have a positive family history or those who are younger than 50 years, especially children.44 However, other clues that should be regarded as an indication for genetic testing include the presence of bilateral pheochromocytomas or multifocal extra-adrenal disease, or the association of pheochromocytomas with other tumours. Direction of genetic testing to one of the suspected genes can benefit from consideration of the clinical picture and biochemical phenotype of the tumour: bilateral epinephrine-producing tumours or the association of pheochromocytoma with medullary thyroid carcinoma indicate the need for RET analysis; predominantly norepinephrine-producing pheochromocytomas—especially if bilateral or in association with renal-cell carcinoma or cysts, retinal or cerebrospinal haemangioblastomas, or pancreatic tumours—indicate VHL testing; and the association between chromaffin tumours and head or neck paragangliomas indicates analysis of SDHB or SDHD genes.

Biochemical diagnosis and localisation
Biochemical testing
All patients with suspected pheochromocytoma should undergo biochemical testing. These include patients with paroxysmal signs or symptoms suggestive of a pheochromocytoma; patients with recent, therapy-resistant, or volatile hypertension; patients with a paradoxical blood pressure response during surgery and anaesthesia; patients with a hereditary predisposition for a pheochromocytoma; asymptomatic patients with an adrenal incidentaloma; and patients with sudden attacks of anxiety. Because of the low prevalence of pheochromocytoma, biochemical screening for the tumour in asymptomatic patients with hypertension is not indicated.

Biochemical presentation of excessive production of catecholamines is an essential step for the diagnosis of pheochromocytoma. Traditional biochemical tests include measurements of urinary and plasma catecholamines, urinary metanephrines (normetanephrine and metanephrine), and urinary vanillylmandelic acid (VMA). Measurements of plasma-free metanephrines (normetanephrine and metanephrine) represent a more recently available test.45 Because of insufficient sensitivity and specificity, chromogranin A has no additional benefit over the use of catecholamines or their metabolites for initial diagnosis of pheochromocytoma.45–47

The potentially fatal consequences of a missed diagnosis justify the need for a high level of reliability of a positive test result in that rare patient with the tumour. This conversely also provides confidence that a negative test result reliably excludes the tumour. The initial examination of a patient with suspected pheochromocytoma should therefore include a suitably sensitive biochemical test. Either blood or urine testing can be used, with each test having its own advantages and disadvantages. Accumulating evidence suggests that
measurements of plasma-free metanephrines or urinary-fractionated metanephrines (normetanephrine and metanephrine separately) are the most sensitive tests for diagnosis, and are the most suitable for reliable exclusion of phaeochromocytoma (table 3).\(^{50-66}\) Increased sensitivity of metanephrines compared with catecholamines is due to the continuous production of O-methylated metabolites in tumours from catecholamines leaking from chromaffin stores. The production of O-methylated metabolites is independent of the highly variable release of catecholamines.\(^{67,68}\) Although tumours produce and metabolise catecholamines, they do not always release catecholamines. Provided appropriate reference intervals are used, the high diagnostic sensitivities of plasma-free or urinary-fractionated metanephrines mean that negative test results exclude the presence of virtually all phaeochromocytomas. Exceptions include asymptomatic small tumours that produce and metabolise negligible amounts of norepinephrine or epinephrine.

As with all biochemical tests of catecholamine excess, a remaining difficulty is that a positive result for plasma or urinary metanephrines does not always reliably indicate a phaeochromocytoma. The many physiological stimuli, drugs, and clinical conditions that cause increases in circulating catecholamines and metabolites compound this problem. The rarity of the tumour in tested patients, many of whom have symptoms associated with increased sympathetic activity and circulating catecholamines (eg, hypertension, heart failure, stroke, baroreflex failure, cardiogenic shock), also means that false-positive results exceed true-positive results.

Most true-positive results can be distinguished from false-positive test results from the magnitude of increases in test results above reference intervals (table 4). Most patients with phaeochromocytoma have increases well above even the highest false-positive results. In these patients, diagnosis is straightforward. Further confirmation can be achieved by the demonstration of similar patterns of increased plasma and urinary normetanephrine and metanephrine or repeat testing with an alternative method.

The major remaining difficulties are patients with smaller increases (less than two to three times the upper reference limits), most of whom will not have phaeochromocytoma. Drugs, dietary interferences, or inappropriate sampling conditions are causes of false-positive results. False-positive results could arise either due to direct interference with analytical methods or pharmacological effects on the disposition of catecholamines (table 3). Interference with analytical methods is highly variable from method to method. Avoidance of such interference needs intimate knowledge of the procedures used and how various drugs can affect them. Pharmacological effects on the disposition of catecholamines are independent of the measurement method and can be avoided by the withdrawal or substitution of drugs known to cause increases in catecholamines and their metabolites. Phenoxybenzamine (used to treat patients with suspected phaeochromocytoma) and tricyclic antidepressants are major causes of false-positive results, in one study accounting for up to 45% of false-positive increments of urinary and plasma normetanephrine and norepinephrine.\(^{69}\) Sampling of blood after overnight fasting and in the supine position can easily avoid the effects of diet and physical activity on plasma measurements. For urine measurements, timed or overnight urine samples or volume correction using additional measurements of creatinine can avoid difficulties with incomplete or overzealous collection of urine.\(^{64}\)

Use of clonidine to suppress catecholamine release from the sympathoadrenal system provides a dynamic

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma-free metanephrines</td>
<td>99%</td>
</tr>
<tr>
<td>Plasma catecholamines</td>
<td>84%</td>
</tr>
<tr>
<td>Urinary catecholamines</td>
<td>86%</td>
</tr>
<tr>
<td>Urinary-fractionated metanephrines</td>
<td>97%</td>
</tr>
<tr>
<td>Urinary total metanephrines</td>
<td>77%</td>
</tr>
<tr>
<td>VMA</td>
<td>64%</td>
</tr>
</tbody>
</table>

Sensitivity values of all tests for familial phaeochromocytoma are lower than that for sporadic phaeochromocytomas; the reverse is the case for specificity values. Table adapted from reference 64.

**Table 3**: Sensitivity and specificity of biochemical tests for diagnosis of phaeochromocytoma

<table>
<thead>
<tr>
<th>Presence of phaeochromocytoma</th>
<th>Unlikely*</th>
<th>Possible</th>
<th>Likely†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urine tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catecholamines (HPLC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norepinephrine (nmol/24 h)</td>
<td>&lt;500</td>
<td>500–1180</td>
<td>&gt;1380</td>
</tr>
<tr>
<td>Epinephrine (nmol/24 h)</td>
<td>&lt;100</td>
<td>100–170</td>
<td>&gt;170</td>
</tr>
<tr>
<td>Fractionated metanephrines (HPLC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normetanephrine (nmol/24 h)</td>
<td>&lt;3000</td>
<td>3000–6550</td>
<td>&gt;6550</td>
</tr>
<tr>
<td>Metanephrine (nmol/24 h)</td>
<td>&lt;1000</td>
<td>1000–2880</td>
<td>&gt;2880</td>
</tr>
<tr>
<td>Total metanephrines (spectrophotometry)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total of normetanephrine and metanephrine (μmol/24 h)</td>
<td>&lt;6</td>
<td>6–12.7</td>
<td>&gt;12.7</td>
</tr>
<tr>
<td>VMA (spectrophotometry)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VMA (μmol/24 h)</td>
<td>&lt;40</td>
<td>40–55</td>
<td>&gt;55</td>
</tr>
<tr>
<td>Blood tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catecholamines (HPLC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noradrenaline (nmol/L)</td>
<td>&lt;3.00</td>
<td>3.00–7.70</td>
<td>&gt;7.70</td>
</tr>
<tr>
<td>Adrenaline (nmol/L)</td>
<td>&lt;0.45</td>
<td>0.45–1.20</td>
<td>&gt;1.20</td>
</tr>
<tr>
<td>Free metanephrines (HPLC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normetanephrine (nmol/L)</td>
<td>&lt;0.60</td>
<td>0.60–1.40</td>
<td>&gt;1.40</td>
</tr>
<tr>
<td>Metanephrine (nmol/L)</td>
<td>&lt;0.30</td>
<td>0.30–0.42</td>
<td>&gt;0.42</td>
</tr>
</tbody>
</table>

HPLC=high-pressure liquid chromatography. Unlikely=number of true-negative results; Possible=number of false-negative results; Likely=number of false-positive results.

