HEALTH TECHNOLOGY ASSESSMENT

VOLUME 23 ISSUE 4 FEBRUARY 2019
ISSN 1366-5278

Ten-year follow-up of a randomised trial of drainage, irrigation and fibrinolytic therapy (DRIFT) in infants with post-haemorrhagic ventricular dilatation

Karen Luyt, Sally Jary, Charlotte Lea, Grace J Young, David Odd, Helen Miller, Grazyna Kmita, Cathy Williams, Peter S Blair, Aída Moure Fernández, William Hollingworth, Michelle Morgan, Adam Smith-Collins, N Jade Thai, Steven Walker-Cox, Kristian Aquilina, Ian Pople and Andrew Whitelaw

DOI 10.3310/hta23040

NHS
National Institute for Health Research

Ten-year follow-up of a randomised trial of drainage, irrigation and fibrinolytic therapy (DRIFT) in infants with post-haemorrhagic ventricular dilatation

Karen Luyt,[1]* Sally Jary,[1] Charlotte Lea,[1] Grace J Young,[2] David Odd,[1,3] Helen Miller,[1] Grazyna Kmita,[4] Cathy Williams,[5] Peter S Blair,[2] Aída Moure Fernández,[2] William Hollingworth,[2] Michelle Morgan,[6] Adam Smith-Collins,[1] N Jade Thai,[7] Steven Walker-Cox,[1] Kristian Aquilina,[8] Ian Pople[9] and Andrew Whitelaw[1]

[1]Neonatal Neurology, University of Bristol, Bristol, UK
[2]Bristol Randomised Trials Collaboration, University of Bristol, Bristol, UK
[3]Neonatal Medicine, North Bristol NHS Trust, Bristol, UK
[4]Faculty of Psychology, University of Warsaw, Warsaw, Poland
[5]Paediatric Ophthalmology, University of Bristol, Bristol, UK
[6]Department of Psychology, Community Children's Health Partnership, Bristol, UK
[7]Clinical Research and Imaging Centre, Bristol, UK
[8]Paediatric Neurosurgery, Great Ormond Street Hospital, London, UK
[9]Paediatric Neurosurgery, University Hospitals Bristol NHS Trust, Bristol, UK

*Corresponding author

Declared competing interests of authors: none

Published February 2019
DOI: 10.3310/hta23040

This report should be referenced as follows:

Luyt K, Jary S, Lea C, Young GJ, Odd D, Miller H, *et al*. Ten-year follow-up of a randomised trial of drainage, irrigation and fibrinolytic therapy (DRIFT) in infants with post-haemorrhagic ventricular dilatation. *Health Technol Assess* 2019;**23**(4).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.513

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 12/35/61. The contractual start date was in September 2014. The draft report began editorial review in November 2016 and was accepted for publication in August 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Ten-year follow-up of a randomised trial of drainage, irrigation and fibrinolytic therapy (DRIFT) in infants with post-haemorrhagic ventricular dilatation

Karen Luyt,[1]* Sally Jary,[1] Charlotte Lea,[1] Grace J Young,[2] David Odd,[1,3] Helen Miller,[1] Grazyna Kmita,[4] Cathy Williams,[5] Peter S Blair,[2] Aída Moure Fernández,[2] William Hollingworth,[2] Michelle Morgan,[6] Adam Smith-Collins,[1] N Jade Thai,[7] Steven Walker-Cox,[1] Kristian Aquilina,[8] Ian Pople[9] and Andrew Whitelaw[1]

[1]Neonatal Neurology, University of Bristol, Bristol, UK
[2]Bristol Randomised Trials Collaboration, University of Bristol, Bristol, UK
[3]Neonatal Medicine, North Bristol NHS Trust, Bristol, UK
[4]Faculty of Psychology, University of Warsaw, Warsaw, Poland
[5]Paediatric Ophthalmology, University of Bristol, Bristol, UK
[6]Department of Psychology, Community Children's Health Partnership, Bristol, UK
[7]Clinical Research and Imaging Centre, Bristol, UK
[8]Paediatric Neurosurgery, Great Ormond Street Hospital, London, UK
[9]Paediatric Neurosurgery, University Hospitals Bristol NHS Trust, Bristol, UK

*Corresponding author Karen.Luyt@bristol.ac.uk

Background: The drainage, irrigation and fibrinolytic therapy (DRIFT) trial, conducted in 2003–6, showed a reduced rate of death or severe disability at 2 years in the DRIFT compared with the standard treatment group, among preterm infants with intraventricular haemorrhage (IVH) and post-haemorrhagic ventricular dilatation.

