Infant wellbeing at 2 years of age in the Growth Restriction Intervention Trial (GRIT): multicentred randomised controlled trial

The GRIT study group

Summary

Background Although delivery is widely used for preterm babies failing to thrive in utero, the effect of altering delivery timing has never been assessed in a randomised controlled trial. We aimed to compare the effect of delivering early with delaying birth for as long as possible.

Methods 548 pregnant women were recruited by 69 hospitals in 13 European countries. Participants had fetal compromise between 24 and 36 weeks, an umbilical-artery doppler waveform recorded, and clinical uncertainty about whether immediate delivery was indicated. Before birth, 588 babies were randomly assigned to immediate delivery (n=296) or delayed delivery until the obstetrician was no longer uncertain (n=292). The main outcome was death or disability at or beyond 2 years of age. Disability was defined as a Griffiths developmental quotient of 70 or less or the presence of motor or perceptual severe disability. Analysis was by intention-to-treat. This trial has been assigned the International Standard Randomised Controlled Trial Number ISRCTN41358726.

Findings Primary outcomes were available on 290 (98%) immediate and 283 (97%) deferred deliveries. Overall rate of death or severe disability at 2 years was 55 (19%) of 290 immediate births, and 44 (16%) of 283 delayed births. With adjustment for gestational age and umbilical-artery doppler category, the odds ratio (95% CrI) was 1·1 (0·7–1·8). Most of the observed difference was in disability in babies younger than 31 weeks of gestation at randomisation: 14 (13%) immediate versus five (5%) delayed deliveries. No important differences in the median Griffiths developmental quotient in survivors was seen.

Interpretation The lack of difference in mortality suggests that obstetricians are delivering sick preterm babies at about the correct moment to minimise mortality. However, they could be delivering too early to minimise brain damage. These results do not lend support to the idea that obstetricians can deliver before terminal hypoxaemia to improve brain development.

Introduction Obstetricians use many tests of fetal wellbeing to predict fetal death and handicap, and in high-risk pregnancies, evidence has shown that use of the umbilical-artery doppler flow-velocity waveform might reduce perinatal mortality.1,2 Recent studies have clarified the order in which alterations in various vessel doppler waveforms, fetal heart rate, and behavioural abnormalities take place, but they cannot indicate when delivery should take place in response to these changes.3,4 No trials have been undertaken to guide the timing of delivery. Datasets relating clinical outcome to gestational age and to the degree of test abnormality are of limited help because they include a heterogeneous group of pregnancies with delivery timing unspecified.5,6 The wide variation in policies is not surprising.7 Delay could increase hypoxia and affect brain development, but early delivery carries the dangers of prematurity and the associated risk of cerebral palsy.8,9

The Growth Restriction Intervention Trial (GRIT) was designed to help time delivery for such babies and to test the hypothesis that early delivery, to pre-empt intrauterine hypoxia, would alter brain development compared with delaying delivery for as long as possible, to gain maturity. Further details of design have been presented previously.10 Interim findings suggested similar mortality before discharge between immediate and delayed delivery groups. Here we consider infant wellbeing at or beyond 2 years of age.

Patients and methods

Patients Details of recruitment have been reported previously.11 The overall recruitment period ran from November, 1993, to March, 2001, in 69 hospitals in 13 European countries. Criteria for entry were singleton or multiple pregnancies for which the responsible clinician was uncertain about whether to deliver the baby immediately, the gestational age was between 24 and 36 completed weeks, and the umbilical-artery doppler waveform had been recorded. Basic clinical data, the umbilical-artery doppler waveform and (if available) results of a biophysical profile score, computerised cardiotocograph analysis, and cerebral-artery doppler waveform were recorded before randomisation.
Methods

An independent programmer organised allocation, using both randomisation and minimisation. A paper-based number sequence with balanced blocks of 8–12 was used except during office hours when a computer-generated sequence was used. The process was designed to mask allocation from participating clinicians, including those with access to the central trial office.

Trial allocations were either immediate delivery (ie, within 48 h to permit completion of a steroid course) or deferred delivery (ie, when safe delivery could be delayed no longer, because worsening test results or the passage of time favoured delivery). Mode of delivery and monitoring strategies for the deferred group were left up to the attending obstetrician. In most cases, caesarean delivery was done without attempts to induce labour.

