Therapeutic trials in childhood ALL: what’s their future?

O B Eden

Evidence based medicine
Since the first randomised controlled medical research trial in the United Kingdom on the use of streptomycin in tuberculosis (1947–1948), such trials have progressively become the gold standard by which the choice of treatment for patients with any particular disease, and the evidence to support such choice, can be judged. In the 1940s there were worries that doctors would be “unwilling to relinquish the doctrine of anecdotal experience” and accept the idea of any individual patient being randomised between a new treatment and a placebo, nothing at all, or what was perceived as standard treatment up until that time. Ultimately the physician always must decide for his or her patient what should be recommended, based on an assessment of the patient’s overall health status, and of course the doctor’s own experience and awareness, modified by external evidence. There is no doubt that the physician either hearing a presentation of a similar case or cases or reading of other peoples’ experiences will be influenced by such evidence. It is therefore essential that such evidence must be as objective, balanced, and unbiased as possible. The perfectly conducted randomised controlled trial should be the optimal way to produce such evidence. The impetus to collate all information for either the treatment or the prevention of disease has been increased by greater public awareness and education on health matters, coupled with easier access to information, which regrettably is not always reliable in the media or over the internet. In the Western world certainly, the public increasingly see through dogma and the “trust me I’m a doctor” approach. Development of the methodology to combine results from randomised controlled trials into systematic overviews is enhancing the ability of researchers to reach conclusions, or confirm evidence from individual trials. Such overview analysis has been particularly important when trials have been relatively small and so unable to provide adequate power, or indeed where there are conflicting results.

Paediatric oncology and trials
Oncology, especially paediatric oncology, has par excellence been one of the specialties that has adopted the approach of randomised controlled trials, and more recently the use of overview analyses. Faith in such an approach has been reinforced by unique evidence in children that outcome has been influenced by participation in trials. Stiller and Eatock recently reported on secular trends in the treatment of children with acute lymphoblastic leukaemia. When the proportion of children treated in paediatric oncology centres was compared for the five year periods 1980–84 and 1990–94, a rise from 77% to 89% was noted, but more interestingly, five year survival improved from 67% to 81%, irrespective of the size of the treatment centre. Entry of patients into national trials had an even more impressive effect on survival. For the 1980–84 cohort, five year survival for those on trials was 70% compared with 64% for those not entered into trials, but this gap had widened by 1990–94 to 84% for trial entrants and 68% for those treated off trial. The recent improvements amount to an increased survival for approximately 50 children a year who would have died in the 1980s. In an earlier study, the Childhood Cancer Research Group had shown an effect of centre size as well as trial entry, as had Meadows et al in 1983. The change in the most recent time period is thought to result from much more standardised care within centres of all sizes.

Why then has participation in acute lymphoblastic leukaemia (ALL) trials resulted in increased survival? Stiller and Eatock suggest that in the MRC UKALLVIII study and trial, 1980–84 survival from disease within the trial was counterbalanced by more early deaths. The induction therapy, especially after the introduction in 1981 of an additional two doses of daunomycin (one of the randomised questions) was associated with greater myelosuppression and more early infective deaths (5% induction deaths). For those receiving three years of continuing therapy there was also an excess of atypical infective deaths (for example measles and pneumonitis). In the next MRC trial (UKALLXX), with experience of protocol use and greater vigilance regarding infection risks (on the part of paediatricians, nurses, patients, and parents), induction mortality fell to under 2%. Consequently intensification of treatment on a trial produced a problem which was overcome with time. So much, then, for the critics of trials who say that toxicity is not recorded and acted upon, and that triallists frantically search for a significant p value at any price!

When you review those reported results you realise how important it is to remember that despite the risks, patients still had a 6% better chance of long term survival even in 1980–84, if they were actually treated in a trial. The excess mortality for those outside the national trial during 1985–89 (UKALLXI) was during continuation therapy from 3–24 months of treatment, while in the next trial between 1990 and 1994 the excess risk appeared to be immediately after diagnosis. This is an interesting finding. In UKALLXI there was no absolute requirement initially to register an intention to treat all patients until the time of first randomi-
sation at the end of induction therapy. There was an excess early mortality in non-trial patients in this time period compared with those registered early and entered into the trial. Even in experienced physicians’ hands it appears that following a set protocol with all its support guidelines does carry a benefit for the individual patient, almost irrespective of the actual randomisation question being asked. Physicians feeling that they always know best, particularly when they create totally ad hoc therapy, appear to provide an extra risk for their patients.

