Kabul diary

Kamran Abbasi

When the charity Child Advocacy International invited the BMJ to experience aid work in war torn Kabul, I took the plunge.

Tuesday 3 March: Manchester, England

Aid work needs complex logistic support, and I'm greeted by workers from the children's charity Child Advocacy International pushing trolleys crammed with medical supplies. In two weeks the charity has accumulated $60 000 (£37 000) of donated medical equipment. The aim is to renovate the paediatric intensive care ward of the Indira Gandhi Children's Hospital in Kabul, the main paediatric hospital in Afghanistan, thus providing the impetus for the refurbishment of the whole hospital.

Marshalling the charity workers is David Southall, professor of paediatrics and founder of Child Advocacy International. We are accompanied by three aid workers from the Duchess of York's children's charity, Children in Crisis, which is working with Child Advocacy International in Afghanistan. David Sogan, projects manager for Children in Crisis, feels that the trip should go well, as the team has negotiated carefully with the British representative of the Taliban, the much criticised rulers of Afghanistan.

I wonder why these people risk their lives in war torn countries. As the plane takes off, I also wonder why I am risking mine too.

Wednesday 4 March: Islamabad, Pakistan

As many internal Pakistani flights are on small jets, with little room for our equipment, we hire a minibus—known locally as a flying coach—and head for Peshawar along the Great Trunk Road. We crawl rather than soar—heavy rains and a bald tyre—enduring five cramped hours on a muddy road. Arriving in Peshawar, Sogan takes his wife Judit, a barrister interested in international children's rights, and Linda Burbridge, a marketing manager whose bank has made donations to Children in Crisis, to the bazaar to attire them for Kabul. The Taliban have ordered that Afghan men must grow their beards and that Afghan women must wear full length body covering. Rules are laxer for foreigners, but it is prudent to wear loose fitting, flesh concealing clothing.

The Taliban were educated in the religious schools of Peshawar. As divisions widened within the mujahedin—who gained power after the Russian withdrawal—the Taliban swept through southern Afghanistan. Northern Afghanistan still remains outside the control of the Taliban and is divided up by various Afghan factions. The Taliban, who are notorious for their treatment of women, have received less foreign aid than necessary to provide basic sanitation, food supplies, health care, and education. Meanwhile, the supply of arms continues, and the war goes on.

Thursday 5 March: Peshawar, Pakistan

We are flying on a jet of the International Committee of the Red Cross. The only other flights to Afghanistan, apart from clandestine night flights, are with the United Nations and Ariana Afghan Airlines. Kabul airport has no radar, and the city nestles on a plateau in the Hindu Kush mountains. Bad weather and regular shelling by opposition forces make flying hazardous. We head west to Afghanistan, and beyond a break in the mountains lies a large plateau and the historic city of Kabul, a once great mountain kingdom. On this clear, cold day we are greeted by the sight of dead MiG fighter planes and bomb damaged buildings. Snow capped mountains ring the city, and the airport terminal is dark and deserted. We are the only visitors to arrive today. As we leave the airport, a sign warns us that there are 10 million land mines buried in Kabul.

News p 369

Kamran Abbasi, *BMJ*

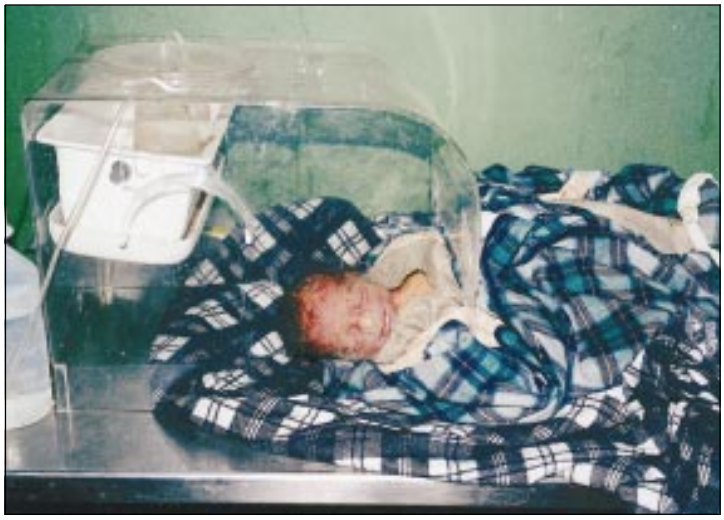
BMJ 1998;317:401–5

Full version appears on our website

The only ambulance is ready for service at the Indira Gandhi Children's Hospital

Time is short, so we hurry to the nearby Indira Gandhi Children's Hospital. It is in a pitiful state, unclean and with blocked toilets so that waste seeps into the corridors and wards. Most of the windows have been smashed, making it bitterly cold inside. The only heating is from old, wood burning stoves that spew as much smoke into the wards as they do outside. Children are crowded into barren wards, usually two to a bed. Medical equipment is in short supply, and the medicine cabinets are almost empty. Needles and syringes are reused, and, with no sharps boxes, even walking around is hazardous. Children are paying the price for the war inflicted on this poor country, and their hardship is being perpetuated by the impasse between the Taliban and foreign agencies, namely Unicef and the European Community, over issues of sexual equality and illegal drugs.

We manage to track down the chief surgeon, who suggests that, in the absence of the hospital president and vice president, we seek final permission for the project from the minister for public health. We rush to the ministry of public health, but the minister, Mullah Abbas, has left and we will have to wait a day to see him.



A baby in the neonatal unit

Friday 6 March 1998: Kabul, Afghanistan

Friday ("Jummah") is the holy day of the week. The Taliban have made Friday prayers mandatory for Muslims, and a register of attendance is kept. I'm surprised to learn that Kabul has a church, and intrigued that the priest has moved Sunday service to Fridays. Kabul also has a large Hindu community, and in the 1970s the Indian government had close links with Afghanistan, and a hand in the development of the children's hospital, subsequently renamed the Indira Gandhi. Those links have long since been severed.