In the UK, we traced participants by flagging the child on the Office of National Statistics register. Local co-ordinators traced children in other countries by use of various methods. All participants were contacted every year. When children were 2 years of age, every child’s family practitioner or paediatrician was asked to complete a short questionnaire about the child’s present health status, which included questions on cerebral palsy, deafness, and blindness. If a child had died, we sought information about the cause of death. The study paediatrician (ML), who was masked to trial allocation, classified deaths using the extended Wigglesworth classification. No extra examination was undertaken to confirm these diagnoses. At or after 2 years of age (corrected for gestational age at birth), every child was visited by a Griffiths-trained assessor who completed a full Griffiths assessment, at the child’s home or in a health facility according to local convenience. Assessors were masked to the child’s group allocation.

The main outcome measure was death or disability at or after 2 years of age. This formulation was more precise than the protocol definitions of survival and development at 2 years, survival to hospital discharge, and Griffiths full scale developmental quotient (DQ) measured at 2 years, and was agreed with the data monitoring and ethics committee during the trial. Disability was defined as a diagnosis of cerebral palsy, little or no vision, requirement for a hearing aid, or a Griffiths DQ of 70 or less. Interim analyses, including survival to discharge and caesarean delivery rates, have already been reported.

Statistical analysis

Although analysis was planned to be Bayesian, the following sample size calculations were made at the design stage. We assumed the Griffiths DQ to have an SD of 15. The aim was to recruit 150 participants at between 24–30 weeks and 450 at more than 30 weeks of gestation.
gestation. This target would give about 95% power at the 5% significance level to detect a difference of ten and five points in the early and later gestational age groups, respectively.

Data monitoring and analysis were Bayesian. Data monitoring and analysis were Bayesian. This trial had no prespecified stopping rules. Interim analyses of accumulating unmasked data were presented to the trial data monitoring committee every year. These unmasked data were also presented to clinical collaborators at biannual trial meetings with the idea that individuals with strong prior beliefs would stop recruiting if they were no longer clinically uncertain, and that other individuals might be encouraged to recruit.

For every outcome, raw data were presented as an odds ratio (OR) with a 95% credibility interval (CrI). CrIs are similar to confidence intervals. Bayesians prefer CrI as an indication of the range within which they believe the true effect to lie. ORs greater than 1·0 indicated an increased likelihood of a specified outcome (death or disability) with immediate delivery. These ORs were used to revise prior beliefs. We interpreted results for people with three archetypal prior beliefs. First, we considered posterior beliefs of someone (a sceptic) who before the study had believed that the most likely effect would be no difference (ie, OR=1·0), and were 95% certain that the effect would be no more than a halving or doubling of the odds (CrI 0·5–2·0). For results that were equivocal, we considered posterior beliefs for an enthusiast who believed beforehand that a halving or doubling of the odds was most likely, but had the same degree of uncertainty (variance) as the sceptic.

Table 2: Trial entry characteristics before randomisation

<table>
<thead>
<tr>
<th></th>
<th>Immediate (n=273)</th>
<th>Delayed (n=275)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age at entry (weeks)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primiparous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous pregnancy loss after 24 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical growth restriction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension (&gt;140/90 mm Hg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteinuria (&gt;0.3 g/L)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple pregnancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steroids already given</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Umbral artery EDF*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table shows number (%) or median (IQR). *Denominator is fetuses undergoing umbilical-artery doppler. Includes one triplet pregnancy.
Since there was a-priori concern that the effect of immediate or delayed delivery would differ with gestational age and degree of fetal compromise, we undertook logistic regression analysis to allow for these effects and to examine possible interactions between each with outcome. All presented ORs were derived from this adjusted model. In this model, the interaction between degree of fetal compromise and allocation group was retained. Multiple pregnancies were included in the results. In the exploratory analysis, multiple births were clearly correlated with survival, with more concordant outcomes than expected by chance. We therefore omitted multiple pregnancies from the data modelling. In practice, inclusion of such pregnancies made no difference to conclusions. All analyses were done by intention-to-treat—ie, data from births were analysed according to their randomised group, irrespective of compliance with that allocation. This study was approved by the UK North West region multicentre research ethics committee (REC), the relevant UK local RECs, and the local RECs in non-UK centres.

The protocol for this study was peer-reviewed and accepted by The Lancet; a summary of the protocol was published on the journal’s website, and the journal then made a commitment to peer-review the primary clinical manuscript. This trial has been assigned the International Standard Randomised Controlled Trial Number ISRCTN41358726.