Consequently the results of the randomised questions in a trial may not turn out to be the most important lesson learnt from any particular study. In UKALLVIII the randomisation between two doses of daunomycin induction produced superior disease control for those who received the drug but excess early infective deaths, but in both arms event-free and total survival was superior to all previous United Kingdom results by over 10%, despite use of the same drugs that had been used previously, albeit in a more sustained and continuous way. Furthermore, in the next trial, UKALLX, when all patients received induction daunomycin early death rate was more than halved. With regard to the second randomisation for duration of treatment there were again fewer relapses in the patients who had longer treatment, but more infective deaths. However, the patients whose treatment stopped at two years and who subsequently relapsed appeared to be more salvageable, albeit at the price of more treatment, so there was no overall survival benefit for the longer treatment.

As a result of that kind of study, shorter treatment became the norm, and most collaborative groups opted for a two year duration of continuation therapy. Interest in longer periods of treatment has been rekindled by meta-analysis showing benefit.

Both current CCG and future MRC trials are readdressing the issue of duration of treatment, especially for boys, who still have an inferior survival to girls. No trial result should ever be carved in tablets of stone. It does appear that when paediatric haematologists and oncologists carefully plan a trial it is essential to ensure that they record toxicity and monitor compliance closely. There is a major benefit for those patients treated within the trial, and improvement in survival can be expected, not always because of better disease control, but sometimes because of a reduction in morbidity.

The advances that have led to survival figures of 70–80% for ALL means that health care planners and even funding bodies have started to question the likely benefit of further large randomised controlled trials to deliver ever continuing improvement in survival. Fortuitously the US Children’s Cancer Study Group (Nachman JB, personal communication) has recently reported event-free survival figures in excess of 90% for standard risk patients (age less than 10 years at diagnosis, with a white blood cell count under 50 x 10⁹/litre). Achievement of such an advance in the United Kingdom would equate to a further 50 lives saved a year. Inadequate primary treatment which then requires salvage with further intensive chemotherapy or bone marrow transplantation, or both, is costly in terms of patient and family physical and emotional sequelae and financially to the nation. For the foreseeable future the United Kingdom requires ongoing trials for children even with the lowest risk of relapse, among whom 30% do not survive.

Planning trials for ALL

One of the first principles of planning for a randomised controlled trial is to make it as universal as possible for any particular condition and the population affected—that is, to make it inclusive and not exclusive. This model has been used in the MRC UKALL trials, with all patients receiving essentially the same treatment in the time period 1980–96 with progressively more intensification. All categories of patients appear to have benefited from first two and then three intensification pulses post-induction. However, sadly these results are still inferior to the current reports from the Children's Cancer Study Group, and there is a need to test whether their experience can be repeated in the United Kingdom.

There have been two exceptions in the United Kingdom to this blanket approach of the “same therapy for all” which has been used since 1984. Both have resulted from observed poor results on conventional treatment in MRC trials. These exceptions are mature B cell ALL and infant ALL (70–80% having blast with rearrangements of the MLL gene).

The adoption of pulsed cyclophosphamide based multiagent lymphoma treatment has transformed an extremely bleak outlook for B cell ALL, where median survival was measured in months, so that cure rates increased first to 50% and then to 80–85%. The rarity of these conditions is limited to only 3–4% of ALL, has led to international collaboration, especially with the French paediatric oncology group for B-ALL. They had produced the most outstanding results. We are now in the difficult position of deciding whether we can in any way lessen treatment to reduce toxicity. For infant leukaemia, induction failure, treatment related mortality, higher rates of CNS relapse, and early overall failure are common. Overall five year event-free survival figures for infants with ALL have ranged between 25% and 40%. However the Berlin, Frankfurt, Munster (BFM) group showed a better event-free survival (53%) for those 19 infants in their 1986 study who had a good prednisolone response after seven days of steroids, compared with 14% in the 14 infants who had a poor response, as defined by having more than 1000 blasts/mm³ in their blood at day 7. Until that observation it had been assumed that all infants with ALL should have a bone marrow transplant once sustained remission was achieved if a cure was to be attained. Such transplantation procedures in infants, whatever the donor source, were associated with high mortality and major long term sequelae.