Saturday 7 March: Kabul

I return to the ministry of health with the two Davids and our translator, Afzal, who also works for the ministry of education. The Taliban authorities' restriction on girls' education has made foreign agencies reluctant to provide the ministry of education with financial support. By limiting access to education for girls and restricting work opportunities for women teachers, the Taliban have alienated themselves from major funding bodies such as Unicef and the European Community. Afghan society is another example of a culture that absorbed Islam and fashioned it to its own ends. But, although rulers regularly distort religion to suppress women and minorities, the true tradition of Islam is one of equality.

On our arrival at the ministry of public health, we are immediately ushered in to meet Mullah Abbas, the minister of public health and a member of the Afghan Shura, the chief law making council of the country. The mullah, who seems to have softened his attitude towards foreign agencies according to David Sogan, is enthusiastic about our proposal, and, according to protocol, gives us a letter authorising our work. He also agrees to help release our equipment, soon to arrive at Kabul airport. Mullah Abbas highlights the lack of a cancer hospital, no pathology service, and a dearth of educational material as major deficiencies. All pathological samples are sent to Peshawar. At the back of his room is a short row of outdated medical books. "These are all we have in the entire ministry of public health," he sighs. Even the *BMJ*, we are told, has not arrived for six months. A lack of aid is making basic healthcare provision an impossibility. He adds, "Unicef say a lot but they are doing very little. They have not even given us a stethoscope."

Authorisation letter in hand, we return to the hospital, and park beside a bullet ridden green van—the only ambulance. Dr Hussain, the hospital vice president, has returned from Peshawar. Pleased that our work can go ahead, he takes us on a tour of the hospital. At the neonatal unit, a filthy blanket is draped over the entrance, hiding a small room with smashed windows and broken incubators. The incubators are unusable. A head box is linked to an ageing oxygen concentrator and is used by the six newborn babies in rotation. The babies are laid on benches and warmed by an electric heater. They look close to death. Shorn of our usual array of high tech medical gadgetry, we feel desperate and impotent and leave to arrange the labourers for the renovation work. Afzal is a trained engineer and is in charge of organising them.

By late afternoon, we are back in our residence, trying once again to set up an email link to the outside world. Suddenly, a loud explosion shakes the building.

Blocked toilets at the Indira Gandhi Children's Hospital

The windows shake but remain undamaged. Outside, we see the terrified faces of four small girls who had been playing across the road, smiling and waving at us. They are now frozen with fear. We learn later that a rocket with a 60 km range landed within a quarter of a mile of us.

Sunday 8 March: Kabul

International Women's Day. The only women in employment are working in hospitals as doctors, nurses, or cleaners, and some who are employed by foreign agencies. Only men work in the government ministries, shops, and schools.

We visit Tahia Maskan, the main orphanage in Kabul, which was built by the Russians. All the children huddle in the warmth of the renovated wing of the orphanage, which, for Kabul, is a comfortable and pleasant place to live. The World Food Programme, a UN organisation, supplies beans, lentils, and rice. Children in Crisis supplements that diet with fruit and vegetables. None the less, the diet is meagre, and many children have vitamin deficiencies. The neglected part of the orphanage is unclean, cold, and damp. The dining room is dark, with a wet floor and dirty tables. Water seeps in from the kitchen, where the tap drips continuously. The renovated section is in marked contrast. The children sleep here and have classes that are mixed sex for those aged under 5. Some of the teachers are women, and there are two women social workers. The girls are confined to the second floor, which has no toilet, so they either use the roof or sneak outside to the overflowing latrines.

On our way home, we manage to lay our hands on the *Kabul Times*, the official Taliban newspaper. An editorial is devoted to women's rights, and I am surprised to read: "The right to knowledge is a mutual right for males and females.... We are truly in need today of women doctors, teachers, and nurses.... Men and women are equal and no one has the right to discriminate between the two sexes, since God Almighty equalled between men and women in all forms of activity and all aspects of work, responsibility, and reward."

Monday 9 March: Kabul

While work continues at the Indira Gandhi Children's Hospital, we visit Malalai hospital, the only tertiary referral obstetric hospital in Afghanistan. We are greeted by Dr Mahmood Nawabi, the only male doctor who works on site here. Eighty usable beds cater for 35-40 deliveries a day. Few of the doctors can perform caesareans, and three male doctors who can, work from home. They operate, but do not see patients before or after procedures. With no on site paediatricians, all neonatal complications are sent to the Indira Gandhi hospital. Dr Nawabi laments: "We tried to get a female paediatrician but no one has been willing to help us. We will be happy to have a qualified paediatric doctor."

The state of the hospital is a testament to the dedication of Dr Nawabi and his team. He is paid 150 000 Afghanis a month ($5) and works a 1 in 4 on call rota. "You must help the staff as well," he pleads. "If there is no help for the doctors, then many more will leave the country." We leave Malalai hospital impressed, thinking that this is a well organised unit where modern equipment and incentives for doctors will have a huge impact.

I arrange a meeting with Unicef's project manager, Robert Biakcin, as I have been surprised by the level of criticism of the organisation's work in Kabul. An initially cordial meeting turns frosty when I ask Mr Biakcin why government ministers, hospital physicians, and aid workers feel that Unicef could be doing more for the health care of children in Afghanistan. "We do a lot but don't say very much. NGOs say a lot but don't do very much," he fumes.

Night flights are becoming more frequent, making it difficult to sleep, or it may just be that I'm more jittery since the rocket blast two days ago.

Tuesday 10 March: Kabul

Work at the hospital is going well, and Afzal's team of carpenters and painters will probably have finished before we leave. All that remains is the medical equipment we sent as freight on Ariana Afghan Airlines, which is arriving today. Brandishing our letter from Mullah Abbas, we try to persuade the customs staff to release our goods. Our hopes are dashed when we learn that we need authorisation from the ministry of finance and an official pharmacist to inspect our medical supplies.

At the ministry of finance we put our case to two ministers with matching beards and turbans—the

The legacy of war

Kabul is ringed by mountains

younger in black, the more senior in white. A local squats behind me on a window ledge as we discuss the release of our equipment. The ministers readily sign our form. Laughter breaks out as we leave, and I'm relieved when it is explained that the man perched behind me has been ticked off by Black Turban: "Don't sit like a donkey, what sort of impression does that give? These people are doctors."