**Table 4**: Likelihood of phaeochromocytoma at different cut-off points for biochemical tests of catecholamine excess.

Seminar
Evidence from one small study indicates that contrast-enhancement with iohexol has no effect on plasma catecholamines. Until this finding is confirmed, patients should be protected from a hypertensive crisis or cardiac arrhythmia by combined blockade by α and β adrenoceptors. Although T2-weighted MRI with gadolinium enhancement has a similar diagnostic sensitivity to CT scanning, it is preferred for the localisation of extra-adrenal tumours or tumours during pregnancy, in children, or in patients with allergies to contrast. No adrenergic blockade is needed. Since CT scanning and MRI have similar sensitivities (90–100%) and specificities (70–80%), MRI is the preferred procedure.

Because of the restricted specificities of CT and MRI, identification of a mass detected by these techniques can be achieved by use of the increased specificity (95–100%) of 123I-metaiodobenzylguanidine (MIBG) scanning. If 123I-MIBG is not available, 131I-MIBG could be used as an alternative, but has poorer imaging qualities than 123I-MIBG. The coupling of functional with anatomical imaging might also be valuable for the detection of additional multifocal or metastatic tumours, which can be important for appropriate guiding of subsequent management and treatment. However, 123I-MIBG scanning might not be warranted before surgery in all patients with biochemically proven phaeochromocytomas. Such imaging is most relevant in patients with extra-adrenal or large (>5 cm) adrenal tumours with increased risk of malignant disease, or in patients with high suspicion of the presence of multifocal disease.

Several drugs (eg, labetalol, tricyclic antidepressants, and specific calcium antagonists) can interfere with tumour uptake or retention of 123I-MIBG. Temporary withdrawal of these interfering drugs (five times their half-lives) or use of other drugs can avoid false-negative test results. In patients whose 123I-MIBG scans are negative, 11C-octreotide scanning might be useful in case no other functional imaging techniques are available. Small recurrent tumours or metastases in the adrenal region can be detected intraoperatively by a γ-detector probe after the isotope is given a few hours before surgery.

PET with 18F-fluorodopamine, 18F-fluorodesoxyglucose, or 18F-fluorodeoxyglucose, or C-hydroxyephedrine are other functional imaging methods that can be used as alternatives to 123I-MIBG or as additional procedures if 123I-MIBG scanning is negative. The only PET imaging compound that is widely available, is not recommended for initial diagnostic localisation, since it is non-specific for phaeochromocytoma and sensitivity is restricted. However, the compound can be useful if other imaging procedures are negative, often in more rapidly growing dedifferentiated tumours that have lost the ability to accumulate more specific drugs. PET with 18F-fluorodopamine PET offers better diagnostic sensitivity than 123I-MIBG scintigraphy, especially in metastatic phaeochromocytomas where the compound

| Table 5: Sources of variable interference with measurements of catecholamines and catecholamine metabolites |
|---------------------------------|---------------------------------------------------------------|
| **Analytical methods** | **Nature of interference** |
| Coffee (including decaffeinated coffee) | HPLC assays: plasma catecholamines |
| Labetalol | Spectrophotometric and fluorometric assays: urinary catecholamines and metanephrines |
| Sotalol | HPLC assays: plasma catecholamines |
| Buspirone | HPLC assays: urinary metanephrines |
| Paracetamol | HPLC assays: plasma-free metanephrines |
| Levodopa | HPLC assays: catecholamines and metabolites |
| α-methylated p-dopa | HPLC assays: catecholamines |
| Symathomimetics (eg, amfetamines, ephedrine) | Spectrophotometric and fluorometric assays: plasma and urinary catecholamines |

HPLC—high-pressure liquid chromatography.

pharmacological test to distinguish increased catecholamine release due to sympathetic activation from increased release due to phaeochromocytomas. Failure to suppress plasma norepinephrine (ie, a reduction of <50% from basal or consistently raised basal plasma concentrations of >3 nmol/L) after clonidine is highly predictive for phaeochromocytoma (97%). By contrast, the negative predictive value of a normal test result is only 75%. If plasma normetanephrine is used (of which a failure to suppress is a reduction of less than 40% from basal or consistently raised basal concentrations of plasma normetanephrine of more than 0·60 nmol/L) instead of plasma norepinephrine, the positive and negative predictive values of this test improve to 100% and 96%, respectively. Thus, although absence of suppression of plasma norepinephrine or normetanephrine both provide strong evidence for phaeochromocytoma, only the suppression of normetanephrine provides reasonable evidence that a phaeochromocytoma is not present.

**Imaging procedures**

Tumour localisation should ideally only be initiated once there is unequivocal biochemical evidence for phaeochromocytoma. CT scans of the entire abdomen (including pelvis), with and without contrast, are most often used for initial localisation of adrenal or possible extra-adrenal abdominal phaeochromocytomas.
showed that use of metirosine as an adjunct to preoperative treatment. Two retrospective studies blocks catecholamine synthesis, is occasionally used for completely. of not causing orthostatic hypotension, but if used alone blockers. Calcium-channel blockers have the advantage than other and suitable for pretreatment than other compounds such as doxazosin, which avoids drug the blockade of (H9251)-adrenoceptors with phenoxybenzamine. Treatment consists of fluid replacement and occasionally intravenous ephedrine. If ephedrine infusion is ineffective, vasopressin might be used. The risk of hypoglycaemia is related to rebound hyperinsulinaemia due to the recovery of insulin release after tumour removal. Pretreatment with an (H9251)-adrenergic blocker can usually be undertaken on an outpatient basis and is safe in most patients. Treatment usually lasts for 10–14 days. The initial dose of phenoxybenzamine is 10 mg twice a day. The dose is increased every 2–3 days by 10–20 mg to a total daily dose of 1 mg/kg, which is sufficient in most patients. Doxazosin is given in increasing doses from 1 to 16 mg once a day. A (H9251)-adrenergic blocker (eg, propranolol 40 mg three times daily or atenolol 25–50 mg once daily) could be included after several days of (H9251)-adrenergic blockade. This addition is especially useful in patients who also have tachyarrhythmias. Blockade of (H9251)-adrenoceptors should never be initiated before blockade of (H9251)-adrenoceptors, since the loss of (H9251)-adrenoceptor-mediated vasodilatation leaves (H9251)-adrenoceptor stimulation unopposed, which could result in hypertensive crises.

To ensure adequate preoperative preparation, several criteria have been proposed: blood pressure should be reduced to below 160/90 mm Hg for at least 24 h; orthostatic hypotension should be present, but blood pressure in the upright position should not fall below 80/45 mm Hg; there should be no more than one ventricular extrasystole every 5 min; and the electrocardiogram should show no S-T segment changes and T-wave inversions for 1 week. Risk of excessive orthostatic hypotension can be kept to a minimum by the increase of salt and fluid intake. The additional advantage of this approach is that it reduces the risk of postoperative hypotension. Should substantial rises in blood pressure still take place during surgery, these can be controlled by bolus or by continuous infusion of phentolamine, sodium nitroprusside, or a shortacting calcium antagonist (eg, nicardipine); tachyarrhythmias can be treated by infusion of a shortacting (H9251)-adrenoceptor blocker (eg, esmolol). After surgery, patients need to be under close surveillance for the first 24 h in an intensive or intermediate care unit. The two major postoperative complications are hypotension and hypoglycaemia. Postoperative hypotension is due to the abrupt fall in circulating catecholamines after tumour removal in the continuing presence of (H9251)-adrenoceptor blockade (by phenoxybenzamine). Treatment consists of fluid replacement and occasionally intravenous ephedrine. If ephedrine infusion is ineffective, vasopressin might be used. The risk of hypoglycaemia is related to rebound hyperinsulinaemia due to the recovery of insulin release after tumour removal.