International collaboration involving Italy, Germany, Austria, the Czech Republic, The
Therapeutic trials in childhood ALL

Netherlands, France, Belgium, all the Scandinavian countries, the United Kingdom, St Jude’s Research Hospital and the Dana Farber Center, and other collaborative groups have all joined together to participate in an international infant protocol. This approach is undoubtedly necessary for such rarer forms of leukaemia, if randomised controlled trials are to be continued. The aims of this international collaboration are to assess the outcome of a hybrid treatment containing elements of both ALL and AML treatment, without irradiation and with only limited amounts of anthracyclines and alkylating agents, and to assess the value of a late intensification course which will form the randomised component to the trial. Bone marrow transplantation will be avoided in first remission for those patients with a rapid steroid response. A group of physicians representing all of the major American and European collaborative groups who first met together at an International Society of Paediatric Oncology meeting in Montevideo five years ago is now working together to tackle the rarer forms of ALL, including infant leukaemia, near haploidy, Philadelphia positive (Ph+) ALL, and those failing to remit on standard treatment. If any or all of these ultra high risk leukaemias are to be brought under control, such an international approach appears essential. The group has been documenting characteristics and outcome measures for these forms of ALL and exploring new approaches to treatment. For the next target, Ph+ ALL, Aricò et al found that an initial steroid response also predicts a more favourable outcome in that group,26 and Ribeiro et al have shown a favourable outcome for those with low initial white cell count.27 There does appear to be heterogeneity in both the biology and response among these rare tumours.

Collectively these ultra high risk ALL cases represent only about 11–12% of all childhood ALL—t(9;22) at 4%; MLL in infancy 2–3%, near haploidy 1%, failure to respond 4%—but they undoubtedly represent a great therapeutic challenge for clinicians. Lower remission rates and early relapse are common. However, in at least one group, those with near haploidy, remission is frequently achieved but with early development of resistance.35 The cellular mechanisms involved in such responses require clarification. This can only be achieved in rare diseases by many groups contributing biological material, as well as patients, to trials.

The increasingly sophisticated methods for defining ALL cases using a panel of antibodies (to distinguish immunological subclones) has shown the tremendous heterogeneity of ALL.28 30 31 However, apart from the broad distinction of B cell precursor ALL from mature B and T cell forms, they have not produced many independently significant prognostic categories. Those rare cases which express lymphoid associated (usually CD2 and CD7) and myeloid restricted molecules29 may do poorly. On the other hand modern era cytogenetic and molecular genetic techniques have revolutionised the definition of subgroups32 (as implied before with reference to infants and those with Philadelphia chromosome positivity). The rising frequency of leukaemias with bcr-abl and, to a lesser extent, MLL rearrangements with increasing age goes some way towards explaining the poorer outcome for adolescents and young adults.33 It is essential to define such high risk patients if we are to understand the molecular mechanisms that lead to primary or secondary resistance. Such analyses have already defined the favourable characteristics of those with high hyperdiploidy (more than 50 chromosomes per blast), and of TEL-AML1 fusions (12;21). There is some overlap of these two features, which are most commonly seen in children aged one to nine years with a low initial white blood cell count.34 Whether they are prognostically significant—indeed, of white cell count, age, and sex—is less clear.

One of the most interesting avenues of research now developing is an attempt to explain the prognostic characteristics in terms of drug responsiveness. For example, those with precursor B cell ALL with hyperdiploidy appear to accumulate higher levels of methotrexate polyglutamates within their cells, which suggests a potential mechanism to explain increased cytotoxicity on treatment, and points to ways in which treatment might be adjusted in these patients.35

The linkage of in vitro studies, especially of blast sensitivity or resistance, with other patient characteristics, and the clinical application of such data, are growth areas in ALL management and are increasingly being incorporated into clinical trials.36 Within the context of clinical trials, where the gold standard has been treatment in a uniform fashion, the adjustment of individual treatment on the basis of in vitro resistance profiles or pharmacokinetic/dynamic profiles presents logistic problems, but could be an important advance. Randomised questions can be asked, but individual patients must be given the optimal dosage and scheduling for their individual characteristics.

The current MRC ALL ‘97 trial is asking as one of its questions whether 6-thioguanine (6TG) might be superior to the traditional 6-mercaptopurine (6MP) as maintenance treatment, and parallels a US Children’s Cancer Study Group trial. It is now possible to measure thiopurine metabolites and hence to monitor for compliance,37 and to detect rare genetic polymorphisms of the enzyme thiopurine methyltransferase (TPMT) and of other pathway enzymes which influence drug handling.38 Since the 1950s 6MP has been used in ALL treatment during the maintenance phase, a phase which appears essential for the cure of children, and the concentration of intracellular 6-thiopurine nucleotides is probably extremely important. 6MP and its active nucleotide metabolites are major substrates for TPMT, whereas 6TG nucleotides are not. After all these years it is now being asked whether 6TG maybe a more effective way to deliver antimetabolite treatment in ALL,40 41 but in view of potential increased toxicity with thioguanine, most notably the occurrence of veno-occlusive disease,42 it is essential to test the 6MP v 6TG question in a randomised
A similar question has been raised about the type of steroids used for ALL (dexamethasone vs prednisolone). The Dutch Childhood Leukemia Group, backed by in vitro sensitivity studies (dexamethasone 35 times more potent), converted to dexamethasone usage and reported an 80% three year event-free survival, compared with 66% in their previous trial. It is felt essential to test the efficacy of dexamethasone vs prednisolone in a randomised fashion because of the reported increased toxicity of dexamethasone (increased disturbance of glucose metabolism, behavioural changes, myopathy, and avascular necrosis). The Children’s Cancer Study of America and the MRC studies testing this question are not mature enough yet for a firm conclusion. Both immediately and for the foreseeable future, optimisation of the delivery of individually effective agents, both globally and for the individual patient, is likely to form the core of randomised trials. Recent United Kingdom studies on asparaginase pharmacokinetics have suggested that route of administration, the product of asparaginase used, and previous exposure (and hence the likelihood of the development of silent inactivating antibodies) are all important. Evans et al have proposed that individualising chemotherapy within the context of trials, although complex, is essential.