Objectives: To compare cognitive function, visual and sensorimotor ability, emotional well-being, use of specialist health/rehabilitative and educational services, neuroimaging, and economic costs and benefits at school age.

Design: Ten-year follow-up of a randomised controlled trial.

Setting: Neonatal intensive care units (Bristol, Katowice, Glasgow and Bergen).

Participants: Fifty-two of the original 77 infants randomised.

Interventions: DRIFT or standard therapy (cerebrospinal fluid tapping).

Main outcome measures: Primary – cognitive disability. Secondary – vision; sensorimotor disability; emotional/behavioural function; education; neurosurgical sequelae on magnetic resonance imaging; preference-based measures of health-related quality of life; costs of neonatal treatment and of subsequent health care in childhood; health and social care costs and impact on family at age 10 years; and a decision analysis model to estimate the cost-effectiveness of DRIFT compared with standard treatment up to the age of 18 years.

Results: By 10 years of age, 12 children had died and 13 were either lost to follow-up or had declined to participate. A total of 52 children were assessed at 10 years of age (DRIFT, $n = 28$; standard treatment, $n = 24$). Imbalances in gender and birthweight favoured the standard treatment group. The unadjusted mean cognitive quotient (CQ) score was 69.3 points [standard deviation (SD) 30.1 points] in the DRIFT group compared with 53.7 points (SD 35.7 points) in the standard treatment group, a difference of 15.7 points, 95% confidence interval (CI) –2.9 to 34.2 points; $p = 0.096$. After adjusting for the prespecified covariates (gender, birthweight and grade of IVH), this evidence strengthened: children who received DRIFT had a CQ advantage of 23.5 points ($p = 0.009$). The binary outcome, alive without severe cognitive disability, gave strong evidence that DRIFT improved cognition [unadjusted odds ratio (OR) 3.6 (95% CI 1.2 to 11.0; $p = 0.026$) and adjusted OR 10.0 (95% CI 2.1 to 46.7; $p = 0.004$)]; the number needed to treat was three. No significant differences were found in any secondary outcomes. There was weak evidence that DRIFT reduced special school attendance (adjusted OR 0.27, 95% CI 0.07 to 1.05; $p = 0.059$). The neonatal stay (unadjusted mean difference £6556, 95% CI –£11,161 to £24,273) and subsequent hospital care (£3413, 95% CI –£12,408 to £19,234) costs were higher in the DRIFT arm, but the wide CIs included zero. The decision analysis model indicated that DRIFT has the potential to be cost-effective at 18 years of age. The incremental cost-effectiveness ratio (£15,621 per quality-adjusted life-year) was below the National Institute for Health and Care Excellence threshold. The cost-effectiveness results were sensitive to adjustment for birthweight and gender.

Limitations: The main limitations are the sample size of the trial and that important characteristics were unbalanced at baseline and at the 10-year follow-up. Although the analyses conducted here were prespecified in the analysis plan, they had not been prespecified in the original trial registration.

Conclusions: DRIFT improves cognitive function when taking into account birthweight, grade of IVH and gender. DRIFT is probably effective and, given the reduction in the need for special education, has the potential to be cost-effective as well. A future UK multicentre trial is required to assess efficacy and safety of DRIFT when delivered across multiple sites.

Trial registration: Current Controlled Trials ISRCTN80286058.

Funding: This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 23, No. 4. See the NIHR Journals Library website for further project information. The DRIFT trial and 2-year follow-up was funded by Cerebra and the James and Grace Anderson Trust.