Wednesday 11 March: Kabul

I'm woken by anti-aircraft fire, which seems near to me, but David Sogan reckons it's at least 10 km away. Rumours suggest that a massive spring offensive is around the corner, as both sides try to inch forward. For now, however, the Taliban are secure in Kabul, and I'm reasonably secure in the knowledge that our flight out of Afghanistan is booked for tomorrow.

Afzal's team should complete the paediatric intensive care ward today. First, I have a meeting with the WHO representative in Kabul, Dr Abdi Momin Ahmed. Arriving at the WHO's new headquarters, I'm amazed by the opulence of the building, a shining marble palace in a city of dilapidated structures. Even as I enter, more marble is being unloaded from trucks. I ask him about the WHO's justification for such plush accommodation in this poor city. Dr Ahmed looks flabbergasted, telling me that the marble is being imported by the owner of the property and that the WHO is renting the house for $1000 a month, which he considers reasonable.

By midday, we gather at the hospital to see the fruition of our project. The paediatric intensive care ward has been washed, sterilised, and painted. A wall has been knocked through to increase space, the toilets have been unblocked, and windows have been repaired. All the beds have been painted and fitted with lights, new mattresses, and blankets. Our medical equipment has finally reached its destination and covers the doctors' area in the centre of the room. A new medicine cabinet stands by the door, brim full of drugs. Afzal's beard can't hide a self satisfied grin.

The medical staff of the hospital are stunned by the transformation. The doctors, mostly women, assemble for training on how to use the new equipment, which we manage to complete before the official opening of the revitalised ward by Dr Abdul Bashir Hassan, deputy minister for public health. Naturally, leaving such expensive equipment is a risk—it may be misused or confiscated by the government for use elsewhere. But medical appliances donated on previous visits are still being used as intended, and, as we leave, Dr Hussain locks the door to the intensive care ward and, reassuringly, begins itemising the equipment.

Thursday 12 March: Kabul

Mission accomplished, we're ready for home. We've set up a modern paediatric intensive care ward that can be used as an example to obtain more funding for the hospital. We cram into three Land Cruisers, and a jostle for the better seats leaves Southall and myself on the bench seats at the back of the smallest vehicle. Our UN convoy heads east out of Kabul, rushing through Taliban checkpoints, where less privileged travellers are being frisked. Soon we are racing down the Kabul gorge, sandwiched between sheer, barren rocks and hurtling white waters. The gentle undulations of the road become a rollercoaster, and, as our trade-off for leg room against head room misfires, we are left holding on for dear life.

The UN drivers don't bother with convoy rules, and, with our colleagues out of sight, we are marooned by a puncture. While we wait for the tyre to be repaired, a crowd quickly gathers, and a slight young man with a pointy moustache picks up a weighty rock and walks towards us. We are perturbed when he raises the rock above his head, but his next move is to throw it away from our vehicle and then draw a line in the ground with his toe. In the absence of Geoff Capes, Southall takes up the challenge of a shot putting contest, which delights the crowd. We are warmly waved off to resume our rollercoaster ride.

Approaching Jalalabad, we see palm trees and camels, and the environment is more typical of a South Asian city, much hotter and more dusty. A Danish crew fly us straight to Islamabad. A good night's sleep lies between us and our flight home to a different world.

On reflection

Afghanistan is a complex country, comprising many poor but fiercely independent races, who tend to distrust each other rather than cooperate. The delivery of health care is similarly blighted: numerous underfunded agencies with differing goals are striving to maintain one of the most basic healthcare systems on the planet. The breakdown in dialogue between the Taliban and leading foreign donors, namely the European Community and Unicef, continues to cripple healthcare provision.

Although recent months have seen the Taliban further restrict women's movements—foreign agencies were banned from employing local women—it is difficult not to feel that the donor community's heavy handed approach to the issue of women's rights has provoked further repression of women rather than liberating them. The Taliban's reluctance to deal firmly with poppy farmers, whose wares eventually supply drug addicts in the West, has raised another barrier. The Taliban can no longer be wished away: after a brief, but ultimately doomed, attempt at a negotiated peace, divisions within the Northern Alliance have allowed the Taliban to secure their hold on the southern two

thirds of Afghanistan and bolstered their ambitions for the rest of the country.

In Kabul, the Indira Gandhi Children's Hospital is struggling on, desperately short of funds. The renovation of the intensive care ward was intended to encourage a grant from the Department for International Development, Unicef, or the European Community Humanitarian Office, to upgrade the rest of the hospital. Instead, the charities' proposal was not supported. To make matters worse, Jack Bell, field worker in Kabul for Children in Crisis, was left to handle the fury of other non-governmental organisations, which felt that our intervention was inappropriate. "Just a lot of envy, I believe," he said. "ICU is going fine. We're still doing the hallway and the other rooms, but the one room is full—two babies per bed." It's nice to know that some traditions don't change. But, reassuringly, on a return visit to Kabul in June, David Southall noted that other non-governmental organisations were refurbishing other wards.

In July, the situation for aid agencies in Kabul took a turn for the worse when the Taliban ordered all non-governmental organisations to move to the old polytechnic building in Kabul by 19 July. Aside from the potential security risk of being confined in the same area, the polytechnic building has no electricity or running water and is within firing range of the front line. UN organisations like the WHO and Unicef were exempt from this edict, as was the International Committee of the Red Cross. Not surprisingly, all but one of the non-governmental organisations decided to leave Kabul rather than face the consequences of not complying with the Taliban's order. Child Advocacy International and Children in Crisis are currently negotiating with the Taliban for a fairer arrangement, so that their aid work can resume.

While the political power struggle grinds on, we can only hope that a few of the countless disadvantaged children in Afghanistan will benefit from the changes that were initiated at its main paediatric hospital.