Surgical treatment
Laparoscopic removal of intra-adrenal and extra-adrenal phaeochromocytomas is now the preferred surgical
technique.\textsuperscript{106} Data from observational studies clearly show that the laparoscopic procedure reduces postoperative morbidity, hospital stay, and expense compared with conventional laparotomy.\textsuperscript{106–112} The complication rate of laparoscopic adrenalectomy (apart from intraoperative blood pressure derails) is probably less than 8% with a conversion rate of 5%.\textsuperscript{109,113,114} Some clinicians recommend a retroperitoneal laparoscopic approach for suprarenal paragangliomas and a transperitoneal laparoscopic approach for infrarenal paragangliomas.\textsuperscript{110}

Because of the high incidence of bilateral adrenal disease in familial pheochromocytoma (eg, von Hippel-Lindau and multiple endocrine neoplasia type 2 syndromes), adrenal-cortex-sparing laparoscopic surgery (partial adrenalectomy) has been advocated to save the cortex and prevent permanent glucocorticoid deficiency.\textsuperscript{115–121}

Currently, after adequate medical preparation, operative mortality is less than 1% if undertaken by an experienced anaesthesiologist and a skilful surgeon.\textsuperscript{114,117,122} The long-term prognosis of patients after resection of a solitary sporadic pheochromocytoma is excellent, although hypertension might persist after surgery in nearly 50% of patients.\textsuperscript{123} Findings from a large study with a long-term follow-up showed a recurrence rate of 17%, with half the patients showing signs of malignant disease.\textsuperscript{27} Recurrences occurred more often in patients with extra-adrenal disease (33%) than in those with adrenal disease (14%), and more often in the familial population (33%) than in the non-familial population (13%).\textsuperscript{124} In patients who have undergone adrenal cortex-sparing surgery, a small piece of tumour could have been left behind or a new tumour could develop in the remnant because of a genetic predisposition. The risk of tumour recurrence in the remnant adrenal gland is 10%, but the metachronous tumour development in the contralateral adrenal gland is 30% in patients with hereditary forms of pheochromocytoma.\textsuperscript{19,119,124} All patients should be followed up every year for at least 10 years after surgery. Patients with extra-adrenal or familial pheochromocytoma should be followed up indefinitely.

**Malignant pheochromocytomas**

Despite the increasing availability of molecular diagnostic and prognostic markers, it remains impossible to predict the subsequent development of malignant disease, based on histological findings in a resected tumour.\textsuperscript{25,126} Not one histological feature predicts or provides unequivocal evidence of malignant derangement. Only the presence of metastases of chromaffin tissue at sites where no chromaffin tissue should be expected establishes a definite diagnosis of malignant pheochromocytoma. The most common metastatic sites are the bones, lungs, liver, and lymph nodes. In general, tumours that are large (>5 cm) or have an extra-adrenal location have a higher risk for malignant disease than tumours that are small or have an adrenal location.\textsuperscript{19} Paragangliomas in patients with SDHB mutations have a particularly high rate of malignant disease.\textsuperscript{25,23} Increased plasma or urinary concentrations of dopamine and dihydroxyphenylalanine (dopa) arise more often in malignant than in benign pheochromocytomas.\textsuperscript{127–129}

When malignant disease is confirmed, the natural clinical course is highly variable in patients with 5-year survival rates of 50%.\textsuperscript{130}

There remains no effective treatment for malignant pheochromocytoma. Radical surgical removal of tumour tissue is the mainstay to improve symptoms and survival, but there are no randomised studies that support this approach. Symptomatic treatment can be obtained with α-adrenergic blockers. α-methyl-paratyrosine can be useful in selected patients with very high concentrations of circulating catecholamines.

Treatment with \textsuperscript{131}I-MIBG or combination chemotherapy (cyclophosphamide, vincristine, and darcabazine) has shown disappointing results, with only shortlasting remissions.\textsuperscript{130–132} Although nearly 80% of patients show symptomatic improvement after \textsuperscript{131}I-MIBG treatment, less than 5% have a complete remission and 30% have a partial tumour remission.\textsuperscript{133,132} Improved long-term survival has been shown with increased doses of \textsuperscript{131}I-MIBG.\textsuperscript{134} However, controlled studies need to confirm these results and establish whether combination treatment of \textsuperscript{131}I-MIBG with myeloablative chemotherapy (and stem cell rescue) can improve prospects of patients with metastatic disease.

**Conflict of interest statement**

We declare that we have no conflict of interest.

**References**

Seminar


Buying results? Contracting for health service delivery in developing countries

Benjamin Loevinsohn, April Harding

To achieve the health-related Millennium Development Goals, the delivery of health services will need to improve. Contracting with non-state entities, including non-governmental organisations (NGOs), has been proposed as a means for improving health care delivery, and the global experience with such contracts is reviewed here. The ten investigated examples indicate that contracting for the delivery of primary care can be very effective and that improvements can be rapid. These results were achieved in various settings and services. Many of the anticipated difficulties with contracting were either not observed in practice or did not compromise contracting's effectiveness. Seven of the nine cases with sufficient experience (greater than 3 years’ elapsed experience) have been sustained and expanded. Provision of a package of basic services by contractors costs between roughly US$3 and US$6 per head per year in low-income countries. Contracting for health service delivery should be expanded and future efforts must include rigorous evaluations.

Background
Substantial improvement in the delivery of health services will be necessary to achieve the health-related Millennium Development Goals (MDGs). For example, 63% of child deaths in developing countries could be prevented through the full implementation of a few effective and low-cost interventions.1 Hence, discovering better ways of delivering these services is critically important. Although many countries undoubtedly need to allocate more resources to health services, experience suggests that simply throwing money at the problem of service delivery is unlikely to have much of an effect.2 Another response to the challenge of improving service delivery has been to use public funds to contract with non-state entities, such as non-governmental organisations (NGOs), universities, individual practitioners, or for-profit companies.

Contracting for health service delivery has some potentially attractive features,3 including the possibility of: (i) ensuring a greater focus on the achievement of measurable results, especially if contracts define objectively verifiable outputs and outcomes; (ii) overcoming the constraints that prevent governments from effectively using the resources made available to them (often referred to as absorptive capacity issues); (iii) using the private sector’s greater flexibility and generally better morale to improve services; (iv) increasing managerial autonomy and decentralising decision-making to managers on the ground; and (v) using competition to increase effectiveness and efficiency; and (vi) allowing governments to focus more on other roles that they are uniquely placed to undertake, such as planning, standard setting, financing, regulation, and the various public health functions.

There are potential difficulties with contracting,4–10 including concerns that: (i) contracts will not be feasible at a sufficiently large scale to make a difference at a country level; (ii) contracts will be more expensive than government provision of the same services, partly indicating greater transaction costs; (iii) contracts might increase inequities in health service delivery; (iv) governments will have limited capacity to manage contracts effectively; and (v) even if successful, contracting will not be sustainable.

This review was undertaken to: (i) examine the effectiveness of contracting taking into account the methodological rigour of the evaluations; (ii) examine the extent to which anticipated difficulties have occurred during implementation; and (iii) make recommendations about future efforts in contracting. We focused on the effectiveness of contracting in terms of health service delivery outputs or outcomes, costs, and scale. Our review did not examine in detail the political economy issues associated with the decision to implement contracting in developing countries.

Methodology
There are several different approaches to contracting for health service delivery, so clarifying definitions will facilitate meaningful dialogue (table 1). As an example, under a management contract (entry 3 in table 1) a government will contract with a non-state entity or an individual to manage existing government services in a specified area. Under a service delivery contract (entry 4 in table 1) the government decides which services the contractor will provide and where, whereas the
contractor will both manage and supply the production infrastructure, such as personnel, equipment, drugs, etc. Table 1 shows the different approaches for contracting, which are not exhaustive—there are clearly hybrids. For example, the line between a management contract and a service delivery contract blurs when the contractor uses government health workers, but pays them substantially more than their civil service salaries. There has been some experience with national governments signing agreements with local governments (entry 2 in table 1) that pertain to achieving specific goals.