Role of the funding source
The EU Concerted Action funded a series of collaborators meetings around Europe, paid for training of paediatricians to perform Griffiths assessments, and supported costs of the 2-year assessments in Slovenia, Poland, Hungary, Czech Republic, Greece, and Cyprus. The Dutch Princess Beatrix Foundation funded the 2-year assessment in the Netherlands. Neither of these sponsors was involved in the study design, collection, analysis, and interpretation of data, writing of the report, or the decision to submit the paper for publication. The UK Medical Research Council (MRC) funded the main trial, and members of the MRC-appointed trial steering committee advised on the protocol design and approved the final paper.

Results
We were not able to record how many women were eligible for the study in participating hospitals during the trial period. 548 women participated in our study and delivered 588 babies. This group included 195 women (205 babies) between 24 and 30 completed weeks and 353 (383) between 31 and 36 weeks of gestation. One woman randomly allocated to delayed delivery (in a deprived area of a large city), discharged herself, never returned, and was lost to follow-up. Three fetuses did not have the doppler waveform recorded before randomisation. Gestational age and umbilical-artery doppler waveform analyses recorded before randomisation are shown in table 1. Figure 1 shows the trial entry profile and losses to follow-up. 290 of 296 babies randomly allocated to immediate delivery and 283 of 292 allocated to delay were included in the primary analysis. Both groups were well balanced at entry (table 2).

Median interval between randomisation and delivery in the immediate delivery group was 0·9 days (IQR 0·4–1·3), and in the delay group was 4·9 days (2·0–11·0). For pregnancies randomly assigned at up to 30 weeks’ gestation, corresponding figures were 0·8 days (0·3–1·1) versus 3·2 days (1·5–8·0). Of those...
Articles

from 31 weeks’ gestation, figures were 0·9 days (0·5–1·6) versus 6·7 days (2·8–14·0), respectively.

Main fetal outcomes by trial allocation are shown in table 3. Table 4 shows causes of death. Types of disability by trial allocation are shown in table 5. Table 6 summarises the overall Griffiths scores and subscale scores for every group above and below 30 weeks’ gestation. Figure 2 contains trial results by gestational age in 2-week bands and the umbilical-artery waveform divided into presence or absence of end diastolic flow velocities.

Figure 3 shows the archetypal prior beliefs, trial data, and archetypal posterior CrIs for the primary outcome overall, and separately by every category of umbilical-artery EDF waveform. Since the data did not show a clear benefit of either policy, they are sufficiently negative to convince either type of enthusiast that there is no substantive difference between policies. In this case, data on death or disability at 2 years were inconclusive (OR 1·1, 95% CrI 0·7–1·8). The sceptic’s posterior belief represented a shift in favour of delay (1·1, 0·7–1·6), but retaining 34% belief that immediate delivery might be superior. The enthusiast for immediate delivery could be sufficiently moved to consider future randomisation, with 23% belief that delay might be superior (0·9, 0·6–1·3). The enthusiast for delay was moderated but remained strongly in favour of delay, with only 6% belief that immediate delivery might be superior (1·4, 0·9–2·0).

Discussion

We had already shown that when obstetricians were uncertain about the timing of delivery, they were prepared to vary the timing by about 4 days, and although such delay caused some stillbirths, earlier delivery resulted in an almost exactly equal number of additional perinatal deaths.21 This finding implies that obstetricians were, on average, delivering babies at the correct point to minimise perinatal mortality. In the present paper, we have shown that at 2 years of age, there is a trend towards more disability in the immediate delivery group, but no overall difference in Griffiths DQ. The main strengths of this study are the randomised comparison, relatively large sample size, and 2-year follow-up period. Previous similar trials of early or delayed delivery for severe hypertensive diseases in preterm pregnancy were much smaller, did not have any follow-up beyond hospital discharge, and were not directly comparable, because babies with fetal compromise were excluded.20,21 Nevertheless, those trials also showed a trend towards lower fetal mortality and morbidity in the group allocated to delayed delivery. The only other comparative study of which we are aware was an observational comparison between two Dutch hospitals with contrasting policies for management of extreme prematurity.22 No major effect resulting from obstetric management was seen.

Some people might argue that the dosage of the intervention, a delay of 4 days, was small. This period was unavoidable, in view of the ethical necessity to deliver when the obstetrician was no longer uncertain. However, 4 days is also the longest delay ever assessed by a randomised trial in this scenario. It clearly made a difference to the babies involved—stillbirths rose by nearly five-fold, and deaths before discharge fell by more than a third.