Getting research findings into practice

Decision analysis and the implementation of research findings

R J Lilford, S G Pauker, D A Braunholtz, Jiri Chard

This is the sixth in a series of eight articles analysing the gap between research and practice

Department of Public Health and Epidemiology, University of Birmingham, Birmingham B15 2TT
R J Lilford, *professor of health services research*
D A Braunholtz, *senior research fellow*

Department of Medicine, New England Medical Centre, Tufts University, 750 Washington Street, Boston, MA 02111, USA
S G Pauker, *professor*

continued over

BMJ 1998;317:405–9

Summary points

Decision analysis reconciles evidence based medicine with patients' preferences

Decision analysis uses Bayesian probabilities together with values assigned to different outcomes to determine the best course of action

Although it is currently unrealistic to do a separate decision analysis for each patient, computer programs may soon overcome this problem

In the meantime, decision analysis can be used to provide guidelines for managing groups of patients with similar clinical features

Calculating specimen decision analyses can be helpful for patients with different values

Evidence based medicine is more than just reading the results of research and applying those results to patients because patients have particular features that may make them different from the "average" patient studied in a clinical trial.[1] There are two types of differences. The first type of differences comprise those that affect probability (for example, the probability that treatments will have the same absolute or relative effects as those measured in the trial). The second type of differences comprise those values (or utilities) that affect how much of a side effect a person is prepared to trade off against the positive advantages of treatment.

Thus it is necessary for doctors to relate the results from a trial to their particular patient. Health professionals usually do this intuitively, but formal decision analysis provides an intellectual framework for developing an explicit decision making algorithm which can be criticised and improved. Although, currently, time constraints make it unrealistic to conduct a separate decision analysis for each patient, computer programs may soon help overcome this problem. It is, however, feasible for decision analyses to be done for categories of patients with similar clinical features and personal utilities. The results of such generic decision analyses provide a sound basis for developing clinical guidelines. Decision analysis thus provides a rational means of allowing health professionals to move from finding evidence to implementing it.

An example of decision analysis

Decision analysis is described in detail elsewhere,[2–5] but we will illustrate it with an example. Megatrials have shown that thrombolytic drugs save lives in cases of suspected myocardial infarction.[6] However, these drugs can cause stroke, which may leave the patient severely incapacitated. Also, there is a choice of drugs; the genetically engineered accelerated tissue plasminogen activator seems more effective than streptokinase in preventing death from myocardial infarction, but it has a higher chance of causing stroke. The risk of causing a stroke does not depend on when treatment is given. However, the probability of preventing death from

MRC Health Services Research Collaboration, University of Bristol, Bristol BS8 2PR
Jiri Chard, *research associate*
Correspondence to: Professor Lilford rlilford@doh.gov.uk

Series editors: Andrew Haines and Anna Donald

myocardial infarction does depend on how soon treatment begins after the onset of symptoms and on whether the patient has actually had a myocardial infarction and on their risk of dying if they have had one. To further complicate the picture, the relative advantage of tissue plasminogen activator over streptokinase in preventing cardiac death dissipates after about six hours, and thrombolytic drugs can cause other complications (such as haemorrhage and anaphylaxis).

How can the clinician account for these factors? We base our example on the work of Kellett and Clarke, who did a systematic review and then modelled all of these variables using decision analysis (fig 1).[7] Probabilities of the main outcomes according to the treatment given are presented in table 1. Specimen utilities are used for the various outcomes, a value of 1 for healthy survival and 0 for death. About half of patients who have a stroke in these circumstances will survive, but often with some degree of impairment; the mean utility of existence after stroke is 0.5.[8] The results of running the base case model (that is, for a 55 year old man with chest pain of recent onset and classical changes on electrocardiography, typical of participants in trials of thrombolysis) are shown in table 2. Clearly, there is much expected utility to be gained by using thrombolytic drugs, and, moreover, tissue plasminogen activator is the drug of choice. Even if we assume a passionate desire to avoid the disability associated with stroke, giving it a utility of −1 (for example, a healthy person who would equate a 20% risk of death with a 10% risk of stroke), the thrombolytic treatments remain optimal (data not shown). However, we get very different results as we move away from the base case. For example, chest pain in a 55 year old man with a normal electrocardiogram is associated with only a 17% risk of myocardial infarction, and thrombolytic drugs would lower the expected utility in these circumstances. The same man with a normal ST wave but an abnormal T wave has about a 24% risk of myocardial infarction; thrombolysis is advantageous, but only slightly and it would be disadvantageous if he was younger (his risk of dying if he had a myocardial infarction would drop to 5% at age 45), if he presented late (after 6 hours), or if he was particularly strongly averse to residual morbidity from stroke.

IAN BARRACLOUGH

Key definitions in decision analysis

Utility
A utility represents a patient's preference for one outcome over others. A utility is given a numerical value which is then used in the decision analysis. Utilities (or values) are quantified on a scale (usually from 0 to 1) that allows meaningful comparison between alternative outcomes.

Probability
Decision analysis is based on bayesian statistics. Probabilities of clinical outcomes are the doctor's best guess, based on indirect evidence such as laboratory studies, updated with data from relevant clinical studies by Bayes's theorem.

Decision tree
The various decision points and their consequences are mapped (with their associated probabilities and utilities) to form a decision tree—a visual representation of the decisions available. The aim of decision analysis is the logical reduction of a decision process into its individual decision points. Probabilities and utilities are assembled logically to determine the optimum treatment (that is, the treatment with the highest expected utility).

Sensitivity analysis
Sensitivity analysis is used to determine the robustness of a choice made using decision analysis. By varying the utilities and outcome probabilities it is possible to see how easily a decision would change—that is, how sensitive it is. This makes it possible to produce guidelines so that treatment can be tailored to groups of patients with similar characteristics; the final choice is, of course, the patient's.

Probability

Effective care in general, and decision analysis in particular, are underpinned by probabilities. Much of clinical research is concerned with providing the necessary probabilistic information.