Although potentially interesting, this arrangement rarely entails a true contract that the parties enter voluntarily, and the contractor can be fired for poor performance (although other rewards and sanctions might be available). There were few examples that we identified in which instances of this approach have been assessed. Government or donor grants (where governments or donors issue requests for proposals and then make grants to the NGOs), to NGOs (entry 5 in table 1), in which the NGOs define where and what services are delivered, are quite common, particularly in HIV/AIDS prevention and treatment. However, these are generally not true contracts, partly because the government generally has little say in what services are delivered, or where, and how they will be assessed. We review experience with management and service delivery contracts, rather than contracts between different levels of government or grants to NGOs.

The focus of this review is instances in developing countries of governments or their agents contracting with identifiable non-state providers to deliver primary health-care services including nutrition (but excluding hospital care or ancillary services such as drug procurement and distribution) in which some explicit form of evaluation was undertaken. To be included in the review, the project had to have been explicitly investigated by measures of quality of care, or outputs, such as increase in the amount of services provided. The evaluations also had to, at a minimum, entail before and after or controlled designs. Hence, evaluations that described the process of contracting but did not measure some tangible outputs were excluded. Furthermore, we excluded studies that only provided after evaluations with no before data or without contemporaneous controls. In instances in which there were before and after data for experimental and control groups, the double differences were calculated. The double difference is the difference between the follow-up and baseline results in the experimental group minus the difference between the follow-up and baseline results in the control group. Wherever possible, differences are expressed in percentage points. Although we did not plan to exclude experiences of contracting with for-profit providers, the search criteria generated only cases with non-profit providers.

**Experience of contracting**

Table 2 summarises ten studies that met the inclusion criteria.11–20 Of these ten examples, four had before and after controlled designs, three had controlled designs with a single measure in time, and the remaining three were before and after assessments. There was only one randomised trial. Three of the studies relied on routinely obtained data from health information systems of unknown accuracy, whereas the remainder relied on information from household and health facility surveys.

From the studies reviewed, contracting with NGOs to deliver primary health or nutrition services seems to be very effective and impressive improvements can be seen.

---

**Table 1: Arrangements of service delivery**

<table>
<thead>
<tr>
<th>Initiator (defines services and area)</th>
<th>Selector (who chooses provider)</th>
<th>Manager</th>
<th>Production infrastructure</th>
<th>Source of financing</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Inter-government agreements</td>
<td>Government-1*</td>
<td>Government</td>
<td>Government-2*</td>
<td>Government†</td>
<td>Transfer of funds from federal to provincial governments</td>
</tr>
<tr>
<td>4. Service delivery contracts</td>
<td>Government</td>
<td>Private sector</td>
<td>Private sector</td>
<td>Government†</td>
<td>Government hires NGO to provide services where none currently exist</td>
</tr>
<tr>
<td>5. Grants to private sector</td>
<td>Private sector</td>
<td>Government or donor</td>
<td>Private sector</td>
<td>Government +/- NGO or community contribution</td>
<td>NGOs submit proposals to Government for needs identified by community or NGO</td>
</tr>
<tr>
<td>6. Vouchers</td>
<td>Government</td>
<td>Consumer</td>
<td>Private sector</td>
<td>Government (and/or donor)</td>
<td>Female sex workers are provided vouchers for curative care which they can redeem at practitioners of their choice</td>
</tr>
<tr>
<td>7. Franchising</td>
<td>Private sector</td>
<td>Consumer</td>
<td>Private sector</td>
<td>Consumer +/- subsidy from Government or donor</td>
<td>Private practitioners join franchise network providing reproductive health services</td>
</tr>
<tr>
<td>8. Private sector services</td>
<td>Private sector</td>
<td>Consumer</td>
<td>Private sector</td>
<td>Consumer or NGO/donor</td>
<td>1. NGO establishes health services in slum areas using its own funds</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. For-profit providers establish private clinic</td>
</tr>
</tbody>
</table>

NGO=non-governmental organisations, †=with or without. *Government-1=higher level of government, *Government-2=local level of government. †Can be supplemented by formal or informal user-charges.
<table>
<thead>
<tr>
<th>Location &amp; type of services (ref)</th>
<th>Type of contract &amp; intervention</th>
<th>Scale &amp; cost</th>
<th>Evaluation methodology</th>
<th>Main results</th>
<th>Subsequent history</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cambodia</td>
<td>Rural PHC &amp; district hospital services&lt;sup&gt;11,12&lt;/sup&gt;</td>
<td>SDC compared to MC and CC</td>
<td>1.5 million</td>
<td>Randomised controlled study with 12 districts as experimental units. HHS and HFS undertaken B&amp;A 2.5 years of implementation</td>
<td>SDC and MC much better than CC. Median double difference on seven indicators for SDC versus CC was 21.3 percentage points compared with 3.1 percentage points in the control districts (a double difference of 18.2 percentage points).</td>
</tr>
<tr>
<td>2. Bangladesh</td>
<td>Rural community nutrition services&lt;sup&gt;14&lt;/sup&gt;</td>
<td>SDC with NGOs compared to control areas with no organised nutrition services (ie, normal government health services with no nutritional component)</td>
<td>15 million</td>
<td>Controlled B&amp;A study with six experimental and two control subdistricts. HHSs conducted by third party</td>
<td>Malnutrition rates declined 1.8 percentage points in SDC subdistricts compared with 13.1 percentage points in controls (double difference=5.5 percentage points). Double difference for vitamin A was 2.7 percentage points.</td>
</tr>
<tr>
<td>3. Bangladesh</td>
<td>Urban PHC&lt;sup&gt;23&lt;/sup&gt;</td>
<td>SDC with NGOs compared to government provision of services, ie, CCC</td>
<td>4 million</td>
<td>Controlled B&amp;A study with 15 contracts compared with a large area implemented by CCC. HHS and HFS survey by third party</td>
<td>Median double difference on 10 HHS indicators was 3.4 percentage points after 2 years. Much larger differences in quality of care indicators from HFS.</td>
</tr>
<tr>
<td>4. Bolivia</td>
<td>Urban PHC&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Limited MC in phase II. MC with expanded authority in phase III. Control area had continued public sector management</td>
<td>250 000</td>
<td>Controlled, B&amp;A design, but data from routine reporting system, only few indicators examined</td>
<td>Double difference for deliveries between MC and control was 21.1 percentage points, 3 percentage points for bed occupancy</td>
</tr>
<tr>
<td>5. Guatemala</td>
<td>Rural PHC in mountainous areas&lt;sup&gt;13&lt;/sup&gt;</td>
<td>MC in selected municipalities &amp; SDC in more remote areas, compared to government provision (control)</td>
<td>3.4 million</td>
<td>Controlled design based on HHS undertaken by third party 3 years after contracting began</td>
<td>Median difference between MC and control on five indicators was 11.5 percentage points (range 5-16 percentage points).</td>
</tr>
<tr>
<td>6. Haiti</td>
<td>Bonuses for NGOs delivering PHC in rural areas&lt;sup&gt;25&lt;/sup&gt;</td>
<td>NGOs with SDCs offered performance bonuses based on agreed targets</td>
<td>534 000</td>
<td>B&amp;A (7 months later) design based on HHS done by third party</td>
<td>Average of follow-up minus baseline ranged from 3.1 percentage points (antenatal care) to 13.2 percentage points (nutrition coverage).</td>
</tr>
<tr>
<td>7. India</td>
<td>Urban TB control services in Hyderabad&lt;sup&gt;19&lt;/sup&gt;</td>
<td>NGO under SDC delivered TB control services in defined population &amp; worked with private providers. Compared to publicly managed area of similar size</td>
<td>500 000 population</td>
<td>Controlled design with after only data from recording system verified by national TB programme officials. Cost data obtained by third party</td>
<td>NGO found 21.2 percentage points more TB cases and had 14.8 percentage points better treatment success rate. Cost per successful treatment $118 for NGO versus $138.</td>
</tr>
<tr>
<td>8. Madagascar &amp; Senegal Community nutrition services&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Madagascar: SDC with 50 NGOs in Senegal. SDCs with NGOs who worked through small groups of unemployed youth</td>
<td>460 000 in Madagascar &amp; 490 000 in Senegal</td>
<td>B&amp;A (17 months) HHS of nutrition status in Senegal. Third party survey of participation in project and control areas.</td>
<td>Severe and moderate malnutrition declined 6.5 percentage points and 4.5 percentage points, respectively. Participation was 72% in project &amp; 35% in control areas.</td>
<td>Continued with NGOs in both countries, albeit in a different format</td>
</tr>
<tr>
<td>9. Pakistan</td>
<td>Rural PHC (data obtained by authors)</td>
<td>MC for the 104 basic health units in one district</td>
<td>3.3 million</td>
<td>Intermittent time series design based on routine recording and reporting system</td>
<td>Nearly a four-fold increase in the number of outpatient visits.</td>
</tr>
<tr>
<td>10. India</td>
<td>Improving quality of care by private practitioners&lt;sup&gt;27&lt;/sup&gt;</td>
<td>SDC for NGO working with private providers to improve MCH services</td>
<td>54 000</td>
<td>B&amp;A (6 months later) design based on HHS by community health workers</td>
<td>Rapid improvement in provider skills ranging from 25 percentage points to 57 percentage points compared with baseline.</td>
</tr>
</tbody>
</table>