<table>
<thead>
<tr>
<th>All gestations</th>
<th>24–30 weeks</th>
<th>31–36 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immediate (n=256)</td>
<td>Deferred (n=251)</td>
</tr>
<tr>
<td>Cerebral palsy</td>
<td>12 (5%)*†</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>Little or no vision</td>
<td>2 (1%)</td>
<td>1 (0%)</td>
</tr>
<tr>
<td>Wears hearing aids</td>
<td>1 (0%)</td>
<td>1 (0%)</td>
</tr>
<tr>
<td>Griffiths DQ &lt;70</td>
<td>13 (5%)</td>
<td>10 (4%)</td>
</tr>
<tr>
<td>Any of the above</td>
<td>20 (8%)</td>
<td>13 (4%)</td>
</tr>
</tbody>
</table>

*One baby also had a DQ <70 and visual impairment. †Six babies also had a DQ <70. ‡Two babies also had a DQ <70.

Table 5: Disability details

<table>
<thead>
<tr>
<th>24–30 weeks</th>
<th>31–36 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected age at assessment (weeks)</td>
<td>Immediate (n=76)</td>
</tr>
<tr>
<td>Overall DQ</td>
<td>116 (107–118)</td>
</tr>
<tr>
<td>Number of babies with DQ &lt;70</td>
<td>9 (12%)</td>
</tr>
<tr>
<td>Number of babies with DQ 71–85</td>
<td>14 (18%)</td>
</tr>
</tbody>
</table>

Subscale scores*  

Motor | 96 (82–108) | 99 (89–108) | 100 (91–110) | 100 (91–112) |
Personal–social | 101 (82–113) | 103 (91–115) | 104 (93–118) | 103 (92–113) |
Hearing and speech | 97 (80–109) | 97 (84–108) | 100 (86–112) | 100 (86–111) |
Eye and hand | 94 (81–103) | 96 (89–108) | 96 (90–107) | 97 (89–108) |
Performance | 96 (83–110) | 101 (89–112) | 104 (92–117) | 100 (93–113) |
Practical reasoning | 102 (92–121) | 102 (90–108) | 103 (91–114) | 103 (96–115) |

*Denominator is those with subscale assessed. Data show median (IQR) unless indicated otherwise.

Table 6: Outcomes of Griffiths assessments
Previously reported findings of overall similar fetal mortality had suggested that, on average, participating obstetricians were delivering such fetuses at the correct time. However, the present trend towards more severe disability in the immediate delivery group suggests that overall, obstetricians are delivering too early to minimise fetal brain damage. This evidence is much stronger than the previously reported slight trend towards more intracerebral haemorrhage in the early delivery group, which could have resulted from differential scanning protocols, and which might not have been translated into more brain damage.

Our results accord with epidemiological evidence that rates of cerebral palsy are high in low birthweight and presumably premature babies. The explanation for that rise has generally been assumed to be improvements in neonatal care, whereby babies who previously would have died are now surviving with cerebral palsy. The present data suggest that part of the rise might include iatrogenic delivery before the optimum moment to minimise cerebral palsy.

Overall risk of cerebral palsy in the immediate delivery group younger than 30 weeks of gestation is similar to that previously reported in babies of similar gestational age. By contrast, the deferred delivery group younger than 30 weeks had no cases of cerebral palsy, which was much lower than expected. This result suggests that deferred delivery could protect against development of the condition, and might be due to a number of reasons. Periventricular leucomalacia was probably the most likely underlying pathology, although we did not obtain information on brain scans from babies with disability. No difference in steroid use before randomisation between the two groups was seen, so this factor would probably have not prevented cerebral palsy in the deferred delivery group, although differential time to responding to steroid administration after randomisation could have been a factor. Risk of cerebral palsy might be related to neonatal complications that have been prevented in the delay group by extra days in utero. Alternatively, the deferred delivery group might have been able to redistribute their cerebral perfusion in this interval between immediate and deferred delivery, so that brains were better protected at delivery, perhaps by an enhanced ability to compensate for changes in blood pressure and other variables.

Participants in the present trial comprised only a small proportion, perhaps a tenth, of babies delivered before 36 weeks’ gestation in the participating units over the study period. Our participants were an even smaller proportion of the total number of pregnancies for which the possibility of early delivery was considered but the obstetrician decided that delay was preferable. Unfortunately, but inevitably, the number of eligible patients who were not recruited was impossible to record. Dividing lines between potentially eligible cases where the obstetrician failed to offer trial entry, cases where patients declined to consider participation (sometimes even before they were eligible), and cases for whom the timing of delivery was clear-cut were inevitably blurred. Results should therefore be generalised beyond the recruitment criteria—ie, those pregnancies with a clinical risk factor, an umbilical-artery doppler recorded, and where the obstetrician was uncertain—with caution.