Epidemiological studies provide baseline estimates of risk (for example, the risk of death from myocardial infarction) and data for revising these risks on the basis of test results (here test denotes any information about a patient, not only results from a laboratory). The ratio of the probability of an observed test result if the patient is or is not affected is known as the likelihood ratio. Given the patient's prior odds of being affected (odds are simply a ratio of probabilities) and the likelihood ratio for the observed test result, the revised (posterior) odds, and hence the probability that the patient is affected by a condition, can be calculated by multiplying the prior odds by the likelihood ratio. For example, a young woman whose only brother has Duchenne's muscular dystrophy has prior odds of being a carrier of the disease of 1:2. A raised concentration of creatine kinase (likelihood ratio 28) gives posterior odds of 14:1 (a probability of 93%). Likelihood ratios of multiple tests can be multiplied together, provided they are relatively independent.

Interventional studies such as clinical trials provide data on the effects of treatments. These studies can give two kinds of probability: conventional and bayesian. Conventional (frequentist) statistical analyses give P values and confidence intervals based on the probability of seeing the observed result (or a more

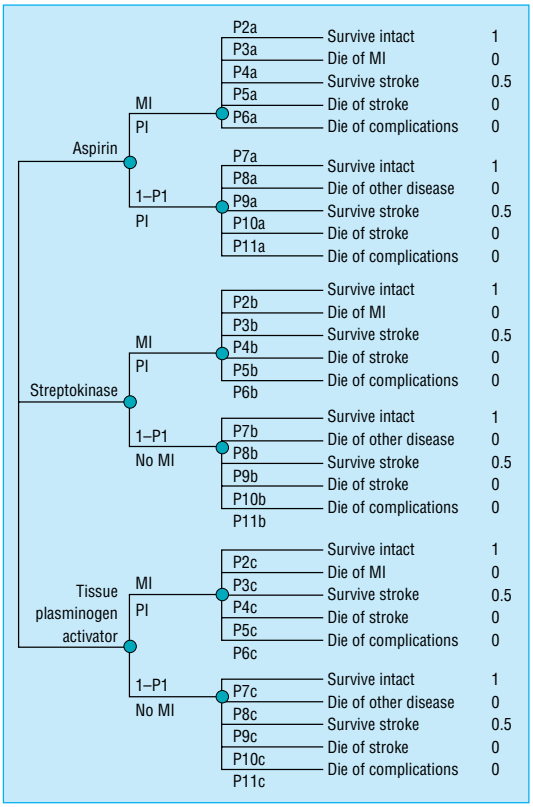


Fig 1 Decision tree for choice of treatment in case of suspected myocardial infarction (MI) of aspirin only *v* aspirin and streptokinase *v* aspirin and tissue plasminogen activator. P stands for probability. For example, P2a is the probability of surviving without impairment for someone who has had an MI and who has been treated with aspirin only. The formulas for expected utilities are aspirin, P1{(1*P2a) + (0*P3a) + (0.5*P4a) + (0*P5a) + (0*P6a)} + (1–P1){(1*P7a) + (0*P8a) + (0.5*P9a) + (0*P10a) + (0*P11a)}; streptokinase, P1{(1*P2b) + (0*P3b) + (0.5*P4b) + (0*P5b) + (0*P6b)} + (1–P1){(1*P7b) + (0*P8b) + (0.5*P9b) + (0*P10b) + (0*P11b)}; tissue plasminogen activator, P1{(1*P2c) + (0*P3c) + (0.5*P4c) + (0*P5c) + (0*P6c)} + (1–P1){(1*P7c) + (0*P8c) + (0.5*P9c) + (0*P10c) + (0*P11c)}. Figures on the far right are values

Table 1 Probabilities of various events occurring after suspected myocardial infarction according to treatment received by a typical patient in the trials and according to whether myocardial infarction has occurred. Probabilities are expressed as percentages

	Treatment		
Probability of outcome	**Aspirin**	**Streptokinase**	**Tissue plasminogen activator**
Given myocardial infarction:			
Dying of myocardial infarction	11.5	11.5×0.75=8.6*	8.6×0.8=6.9*
Surviving cerebrovascular accident	0.2	0.5	0.7
Dying of cerebrovascular accident	0.2	0.5	0.7
Dying of haemorrhage or anaphylaxis	0	0.2	0.18
Surviving without stroke†	88.1	90.2	91.5
Given no myocardial infarction:			
Dying of another cause	2.0	2.0	2.0
Surviving stroke	0	0.4	0.6
Dying of stroke	0	0.4	0.6
Dying of complications	0	0.08	0.06
Surviving without impairment†	98	97.1	96.7

*Streptokinase reduces the risk of dying by 25 percentage points when compared with aspirin alone; tissue plasminogen activator may reduce the risk of dying by a further 20 percentage points.[7]
†Calculated as 100 minus the sum of the four probabilities above.

extreme result) given a particular state of the world, typically that different treatments are equally effective. However, decision analysis, and bedside decisions generally, requires not the probability of already observed results given some assumed particular treatment effect, but rather the posterior probabilities of particular differences in the effects, given the observed data.[9]

If a trial comparing treatments X and Y finds an improvement in survival of 10 percentage points with treatment Y, a patient who is similar in relevant characteristics to those in the trial does not want to know that this observed improvement had only a 2.5% chance of occurring if the treatments are equivalent. The patient needs to know, for example, what the probability is that survival with treatment Y really is better than with treatment X. Probabilities that describe beliefs about the size of true effects are known as bayesian, and their calculation requires that a prior belief, expressed as a probability distribution, is updated according to the results of research. Because bayesian probabilities

Table 2 Relative expected utilities of aspirin *v* aspirin and streptokinase *v* aspirin and tissue plasminogen activator determined for different probabilities of myocardial infarction and death given myocardial infarction. Probabilities are expressed as percentages