MC=management contract, SDC=service delivery contract, CC=control-comparison, double difference=difference between follow-up and baseline results in the experimental group minus the difference between follow-up and baseline results in the control group; B&A=before and after; HHS=household survey; HFS=health facility survey; TB=tuberculosis; NGO=non-governmental organisations; PHC=primary health care; CCC=Chittagong City Corporation, %p=percentage points, MCH=maternal and child health. All costs are in $US dollars.

Table 2: Summary of contracting experiences

achieved rapidly. Good results have been achieved in various settings and for many different services ranging from nutrition services in Africa to primary health care in Guatemala.<sup>17,20</sup> All the studies found that contracting yielded positive results; however, the most rigorously assessed cases tended to show the largest effect. For example, the service delivery contracts in Cambodia increased immunisation coverage by 40 percentage points compared with 19 in the control districts (a double difference of 21 percentage points). In the four studies with controlled before and after designs, the median double differences ranged from 3.4 to 26.0 percentage points (figure 1, table 3). These cases combined examined 31 main indicators, and all but one of the double differences was positive (ie, favoured contracting). Larger double differences were seen for those factors that are easier to change such as immunisation, vitamin A, and antenatal care coverage. Smaller changes were recorded for factors that require important behavioural changes, such as family planning and institutional delivery.

Six of the ten studies compared contractor performance with government provision of the same services. All six showed that the contractors were more effective than the government, on the basis of several measures related to both quality of care and coverage of services. In the studies reviewed here, the differences between contractor and government performance tended to be large. For example, in India an NGO achieved a treatment completion rate that was 14 percentage points higher than the public services in a nearby area, and at a lower cost.<sup>19</sup>
Several possible difficulties have been raised about contracting, some of which can be addressed by the examples reviewed.

First, contracting is able to provide services on a large scale. Half the examples studied involved populations of millions of beneficiaries, and in one example contracts now cover a third of rural Bangladesh—equating to more than 30 million people. Second, contracting can be more cost effective than directly provided government services. The studies from Pakistan, urban Bangladesh, Hyderabad, India, and the management contracts in Cambodia, indicate that non-governmental entities did better even when they had the same or fewer resources than public institutions. Although the services provided under the different contracts are not strictly comparable, a basic package of primary health services in rural areas ranged from US$2–82 per head per year in Cambodia to US$6–25 per head per year in Guatemala. These amounts represent less than 1% of the gross national income.

Third, contracting can increase coverage, even in poor, remote areas. With the resources and the explicit responsibility, many contractors were willing and able to work in difficult areas that had previously been under served. However, only the assessment in Cambodia explicitly addressed the issue of whether contracting could improve equity. It found that when contracts explicitly included targets for reaching the poor, contractors were able to greatly improve health services for the most marginalised groups. The results from this study also showed that contractors were considerably better than the government at reducing inequities (figure 2).

Last, contract management is often difficult for governments, but does not seem to prevent improvements in service delivery. Even in rural and urban Bangladesh and Guatemala, where observers felt contract management was not done well, contractors were still successful at implementing large-scale programmes. The fact that in some situations, such as Senegal, contract management was done well, suggests that the problem is tractable. The cases with successful contract management seem to have benefited from either external management support or having only a few contracts.

In view of its apparent success, the sustainability of contracting is a genuine concern. Nine of the examples of contracting reviewed had more than 3 years of elapsed experience to judge whether they were sustained or not. Seven of the nine contracts have been continued and expanded, often substantially. (Information about two examples is not available). In Guatemala, Cambodia, rural Bangladesh, Haiti, Pakistan, and India, the scope of contracting has more than doubled from what it was initially. However, it might still be too early to say whether the approach is sustainable in instances in which the donors introduced the contract.

### Discussion

Under real world conditions and at a large scale, contracting has achieved impressive and rapid results. The cases we review suggest but do not prove that the most successful approaches to contracting maximise the amount of autonomy given to contractors. This factor is shown clearly in the studies from Bolivia and Cambodia, especially in Cambodia where service delivery contracts did better than management contracts. This finding is consistent with the experience in hospital autonomy in which the ability to manage labour seems to be critically important to improved performance.13

The successful approaches also focus on outputs and outcomes, rather than inputs, a finding that accords with those of other studies.24 In practice this focus on outcomes requires careful attention to monitoring and assessment. These approaches use contracts of a fairly large size. To obtain economies of scale, reduce the contract management burden on the government, and facilitate monitoring and evaluation, each contract should probably include more than 500 000 beneficiaries.

There are several methodological limitations with this review: (i) half the cases were based on reports in the grey literature, some of which had not been peer reviewed: (i) half the cases were based on reports in the grey literature, some of which had not been peer

### Table 3: Full immunisation coverage in Cambodia—calculation of double difference

<table>
<thead>
<tr>
<th>Type</th>
<th>Baseline</th>
<th>Follow-up</th>
<th>Difference (follow-up – baseline)</th>
<th>Double difference versus control</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDC</td>
<td>25.5%</td>
<td>65.8%</td>
<td>40.3 (a)</td>
<td>21.3 (a–c)</td>
</tr>
<tr>
<td>MC</td>
<td>29.9%</td>
<td>54.4%</td>
<td>24.5 (b)</td>
<td>5.5 (b–c)</td>
</tr>
<tr>
<td>Control</td>
<td>34.0%</td>
<td>53.0%</td>
<td>19.0 (c)</td>
<td>--</td>
</tr>
</tbody>
</table>

Double difference is follow-up minus baseline in the experimental group minus follow-up minus baseline in the control. The range and median double differences are for the main indicators identified in each study: SDC = service delivery contract, MC = management contract.

### Figure 1: Double differences (in percentage points) in coverage rates from studies with controlled before and after methodology

SDC = service delivery contract, MC = management contract, PHC = primary health care.
in describing the distribution of income, in which the concentration index is referred to as the Gini coefficient. A Lorenz curve has the cumulative distribution of some benefit/service on the vertical access and income percentiles on the horizontal axis. This approach is commonly used.

Technically, it represents twice the area between the perfect equality line and the actual cumulative distribution in a Lorenz curve and, can range from −1 to +1. By convention a negative concentration index for services represents greater equity —ie, the distribution is pro-poor. A concentration index summarises the extent to which services are equitably provided across income groups.²²

Figure 2: Change in equity (concentration index) of services in Cambodia
FIC=fully immunised child coverage in children aged 12–23 months, VIT A=Vitamin A coverage in the past 6 months in children aged 6–59 months, ANC=antenatal care coverage, HF DEL=proportion of mothers delivering in a health facility, MBS=use of modern birth spacing, USE=utilisation of publicly-financed health facility by people who were sick in the past month, CC=control/comparison, MC=management contract, SDC=service delivery contract. A concentration index summarises the extent to which services are equitably provided across income groups.²³ Technically, it represents twice the area between the perfect equality line and the actual cumulative distribution in a Lorenz curve and, can range from −1 to +1. By convention a negative concentration index for services represents greater equity —ie, the distribution is pro-poor. A concentration index summarises the extent to which services are equitably provided across income groups.²³