Speed of response to treatment has already been alluded to in the context of infant ALL. Increasingly it is emerging as the single most important treatment-related prognostic characteristic, and may indeed be the most important overall prognostic factor. Pieters et al, in a series of studies using the MTT assay, have shown a difference in outcome for patients who are above and below the prednisolone LC50 median, and concluded that LC50 may be a better marker of in vivo sensitivity than early bone marrow response. These laboratory data have provided corroborative and explanatory evidence for the clinically observed result obtained in a series of BFM trials to a week long prephase of steroids. Such a prephase and the estimation of the peripheral blast cell count is clearly much less invasive than bone marrow review at either day 7 or day 14, which has been the mainstay of early response assessments over the last 20 years, since the original CCG studies proved its importance. It has been reported that the detection by PCR of leukemia associated clonal rearrangements of T cell receptor and immunoglobulin heavy chain genes at particular phases of treatment beyond induction is linked to poor prognosis. This is now at last being assessed in clinical trials with therapeutic alteration randomised in response to observation of persistent disease.

The more complex molecular methods of both subtyping and monitoring for disease are making stratification of patients more “scientific.” Giving all patients the same treatment was replaced very early on in the CCG and BFM studies by stratified treatment. Despite many previous attempts, only recently has a real consensus been reached on an overall scheme of risk classification, with universal agreement on those with ultra high risk of relapse, constituting 11–12% of cases; we are now moving towards distinguishing those high risk and intermediate risk groups, as defined by white blood cell count, age, and poor initial response to treatment, who do benefit from extra intensification of therapy. Nachman et al have shown that those with a high initial white cell count (and who are therefore by definition in the high risk group), who had more than 25% blasts in the bone marrow at day 7, and who previously would have fared very poorly did much better when given augmented treatment with longer, high dose, more sustained consolidation, interim maintenance involving an escalating dosage schedule of methotrexate and asparaginase, and two double delayed intensification modules. Although toxic, the treatment was tolerable and yielded a 20% improvement in event-free survival over the control arm, which in itself was superior to historical controls.

With groups such as the Children’s Cancer Study Group of America and the BFM group reporting such good results for standard risk and even conventionally high risk patients (though sadly no group is yet reporting much global improvement for the ultra high risk group), it looks as though international collaboration in studies is essential. We must not forget that most relapses still occur in patients with what we term “standard risk.” They are relapses which we could not predict from initial patient characteristics or early response to treatment. It is in these patients that optimisation of the drugs given and their delivery may really matter. Detection of resistant disease (and ways to overcome that) must be our focus in trials. For the largest group of patients we can probably continue to use identical or at least very comparable basic treatment worldwide. For the ultra high risk patients the small numbers involved mean that we have to collaborate internationally if we are to ask important therapeutic questions in the context of randomised clinical trials, based on earlier phase I and phase II studies of drug responsiveness.

New agents

Completely new agents do not often come on to the market for treatment of lymphoblastic leukaemia, but where they do, it is essential that phase I and phase II studies should be conducted in childhood, even though they are difficult to organise; it is, however, more likely that they will initially be conducted in adults, whose response may be less noticeable. Areas of interest at the present time include the use of arabinosylguanine for refractory T cell ALL, targeted immunotherapy, and the use of antisense oligonucleotides.

Conclusions

Until we approach 100% cure for all children with ALL there are questions to be asked, and treatments to be modified, all ultimately within the context of randomised controlled trials. There appear to be real benefits from participation in such trials, beyond the specific ques-
tions being asked. Increasingly the stratification of treatment in national trials is giving way to international collaboration. Running trials across national boundaries is not easy, but certainly feasible if people try hard to make them work. The life of the individual child is too precious for us to fail to attempt to cure all patients who present to us, and we should resist pressure from health service planners to “give up on 30%” of our patients.