Probability of myocardial infarction	**Probability of death given myocardial infarction***	**% Improvement with tissue plasminogen activator compared with streptokinase†**	**% Improvement with streptokinase compared with aspirin**	**Streptokinase *v* aspirin‡**	**Tissue plasminogen activator *v* aspirin‡**	**Guidelines**
17	11.5	20	25	–0.00186	–0.00173	At low probabilities of myocardial infarction treat with aspirin only.
17	5.0	20	25	–0.00463	–0.00615	
17	11.5	0	15	–0.00382	–0.0662	
24	11.5	20	25	+0.00017	+0.00152	At moderate probabilities of myocardial infarction treat with thrombolytic drugs only if patient presents in first 6 hours after onset and if risk of death given myocardial infarction is in the moderate to severe category. Treat with tissue plasminogen activator.
24	5.0	20	25	–0.00373	–0.00473	
24	11.5	0	15	–0.00259	–0.00539	
90§	11.5	20	25	+0.01935	+0.03207	At high probabilities of myocardial infarction treat with thrombolytic drugs and use tissue plasminogen activator only if history of symptoms <6 hours. Even at these high probabilities the benefits disappear if survival with stroke is given a value of –1 and the prognosis for survival is high (95%) or delay is considerable (data not shown).
90	5.0	20	25	+0.00472	+0.00867	
90	11.5	0	15	+0.00899	+0.00619	

*Probability increases with age.
†These figures are dependent on duration of symptoms.
‡Negative values indicate that treatment with aspirin only is preferable.
§Base case (person typical of participants in the trial).

relate to the probability of the true state of the world they are the rational basis for the implementation of a finding and for calculating parameters, such as the number of patients who must receive a new treatment to help (or harm) one patient. When studies are very large, as in trials of thrombolytic drugs, the two types of statistics will give similar results.

In our example of a myocardial infarction we considered two kinds of patient variables. We first considered features that affect absolute risk, such as age, but not relative treatment effects, such as the 25% improvement in cardiac mortality that occurs over a wide range of ages when patients are treated with thrombolytic drugs. The second type of patient variable has an influence on relative treatment effects (for example, the duration of symptoms affects the effectiveness of thrombolysis). Of course, if trials were infinitely large we could look up the precise relative treatment effect for any given category of patient. However, even when overall effects are measured precisely the effects in subgroups of patients (strata) are typically imprecise.

Should clinicians take the overall effect and apply it to the subgroup, or should they use the imprecise measurement made in the subgroups? For example, the second international study of infarct survival (ISIS 2) trial of thrombolytic drugs was analysed in subgroups. Unsurprisingly, this showed a null effect for people who had had their pain for a long time, but it also unexpectedly found a null effect for those born under the star sign Gemini. On what basis can we believe one analysis of a subgroup and not the other? In a bayesian analysis of subgroups we must state our prior beliefs for how the effect in the subgroup may relate to the effect in the remainder of the group.[10 11] This prior belief would be that there is little or no difference between Geminis and non-Geminis. The observed difference will therefore fail to shift our prior belief, and our posterior belief will remain that Geminis and non-Geminis benefit similarly. Our prior belief in the difference between patients with prolonged pain and others would be less precise and would reflect our belief that those with prolonged pain will benefit less than patients with a shorter duration of pain; for example, from our knowledge of drugs and infarcts, we expect the benefits to be largest when these drugs are administered quickly. In this case the data reinforce our prior belief and enable us to be more precise about how benefit is reduced as delay increases.

Values (utilities)

The great strength of decision analysis is that it is based not just on probabilities but also on the value placed on various outcomes. It therefore represents a method for synthesising both medical facts (probabilities) and human values (utilities), which together determine the best course of action—that is, the course that stands to maximise expected utility.[12] Decision analysis reconciles evidence based medicine with patients' preferences.

There is debate about the best way to obtain these utilities. Utilities imply a trade off: the extent to which the disadvantages of one outcome can be offset by the advantages of another. For example, for certain patients with cancer of the larynx survival is better for those who have radical surgery than for those treated with radiotherapy. However, radical surgery limits the ability to speak, at least in the short term. There is then a trade off between survival (maximised by surgery) and the ability to communicate (which is retained to a much better degree with radiotherapy). If a patient would run a 10% chance of dying to avoid losing the power of speech then the patient values life with this impediment at 0.9 on a scale from 1.0 (healthy life) to 0 (death). The subject of utilities and how they can be elicited is discussed in more detail elsewhere.[13–16]

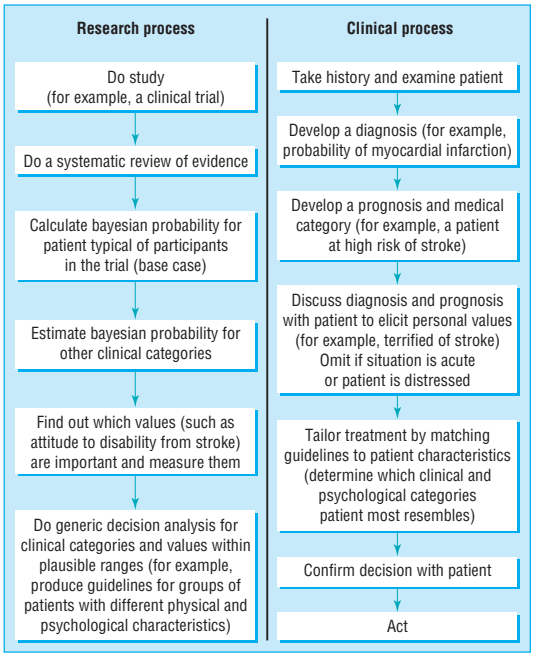


Fig 2 The sequence of events followed in performing a decision analysis, developing a clinical guideline based on the analysis, and implementing the guideline

Sensitivity analysis, generic decision analysis, and the individual patient

When consulting with individual patients it is important to elicit their personal values or at least to get a sense of them. However, it is not essential to redo an analysis for every patient in a busy clinic. Decision analysis may also be done outside the consulting room using a selection of different probability and utility figures within a reasonable range; this is known as sensitivity analysis. We used this technique to see how the expected utility of thrombolytic drugs might vary according to a patient's medical and psychological characteristics to produce the guidelines in table 2. The sequence of events followed in performing a decision analysis, developing a guideline based on that analysis, and implementing the guideline are shown in figure 2.