Panel: Contractor versus government performance in middle-income countries
There are only a few examples of initiating contracting for primary health-care services in developed countries. Most countries that contract have done so. Nevertheless, a few opportunities have arisen to assess the effect of initiating contracting. Several central European countries have initiated contracting for packages of primary health-care services. Where contracted services have been compared with those that continued to be provided by salaried doctors—results have generally been favourable. In Croatia, evidence of higher productivity was noted in contracted practices, including indicators of patient accessibility.²⁴ In Estonia, where salaried doctors converted to a contracted status, a before and after analysis showed allocated efficiency indicators improved; technical efficiency indicators, such as annual number of visits per doctor and number of visits per inhabitant, improved; and, immunisation rates rose from 74% to 88%.²⁵

Contracting examples were done on a very large scale and provided services to many millions of beneficiaries. The history in the USA and Australia of contracting for social service delivery suggests that the initial experiences were problematic and that results improved as governments and contractors ironed out the difficulties they encountered.²⁶,²⁷

We should also keep in mind that all the cases discussed here focused on primary care and nutrition services (although two included first level hospital care as well)—services for which outputs are fairly easy to measure. Other health services, especially specialist inpatient care, present larger measurement challenges. Furthermore, the providers in these cases were non-profit organisations. Although contracting with for-profit entities, especially self-employed doctors, is common in developed and middle-income health systems (panel)—there is little experience in low-income countries with for-profit providers being given contracts for primary health care.

With the methodological concerns about the cases studied, there is still a need for future contracting efforts to include rigorous evaluations. However, the current weight of evidence suggests that contracting with non-governmental entities will provide better results than government provision of the same services. Contracting should no longer be considered an untested intervention or a so-called leap of faith.

We realise that our findings will be controversial. Contracting with non-state providers is often seen as arising out of an ideological desire to privatise publicly financed health services and ultimately to limit or end governments’ involvement in health care.²⁸,²⁹ However, the discussions held during the preparation of this review indicated that the impetus for all the contracting initiatives studied was the inadequate quality and coverage of government services, especially for poor people. For example, the case from Pakistan arose because a district governor and his advisers became frustrated with the poor quality of existing government provision of primary health care. Their approach involved no reduction in health care expenditures.

Far from limiting government involvement in health care, contracting may be one way of keeping publicly financed health care relevant. Governments in developing countries are currently responsible for only a modest role in providing curative services, even for the poor. For example, in south Asia, 80% of children in the lowest income quintile who are brought to care for acute respiratory tract infections use a private provider.³⁰

Although some argue that long-term government provision of services is essential, contracted provision is a well established model for delivering primary health-care services. In most of continental Europe, for example, social health insurance funds contract with independent providers. In Canada, the UK, and New Zealand, tax-based funding bodies similarly contract...
with independent providers (or groups) for provision of virtually all primary health-care services. Even in Scandinavia, where the government role in service provision is the greatest among Organisation for Economic Co-operation and Development (OECD) countries, private providers deliver a substantial amount of primary care. In Norway in 2001, 66% of primary care services were provided by private, contracted doctors, whereas 19% were delivered by salaried doctors.¹¹

**Recommendations**

On the basis of the success thus far, contracting frequently merits consideration in developing countries that are seeking to rapidly improve service delivery and achieve the MDGs. There is enough evidence supporting contracting to have it, at least, tried on a larger scale. Future efforts at contracting should continue to include rigorous evaluations to better determine its effectiveness, obtain robust estimates of the effect size, and test it under various conditions. Such operational research should also address remaining issues such as the effects of contracting on equity, the usefulness of performance-based bonuses, its cost-effectiveness compared with grants to NGOs, and different approaches to establishing the price of contracts. Ultimately, any debate about the effectiveness of contracting must be settled by systematically obtained evidence.

**Contributors**

B Loevinsohn conceived the idea of the paper, did much of the research, and helped write the article. A Harding co-authored the report.

**Conflict of interest statement**

B Loevinsohn was instrumental in setting up and funding the Cambodia and Bangladesh Urban PHC contracting initiatives. A Harding declares no conflict of interest. No external financing was received for this study.

**References**

End-of-life: a Hindu view

Shirley Firth

Introduction

The term Hindu was a Persian-Arabic label for dwellers of the Indus area. Hinduism does not have an institutional framework nor demand adherence to particular doctrines. Rather, it is a diverse umbrella or family of beliefs and practices. Nevertheless, Hindus have specific beliefs in common that influence their attitudes to death. These beliefs relate not to the end-of-life period, but to transition to another life, whether by reincarnation, life in heaven with God, or absorption into Brahman (the ultimate reality). They provide direction through the whole of life by appropriate religious rituals and meritorious acts, which relate to the dharma of one’s class, and by prayer and meditation. Of particular importance is the notion of a good death, which provides a model of how to die; a bad death is greatly feared.

Here, I discuss the key concepts for understanding Hindu attitudes to death; beliefs about the soul, heaven, hell, and liberation; and good and bad deaths; the roles of the family and priests in a good death, which are important for medical professionals to understand; and disclosure and ethical issues about withdrawal of treatment and euthanasia.

General terms

Hindus sometimes refer to their religion as sanatanadharma, the eternal religion or law. All Hindus are born into a particular class or varna. The four varnas are Brahmén (priests), kshatriyas (warriors and kings), vaishyas (merchants and traders), and sudras (peasants, labourers). These varnas are subdivided into occupation-based castes (jatis). Each caste has its own dharma, or social and ethical code for behaviour. Thus the warrior’s dharma might involve killing others in battle,7 which could also determine his attitude to religious suicide.

Dharma, “righteousness, morality, or virtuous conduct”,1 also informs personal morality and family roles: “Thus . . . all the activities of the individual are fundamentally religious activities, and there is no aspect of life which can be divorced from Dharma.”4 Closely related to dharma is the concept of karma, a causal law in which all moral or immoral acts and thoughts have consequences in the next life. Good karma leads to a good rebirth or release, and bad karma to a bad rebirth. Suffering can be explained in terms of past karma.

Another relevant term, ashrama, refers to four stages of life: student, householder, forest dweller (when one has grandsons), and ascetic or sannyasi. These stages were for the top three classes, known as the twice born because they were reborn when initiated at puberty with a sacred thread. Although few people follow this model, it is socially and psychologically important. Many elderly people metaphorically withdraw into the “forest” in their homes, detaching themselves from material and emotional concerns and preparing for death through prayer, scripture reading, and meditation.

Many Hindus belong to traditions devoted to one god (bhaktimarga), such as Krishna. They believe that those who are virtuous, with the grace of God, will go to heaven and obtain liberation from samsara—the cycle of birth and death. Such movements (sampradaya) have their own discipline, often involving monastic life and celibacy, and have specific teachings about the way to approach the end-of-life period. These movements include ISKCON (the International Society for Krishna Consciousness, Hare Krishna), Pushtrimarga, and Swaminarayan, which is very popular among Gujarati people. Followers of Swaminarayan, an avatar (incarnation) of Vishnu, will reside in their abode, Akshardham. Krishna’s heaven is Go-loka and Shiva’s heaven is Kailasha. A contemporary Guru, Sathya Sai Baba, has many followers in India and the west who see his teaching as universal. The Brahma Kumaris also have followers in the west. A reform movement, Arya Samaj, rejects the use of images (murtis) and Brahmanical rituals in favour of the Vedic fire ritual havan. After death, instead of the long and complex rituals, which last from 10 to 16 days or more, there is a havan after the disposal of the ashes on the third day.

Beliefs about life and death

Beliefs about life after death are derived both from the ancient Hindu texts and from popular belief. Most Hindus believe that there is a soul (atman) in all living beings, which transmigrates from one life to another, including animal life. In the Upanishads8 the soul within man is identified with ultimate reality, Brahman. Liberation (moksha) from birth and death can be
obtained through austerity and meditation, leading to mystical realisation of unity with Brahman in this life, and absorption into Brahman in the next. Brahma (the creator), Vishnu (the preserver), and Shiva (the destroyer), are three manifestations of the impersonal, neutral absolute Brahman. Vishnu and Shiva are sometimes identified with Brahman alone, and are at the centre of devotional sects.