Nevertheless, the onus is now on those who advocate early delivery in other specific groups of patients to test their beliefs in a similar trial with different recruitment criteria. The present data suggest that trials at slightly later gestational ages for every type of doppler waveform result are both feasible and worth testing. To derive a consensus on which patients should be recruited would be easy.

Although the present data will be useful to inform obstetricians about delivery timing, our study cannot give detailed guidance on the exact point at which compromised fetuses should be delivered. First, few obstetricians would time delivery on the basis of gestational age and umbilical doppler waveform alone—all would also take into account the clinical history, fetal size, and presence or absence of coexisting hypertension or haemorrhage. Most individuals would also use serial fetal heart-rate analysis, biophysical score, or analysis of the ductus venosus doppler waveform, to indicate whether the fetus was in or near preterminal hypoxaemia. In the present study, factors that provoked delivery in patients who were randomly allocated to the delayed group included maternal
request, worsening hypertension or bleeding, fetal heart rate decelerations and venous doppler waveform changes, and a low or falling biophysical score, or a combination of these. In some cases, the final decision to deliver was made by a doctor unininvolved with the trial.

Clinicians faced with a patient at a known gestational age and with a specific umbilical-artery doppler waveform will be tempted to look for the outcomes of babies entered in the GRIT trial with the same gestational age and doppler results. Unfortunately, despite the relatively large size of our study, the number of participants entered at every particular gestational age or with particular doppler waveform results was small (figure 2). Although no subgroup showed that an immediate or delayed delivery policy was clearly superior, CI’s of any estimate of effect broken down into subgroups in this way would be very wide.

Nevertheless, the present study should discourage doctors who deliver such fetuses before the point at which they believe delivery can be delayed no longer. This caution applies particularly to pregnancies for which early delivery is considered before 30 weeks. In this situation, we believe that the obstetrician should delay.

Contributors
J G Thornton conceived, designed, and ran the study, analysed data, and wrote the paper. J Hornbuckle ran the study, analysed data, and wrote the paper. A Vail conceived and designed the study, and analysed data, and wrote the paper. J D Spiegelhalter conceived and designed the study and analysed data. M Levene conceived, designed, and ran the study, and wrote the paper.

GRIT study group
Belgium—Algemeen Centrum Ziekenhuis, Antwerp, B Van Bulck; Cypres—Arch Makarios III Hospital, Nicosia, G M Kalakoutis; Czech Republic—Ceske Budejovice, Hradec Kralove, P Sak; Germany—Universitat Munchen, K T M Schneider; Greee—Alexandra, Athens, S E Karpathios; Hungry—Debrecen, T Major; Italy—Tirrin, T Todro; Fatebenefratelli, Rome, D Arduini; Ospedale Salesi, Ancona, A Tranquirli, C Flumini; Ospedale Valduce, Como, A C Tenore; University Degli Studi Di Milano, Bicocca, Monza, N Roncaglia; University of Brescia, T Frusci; IRCCS Stella Maris, Pisa, G Ferretti; Florence—A Z Groningen, M J Neiwans; A Z Leiden.

J van Roosmalen; De Heel, Zaanmond, J W van der Slkke; Free University, Amsterdam, H van Geijn; Kemenner Gasthuis Lokatie Deo, Haarlem, P J M Pernet; Amsterdam Medical Centre, H Wolf; A Z Utrecht, R H Stigler; Poland—Polish Mother’s Memorial Hospital, Lodz, J Wilczynski; Portugal—Hospital Universidad Coimbra, E Vasco; Saud Arabia—Riyadh Security Forces Hospital, M Abdullah; Slovenia—Ljubljana, Z Novak-Antolic; UK—Aberdeen, P Danielian; Redditch, S D Jenkins; Leeds, J Walker; St Mary’s, Manchester, M Maresht; St Mary’s, Isle of Wight, A Yoong; St Michael’s, Bristol, P Soothill; Sunderland, H Cameron; the Royal Cornwall, Truro, D Byrne; Heath Cardiff, B Beattie; West Cumberland, S Bober.

Poediatric assessors

Trial offiice staff: B Feller, E Isaac, L Liddle (principal trial coordinator), T McGhee, P Rowse.

Data monitoring ethics committee
P Howie, Max Parmar, D Field.

Trial steering committee
A Grant, P Steer, M Levene, D Torgeson, S Kitzinger.

Poediatric steering group
M Levene, N Marlowe, D Wolke, L De Vries, A Johnson, G Ferretti.

Conflict of interest statement
None declared.

Acknowledgments
Richard Lilford had the original idea and contributed many ideas to the trial design.

References