Information on short term outcomes is often available from clinical trials but long term outcomes must be derived from observational studies. Since long term outcomes are often more important to the patient and third party payer, these should be modelled by decision analysis. For example, modelling was required to extrapolate the results of a trial evaluating the short term effects of different types of angioplasty from the information collected in the trial.[17] Decision analysis is also useful when a clinical problem requires input from more than one set of study results; the effects of

hormone replacement therapy have been analysed in many different studies, each concerned with different outcomes and with each outcome being valued differently.[18] Furthermore, observational studies have shown that women have different risks at baseline (for example, thin women are at higher risk of fractures). Decision analysis has shown how these factors may be integrated to optimise individual treatment.[19]

Decision analysis is used to determine how to maximise an individual's expected utilities. By obtaining the median values of utilities from a large number of people the methodology can also be used to derive expected utilities for the community. When the costs of various options are included this is called a cost utility analysis. However, using decision analysis to make decisions for groups of patients creates some thorny ethical issues, especially when there is a conflict between maximising utility and maximising equity.

Conclusions

Decision analysis depends on probabilities and values, neither of which can be measured with certainty. These problems are not lessened when health professionals approach them intuitively; decision analysis makes these uncertainties explicit. The attempt to make complex decisions intuitively inevitably results in gross oversimplifications because it is impossible to incorporate and consider several components of a decision simultaneously. There is a large amount of empirical literature on the limitations of intuitive reasoning that is summarised by Dawes et al.[20]

Most research findings are applied unsystematically and intuitively. If evidence based medicine is to be seen through to its logical conclusion and if both empirical evidence and human values are to be incorporated into decision making, then this duality (the explicit collection of data *v* its implicit use) must be addressed.[21 22]

1 Glasziou PP, Irwig LM. An evidence based approach to individualising treatment. *BMJ* 1995;311:1356-9.
2 Keeney RL, Raiffa H. *Decisions with multiple objectives*. London: Wiley, 1976.
3 Weinstein M, Fineberg HV. *Clinical decision analysis*. London: Saunders, 1980.
4 French S. *Readings in decision analysis*. London: Chapman and Hall, 1989.
5 Sox HC, Blatt MA, Higgins MC, Marton KI. *Medical decision making*. Boston: Butterworth Heinemann, 1988.
6 Second International Study of Infarct Survival (ISIS 2) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected myocardial infarction: ISIS 2. *Lancet* 1988;ii:349-60.
7 Kellett J, Clarke J. Comparison of accelerated tissue plasminogen activator with streptokinase for treatment of suspected myocardial infarction. *Med Decis Making* 1995;15:297-310.
8 Tsevat J, Goldman L, Lamas GA. Functional status versus utilities in survivors of myocardial infarction. *Med Care* 1991;29:1153-9.
9 Lilford RJ, Braunholtz D. The statistical basis of public policy: a paradigm shift is overdue. *BMJ* 1996;313:603-7.
10 Donner A. A bayesian approach to the interpretation of sub-group results in clinical trials. *J Chronic Dis* 1982;34:429-35.
11 Oxman AD, Guyatt GH. A consumer's guide to sub-group analyses. *Ann Intern Med* 1992;116:78-84.
12 Swales J. Science in a health service. *Lancet* 1997;349:1319-21.
13 Lilford RJ. Trade-off between gestational age and miscarriage risk of prenatal testing: does it vary according to genetic risk? *Lancet* 1990;336:1303-5.
14 Johnson N, Lilford R, Jones SE, McKenzie L, Billingsley P, Songane FF. Decision analysis to calculate optimum treatment for microinvasive cervical cancer. *Br J Cancer* 1992;65:717-22.
15 Lilford RJ, Jackson J. Equipoise and the ethics of randomisation. *J R Soc Med* 1995;88:552-9.
16 Pauker SP, Pauker SG. Prenatal diagnosis: a directive approach to genetic counselling using decision analysis. *Yale J Biol Med* 1977;50:275-89.
17 Sculpher M, Michaels J, McKenna M, Minor J. A cost-utility analysis of laser-assisted angioplasty for peripheral arterial occlusions. *Int J Technol Assess Health Care* 1996;12:104-25.
18 Johnson N, Lilford RJ, Mayers D, Johnson GG, Johnson JM. Do healthy asymptomatic post menopausal women want routine cyclical hormone replacement? A utility analysis. *J Obstet Gynecol* 1994;14:35-9.
19 Col NF, Eckman MH, Kara RH, Pauker SG, Goldberg RJ, Ross EM, et al. Patient-specific decisions about hormone replacement therapy in postmenopausal women. *JAMA* 1991;277:1140-7.
20 Dawes R, Faust D, Meehl PE. Clinical versus actuarial judgement. *Science* 1989;243:1668-74.
21 Dowie J. "Evidence-based," "cost-effective" and "preference-driven" medicine: decision analysis based medical decision making is the pre-requisite. *J Health Serv Res Policy* 1996;1:104-13.
22 Dowie J. The research-practice gap and the role of decision analysis in closing it. *Health Care Analysis* 1996;4:5-18.

The articles in this series are adapted from *Getting Research Findings into Practice*, edited by Andrew Haines and Anna Donald, and published by BMJ Books.

Statistics Notes

Generalisation and extrapolation

Douglas G Altman, J Martin Bland

All medical research is carried out on selected individuals, although the selection criteria are not always clear. The usefulness of research lies primarily in the generalisation of the findings rather than in the information gained about those particular individuals. We study the patients in a trial not to find out anything about them but to predict what might happen to future patients given these treatments.

A recent randomised trial showed no benefit of fine needle aspiration over expectant management in women with simple ovarian cysts.[1] The clinical question is whether the results can be deemed to apply to a given patient. For most conditions it is widely accepted that a finding like this validly predicts the effect of treatment in other hospitals and in other countries. It would not, however, be safe to make predictions about patients with another condition, such as a breast lump. In between these extremes lie some cases where generalisability is less clear.