In the Bhagavad Gita, the Supreme Lord is Krishna, incarnation (avatar) of Vishnu. He is creator of the universe, but also a personal God with whom one can have a relationship. Krishna assures the warrior, Arjuna, that he has a duty to fight his kinsman, but he cannot destroy the soul of those he has killed:

"Weapons do not cleave this self: fire does not burn him; waters do not make him wet; nor does the wind make him dry... For to the one that is born death is certain, and certain is birth for the one that has died."9

All action should be selfless, without any thought for its fruits (karma), motivated only by the love of God. Those who love God and think of Him at the time of death will come straight to him:

"And whoever, at the time of death, gives up his body and departs, thinking of Me alone, he comes to Me; of that there is no doubt."10

For many Hindus this passage is the key to a good death. A later work, Garuda Purana, describes elaborate rituals performed before and after death, still observed today, influencing popular beliefs about death and the afterlife."10 Yama, king of the dead, is a terrifying judge, Dharma Raja, who rules over a series of temporary hells.11 The dying are fetched by Yama’s servants, the Yamaduts.12 After cremation, the disembodied souls travel through various hells expiating their sins, depending on relatives’ offerings to release them. There are also temporary heavens. During the first 10 days the relatives have to create a new ethereal body.13 On the 12th day (symbolising 12 months), in a powerful ritual called sappindikarna14,15 the deceased becomes an ancestor or pitr. There is an apparent contradiction between the creation of an ancestor in this ritual, going back to Verdic times, and the belief in rebirth, which is related to different historical traditions. Until this time, the relatives are in severe ritual impurity and live austerely.

Good and bad deaths

A good death (su-mrtyu) occurs in old age, at the right astrological time, and in the right place (on the ground at home if it cannot be on the banks of the sacred Ganges). Hence Ganges water is always kept at home in a small container and offered to the dying person and placed on the lips of the corpse. A good death should be prepared for throughout life, and entered into consciously and willingly (iccha-mrtyu).17 All affairs should be set in order; unmarried daughters or granddaughters should have marriages arranged, conflicts should be resolved, and gifts of money and land should be made. People find it important to say goodbye to relatives and friends, and last words are highly treasured, becoming part of the way the tradition is passed on:

“My mother was like a saint and she died in just 5 minutes. She was 103 and she could put thread into needle, she walked without stick. She asked for bed on the floor. After that there was no one to give her a light—when people die we give a diva, like a candle, to show him or her a way to God. Then my sister’s son came. He said, ‘What’s happening, Bibi?’ She said, ‘O thank God you came. Come and give me diva on my hand.’ My sister started crying and she said, ‘Don’t cry, your tears will make a river for me to cross. I’m going to God. Let me go first. Don’t stop my way.’ He did everything, [then] she said, ‘Put my head in your lap, I want to go to God.’”(Panjabi woman)

A Brahmin priest may lead an act of penance. To ensure the dying person focuses on God, devotional hymns, bhajans, are sung; or “Ram Ram”, “Om”, or the Gayatri Mantra are chanted. Just before death the person is laid on the floor or ground, symbolising Mother Earth, with the head to the north,16,17 and Ganges water and a tulasi (basil) leaf are placed in the mouth.18 Signs of a good death are a shining forehead and a peaceful expression, with the eyes and mouth slightly open, indicating that the soul has left from these orifices. In holy individuals the soul leaves from the brahmaraundra, the fontanel at the top of the head.

Bad deaths (akal mṛtyu) are violent, premature, and uncontrolled deaths in the wrong place and at the wrong time, signified by vomit, faeces, urine, and an unpleasant expression.21,22 A Gujarati woman who died in the lavatory with a horrible expression on her face had a particularly bad death. The worst death is said to be suicide for selfish reasons. A good death needs the right rituals to see the soul on its way. For a son this is a sacred debt.” The failure to do so can cause a bad death because the soul fails to move on, haunting the relatives or causing bad luck, nightmares, illness, and infertility:

"An aunt was dying… the doctors told the family, and the whole family was present at the death. But when the doctors switched off the life support machine they wouldn’t let the family give Ganges water or perform any last rites. Today, after ten years it still affects the family. If they want to have a social occasion like a wedding in the family or something they must do some penances because she died without water, therefore her soul is still not free, and her family is not free. They’ve got to keep performing all these rites [for seven generations] that they weren’t able to during the death, until the soul is free… The doctors said that she would live a little longer, but there was no point, she was dying anyway. They switched off the machine, and they said they musn’t give her anything that would give her a shock and kill her straight away, that would choke her. (Gujarati couple, UK)."7
Disclosure and last rites
Although there is a strong tradition of being prepared for death, undergirded by a belief in continuity, unfortunately, not all deaths are those of elderly people who have fulfilled their life ambitions. This situation can create problems of disclosure and withdrawal of care. In practice, there is a tradition of non-disclosure, which creates a tension between autonomy and knowledge of the outlook to prepare for a good end, and being protected from the knowledge by relatives in case the person gives up hope and dies prematurely. Furthermore, modern medicine often provides hope, however unrealistic, that a cure is possible. As the psychiatrist Sudhir Kakar observed, the death of an individual, especially a parent, ruptures the extended family system. “[This] not only brings a sense of insecurity in a worldly social sense, it also means the loss of significant others who guarantee the sense of sameness and affirm the inner continuity of the self.”

The difficulty of open disclosure is shown by an example of a Gujarati man, Suresh, and his son, Ramesh, in the UK. The Hindu family practitioner warned Ramesh that his father was terminally ill with prostate cancer and tuberculosis. Ramesh resisted this idea, and did not want his father told his prognosis. Unfortunately, the hospital staff kept reassuring Ramesh that Suresh would recover, and his impending death was never discussed with Suresh. As often happens, he was clearly aware that he was dying, as he gave his books away, talked about dying, and obtained a gold chain for his granddaughter’s marriage, but he colluded with his son’s silence. When he died unexpectedly in his son’s absence, Ramesh was racked with guilt because he had not been present to give him the last rites or to say goodbye.

In India, the patient can be taken home, which implicitly discloses to the patient that death is imminent. The dying person can be placed on the floor and the appropriate rituals performed. In the diaspora this is more difficult, especially if the professionals are not aware of the importance of the family’s tradition.

Withdrawal of treatment and euthanasia
In a conscious and willed death the body is relinquished voluntarily. There has long been a tradition of voluntary death, and indeed of religious suicide in the Hindu community, which could owe its roots in part to the Jain tradition, in which spiritual adepts were encouraged to fast to the death. Such a self-willed death was “linked to a specific purpose: to obtain freedom (heaven or liberation) through an act of omnipotence involving the sacrifice of the self.” For a terminally ill person, fasting has several functions—as a spiritual purification, to promote detachment, and to ensure that there are no signs of bad death (faeces, vomit, or urine), Justice, who spent some time in a bhavan (home for the dying) in Varnasi, noted that: “Not eating or drinking during the time of dying can be considered an aspect of general detachment from the material world, a spiritual goal of classical Hinduism.” A dying person can refuse medication to die with a clear and unclouded mind, and view pain as a way of expurgating sin. This belief can cause problems for non-Asian professionals whose training makes them want to maintain life and relieve suffering.

A more difficult issue is the withdrawal of treatment from those unable to give consent, especially if the family resists. With the aunt described here, death was believed imminent because of her stage of life. However, there can be unfamiliarity with medical terminology and technology. For example, when a 3-year-old Panjabi child had been knocked down by a car, it took 3 days for the family to understand the idea of brain death and give permission for withdrawal of life support.

Euthanasia
There is a distinction between the willed death of a spiritually advanced person, and someone in great pain wishing to end an intolerable life. Suicide for selfish reasons is morally wrong and leads to hell. Such a death cannot have shraddha, the all-important post-mortem rites. However, some lawmakers made exceptions:

“If one who is very old (beyond 70), one who cannot observe the rules of bodily purification (owing to extreme weakness...) one who is so ill that no medical help can be given, kills himself by throwing himself from a precipice or into a fire or water, or by fasting... shraddha may be performed for him.”