For example, when trials showed the benefits of β blockers after myocardial infarction the studies had been carried out on middle aged men. Could the findings reasonably be extrapolated to women, or to older men? It is probably rare that treatment effectiveness truly varies by sex, and claims of this kind often arise from faulty subgroup analysis.[2] Age too rarely seems to affect the benefit of a treatment, but clinical characteristics certainly do. Treatments that work in mild disease may not be equally effective in patients with severe disease, or vice versa. Likewise the mode of delivery—for example, oral versus subcutaneous—or dose may affect treatment benefit. Clinical variation is likely to affect the size of beneft of a treatment, not whether any benefit exists.

ICRF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, Oxford OX3 7LF

Douglas G Altman, *head*

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE

J Martin Bland, *professor of medical statistics*

Correspondence to: Mr Altman.

BMJ 1998;317:409–10

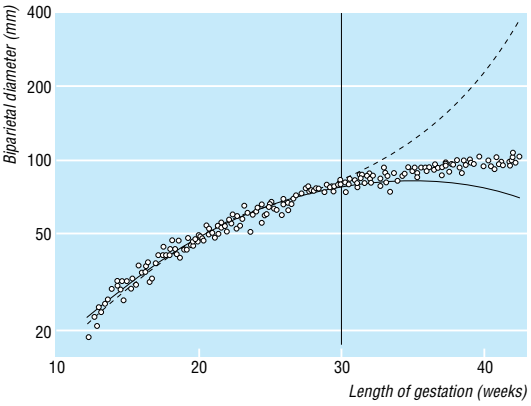
The extent to which it is wise or safe to generalise must be judged in individual circumstances, and there may not be a consensus. Arguably many studies (especially randomised controlled trials) use over-restrictive inclusion criteria, so that the degree of safe generalisability is reduced.[3] Even geographical generalisation may sometimes be unwarranted. For example, BCG vaccination against tuberculosis is much less effective in India than in Europe, probably because of greater exposure in India.[4] For the clinician treating a patient the question can be expressed as: "Is my patient so different from those in the trial that its results cannot help me make my treatment decision?"[5]

In a clinical trial we are interested in the difference in effectiveness between two treatments. There is no need to generalise the success rate of a particular treatment. In some other types of research, such as surveys to establish prevalence and prognostic or diagnostic studies, we may be trying to estimate a single population value rather than the difference between two of them. Here generalisation may be less safe. For example, the prevalence of many diseases varies across social and geographical groups. Results may not even hold up across time. For example, changes in case mix over time can affect the properties of a diagnostic test.[6]

Many studies use regression analysis to derive a model for predicting an outcome from one or more explanatory variables. The model, represented by an equation, is strictly valid only within the range of the observed data on the explanatory variable(s). When a measurement is included in the regression model it is possible to make predictions for patients outside the range of the original data (perhaps inadvertently). This numerical form of generalisation is called extrapolation. It can be seriously misleading.

To take an extreme example, a linear relation was found between ear size and age in men aged 30 to 93, with ear length (in mm) estimated as $55.9 + 0.22 \times$ age in years.[7] The value of 55.9 corresponds to an age of zero. A baby with ears 5.6 cm long would look like Dumbo.

Extrapolating may be especially dangerous when a curved relation is found. Figure 1 shows fetal biparietal diameter (on a log scale) in relation to gestational age. Also shown are quadratic and cubic models fitted to the log biparietal diameter measurements from only those fetuses less than 30 weeks' gestation. Both curves fit the data well up to 30 weeks, but both give highly misleading predictions thereafter. The quadratic model shows a spurious maximum at around 34 weeks, while the cubic curve takes us again into elephantine regions.



Fetal biparietal diameter (on log scale) in relation to gestational age[8] with quadratic (solid line) and cubic (broken line) regression models fitted to data from only those fetuses less than 30 weeks' gestation (n=119)

When we have two explanatory variables it will not usually be apparent (unless we examine a scatter diagram) when a patient has a combination of characteristics which do not fall within the span of the original data set. With more than two variables, such as in many prognostic models, it is not possible to be sure that the original data included any patients with the combination of values of a new patient. Nevertheless, it is reasonable to use such models to make predictions for patients whose important characteristics are within the range in the original data.

Clearly patient characteristics, including the criteria for sample selection, need to be fully reported in medical papers. Yet such basic information is not always provided.

1 Zanetta G, Lissoni A, Torri V, Dalle Valle C, Trio D, Rangoni G, Mangioni C. Role of puncture and aspiration in expectant management of simple ovarian cysts: a randomised study. *BMJ* 1996;31:1110-3.
2 Matthews JNS, Altman DG. Interaction 1: Heterogeneity of effects. *BMJ* 1996;313:486.
3 Ellenberg JH. Selection bias in observational and experimental studies. *Statistics in Medicine* 1994;13:557-67.
4 Anonymous. BCG: bad news from India. *Lancet* 1980;i:73-4.
5 Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine. How to practice and teach EBM.* London: Churchill-Livingstone, 1997:167.
6 Begg CB. Biases in the assessment of diagnostic tests. *Statistics in Medicine* 1987;6:411-23.
7 Heathcote JA. Why do old men have big ears? *BMJ* 1995;311:1668.
8 Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 2. Head measurements. *Br J Obstet Gynaecol* 1994;101:35-43.

A memorable meeting

A view from the man in the seat opposite

The train journey, the pile of papers, and the *BMJ*. The man opposite catches my eye.

"I see you're a doctor."

Self—aiming at closure: "Yes, I've a lot of work to get through."

"I'm on my way back from London. I was in Harley Street being interviewed as a volunteer for a drug trial of antidepressants. I had to use some 'kidology'."

"Yes?"

"I'm not depressed but they pay better in London than they do in Scotland or Manchester but I don't think I succeeded. The doctor advised me to use cognitive therapy. They said they would pay my expenses by cheque but eventually agreed to give me cash."

"Oh?"

"Yes, the trials are advertised, the best pay about £100 a day to volunteers. For a 20 day trial that's £2000. The worst trial was when I had to be woken up every hour to do mental tests, but usually it's like being on a health farm."

"What about making sure you don't come to harm?"

"Oh they have a committee of vicars and lawyers to decide it's all right, and it's nice to see your regular friends."

My train drew into the station; I was no longer irritated at the interruption.

Robert Boyd, *principal, St George's Hospital Medical School*