Some recent authorities have argued that only God can take life, and human beings should not do it because of the karmic effect on the next life. Subramuniaswami states that “a lethal injection severs the astral silver cord connecting the astral body to the physical. Those involved then take on the remaining karmas of the patient.” Instead, suffering should be seen as purifying and cleansing. Crawford, however, suggests that Ayurvedic medicine “allows room for human efforts to curtail the effects of ordinary non-moral actions by the use of intelligence, wisdom, balanced conduct, and recourse to medicine. Only the fruits of immensely good or bad moral actions cannot be averted by these means.” As an enlightened person in the past was allowed to choose the time of his death, so it is morally permissible for a terminally ill person who is suffering greatly. Otherwise he may lose “the equanimity he so cherishes for his final moments of life. Euthanasia ensures a merciful death because he can leave this life with consciousness unclouded by the stropor of drugs, and without fear that some unexhausted karmas will plunge him back into mortal existence.”

A difficulty arises for those who are not spiritual adepts, but the arguments below might be adapted for such cases.

Hindu ethics on the whole come out strongly against involuntary euthanasia, because it contradicts the
principle of autonomy and can lead to abuse. However, the primary concern is the relief of suffering. Mahatma Gandhi advocated *ahimsa*, non-violence, which is utmost selflessness, refraining from any harmful act. Yet it is necessary to destroy some life to live at all, and sometimes one has to take life to protect others. Gandhi said that he could not abide the thought of allowing a rabid dog to die a painful and slow death, and he would hope there were more hopeful remedies for human beings in a similar situation. However:

“Should my child be attacked with rabies and there was no helpful remedy to relieve his agony, I should consider it my duty to take his life. Fatalism has its limits. We leave things to Fate after exhausting all remedies. One of the remedies and the final one to relieve the agony of a tortured child is to take his life.”

Thus there are different levels of *ahimsa*, wherein evil can be eclipsed by a compassionate act performed selflessly, because ultimately it is motivation that counts. Such an approach might also be applied to a severely disabled baby with a poor outlook. Crawford comments: “*Karma* does not give us the right to keep such people alive and in pain when all they want is a peaceful death. Their *karma* is our *dharma*. We have a duty to our fellow human beings. If they are suffering because of some sin, it is not less a sin to let them suffer. Mahatma Gandhi had said, ‘God comes to a hungry man in the form of a slice of bread!’ In what form does God come to a person begging to die?”

There is thus not a single moral position on the issues of involuntary euthanasia, and there is a longstanding tradition of voluntary suicide in certain carefully defined circumstances.

**Conclusion**

Generalisation about Hindu patients at the end of life is difficult because their beliefs and attitudes will depend so much on education, class, and religious tradition. However, the key factors are to find out the patient’s and family’s particular religious position, and to remember the importance of enabling him or her to have a good death with the help of family. They will need to be actively involved in care. The long-term importance of allowing the right death-bed rituals to take place, recognising the tension between allowing for patient autonomy and disclosure and the need for the family to make decisions on behalf of the patient should be acknowledged. The Hindu good death provides a valuable model for how death can be approached positively and without apprehension.

**References**

It was one of our first lectures on history and examination, and a number of professors were sat at the front. The speaker was telling us about the diagnostic mistakes we were going to make in our working lives. “You will miss a perf, you will miss a pulmonary embolism. Who here,” he indicated the front, “hasn’t missed a perf?” Not one of them budged.

Predictably, I determined that whatever else I did or didn’t do, I would not miss a perf.

5 years later as a senior house officer in general surgery, our ward round progressed to a 50-year-old man with toxic colitis, who had been on high dose steroids for 3 days.

The consultant examined him. I indicated that his biochemical parameters were unfavourable, but was told, “Sometimes it’s beyond biochemistry. Sometimes you just know by the way things look. He’ll be fine”.

The consultant’s words were very much in my mind when I was called back to see the patient during the night, after an acute exacerbation of pain. And by the time I got there, he had settled down. His vital signs had not changed. He told me about a sudden intense cramping pain he’d felt, indicating his left upper quadrant, which had gradually relaxed over the next 5 minutes. His abdomen, though tender, was soft, and there was no overt peritonitis.

Steroids mask the acute abdomen. It’s classic. I missed his perforation. His subtotal colectomy and ileostomy was delayed by 24 hours. He survived after a torrid time in the intensive care unit.

It’s a mixed message—I was wrong in both lecture theatre and the ward. Both situations involved authority, to which we often defer in medicine. We must respect experience (history is the only guide to the future). We must not forget that hindsight (a special form of experience) shows that experience may get it wrong. Finally, when situations change, we must retain our own judgment, especially at night.
In June, 2003, a 29-year-old woman had a routine abdominal ultrasound at 12 weeks' gestation, which showed a possible molar pregnancy. Dilatation and curettage (D&C) showed a partial hydatidiform mole (figure, A). Flow cytometry showed that the tissue was triploid. We monitored the patient’s β-human chorionic gonadotropin (βhCG), which returned to normal (<5 IU/L) in November, 2003. In April, 2004, her βhCG was increased, and an abdominal ultrasound showed a gestational sac. 1 month later, she developed vaginal bleeding; miscarriage was suspected and D&C showed decidua but no fetal tissue. The bleeding continued, her βhCG remained raised, and in June, 2004, a further D&C showed gestational trophoblastic disease. We reviewed the histological specimen, and diagnosed placental site trophoblastic tumour (PSTT) (figure, B). Her βhCG was 1924 IU/L, and a uterine doppler ultrasound showed a 105 mL uterus with a 4 cm vascularised mass extending into the left adnexa and ovary. Full-body CT showed bilateral pulmonary metastases. MRI of the brain showed no abnormalities. We started her on systemic chemotherapy in conjunction with intrathecal methotrexate as prophylaxis for occult central nervous system disease involvement. βhCG returned to normal after 11 weeks of treatment, and the left adnexal and thoracic lesions disappeared. She then had a hysterectomy and left sided salpingo-oophorectomy to debulk any residual tumour. Histopathology showed that the only viable tumour consisted of isolated cells in the left ovary (figure, C).

We genotyped the patient, her partner, and the histological specimens. There were four fully informative micro-satelitite polymorphisms, and the partial mole showed both maternally and paternally derived alleles. For two of these markers the molar tissue was trisomic with one maternal and two paternal alleles, and for the other two markers the ratio of paternal to maternal DNA was also consistent with trisomy from one maternal and two paternal contributions. Genotyping therefore confirmed a partial mole of dispermic origin. Analysis of the DNA from the subsequent PSTT also showed paternal alleles. Both maternal alleles were present for all markers in the DNA prepared from the trophoblastic tissue due to contamination of the tumour sample with infiltrating maternal cells. The proportion of maternal and paternal DNA showed that the PSTT originated in the previous partial hydatidiform mole. We gave the patient chemotherapy for a further 8 weeks; when last seen in May, 2005, she was healthy and had a normal serum concentration of βhCG.

Gestational trophoblastic neoplasia is a group of pregnancy-related disorders ranging from the pre-malignant complete hydatidiform mole and partial mole to the malignant invasive mole, choriocarcinoma, and PSTT. Complete moles occur in about 1 per 1000 pregnancies and partial moles in 3 per 1000 pregnancies.1,2 Although both secrete βhCG, complete moles are diploid and nearly always androgenetic in origin. By contrast, partial moles are triploid, consisting of one maternal and two paternal sets of chromosomes.3 After uterine evacuation, 16% of complete and 0·5% of partial moles undergo malignant change.4 For complete moles this malignant change includes invasive mole, choriocarcinoma, and PSTT. Partial moles can also change into choriocarcinoma but have never been shown to progress to PSTT. PSTT is usually curable if diagnosed within 4 years of the associated pregnancy.5 Delay in recognising malignant change after molar pregnancy can result in haemorrhage, uterine perforation, metastasis, and death. Since partial moles can become choriocarcinomas and PSTTs, and this malignant change can be effectively detected by simply monitoring serum concentrations of βhCG, patients should be followed up after partial molar pregnancy.

References

Figure: Photomicrographs of (A) the initial products of conception showing hydropic chorionic villi with irregular outlines and abnormal trophoblastic proliferation with a focally vacuolated phenotype (H&E ×40), (B) D&C specimen showing myometrial infiltration by sheets of monomorphic cells with focal vacuolation (H&E ×40), and (C) left ovarian section showing single infiltrating cells staining strongly (Mel-CAM × 200).