Contents

List of tables

List of figures

List of abbreviations

BAS III	British Ability Scales version three
BRTC	Bristol Randomised Trials Collaboration
BSID III	Bayley Scales of Infant and Toddler Development version three
CI	confidence interval
CP	cerebral palsy
CQ	cognitive quotient
CRICBristol	Clinical Research and Imaging Centre Bristol
CSF	cerebrospinal fluid
CSRI	Client Service Receipt Inventory
CVI	cerebral visual impairment
DAE	developmental age equivalent
DfE	Department for Education
DQ	development quotient
DRIFT	drainage, irrigation and fibrinolytic therapy
EBD	excess bed-day
ED	emergency department
EQ-5D	EuroQol-5 Dimensions
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
EQ-5D-Y	EuroQol-5 Dimensions – Youth
GLM	general linear model
GMFCS	Gross Motor Function Classification System
HDU	high-dependency unit
HES	Hospital Episode Statistics
HRG	Healthcare Resource Group
HRQoL	health-related quality of life
HUI3	Health Utilities Index – 3
ICER	incremental cost-effectiveness ratio
ICU	intensive care unit
INMB	incremental net monetary benefit
IQR	interquartile range
ITT	intention to treat
IVH	intraventricular haemorrhage
KS1	Key Stage 1
KS2	Key Stage 2
LP	lumbar puncture
MAR	missing at random
MDI	Mental Development Index
Movement ABC	Movement Assessment Battery for Children-2
MRI	magnetic resonance imaging
NASS	National Association of Independent Schools & Non-Maintained Special Schools
NICE	National Institute for Health and Care Excellence
NICU	neonatal intensive care unit
NNT	number needed to treat
ODN	Operational Delivery Network
OLS	ordinary least squares
OR	odds ratio
PHVD	post-haemorrhagic ventricular dilatation
QALY	quality-adjusted life-year
QoL	quality of life
RCT	randomised controlled trial
SAP	statistical analysis plan
SCU	special care unit

SD	standard deviation
SDQ	Strengths and Difficulties Questionnaire
SEN	special educational needs
SLT	speech and language therapy
TPA	tissue plasminogen activator
VAS	visual analogue scale
VP	ventriculoperitoneal

Plain English summary

Bleeding into the fluid spaces of the brain is a common complication of being born very early. Such bleeds often block the normal fluid flow around the brain, causing expansion of the fluid spaces (ventricles), pressure on the brain and serious disability.

The standard treatment is to drain off excess fluid with a needle. This may need to be repeated and often leads to further complications. An alternative treatment is to wash out the blood clot and clear the effects of bleeding. This is called drainage, irrigation and fibrinolytic therapy (DRIFT).

Fifteen years ago (2003–6), the new DRIFT washout treatment was compared with standard treatment in a randomised trial. A total of 77 premature babies with bleeding in, and expansion of, the brain spaces were studied. At age 2 years, the babies in the washout (DRIFT) group were doing slightly better. Fewer of them had severe learning problems than those in the standard treatment group.

In the present study, we followed up those babies we could trace to school age. We managed to examine 52 out of the 66 surviving babies.

DRIFT improved cognitive function and reduced the need for special education at age 10 years. This is the first treatment to show improved brain function in premature babies with this condition.

DRIFT treatment was slightly more expensive. However, the long-term benefits were such that, after taking into account the costs of special schooling, the treatment was probably cheaper overall.

Despite these results, it is not possible to implement DRIFT in the NHS right away. Few centres have the skills and expertise to deliver DRIFT safely and, with improvements in survival of the most immature babies, those who would be eligible for DRIFT are now even more premature than those included in the original trial. Therefore, we recommend the development of DRIFT for the NHS through a further implementation trial based in a few specialist centres.

Scientific summary

Background

Severe intraventricular haemorrhage (IVH) with post-haemorrhagic ventricular dilatation (PHVD) is a serious neurological complication seen in preterm infants, with significant neurodisability in survivors. No medical intervention has been proven to reduce neurodevelopmental disability in infants with PHVD.

Objectives

Our primary hypothesis was that drainage, irrigation and fibrinolytic therapy (DRIFT) will reduce severe cognitive disability in children assessed at school age.

Our secondary hypotheses were that DRIFT will:

- improve cerebral visual dysfunction
- improve sensorimotor ability
- improve education outcomes
- improve emotional/behavioural difficulties
- improve preference-based measures of health-related quality of life (HRQoL)
- reduce the health, social care and broader societal costs at 10-year follow-up.

The aims of this study were to:

1. compare cognitive function, visual function, sensorimotor ability and emotional well-being between the two treatment groups in the DRIFT trial at school age
2. explore the use of specialist health/rehabilitative and educational services
3. estimate the economic cost and outcomes of the DRIFT intervention by age 11 years and model longer-term costs and outcomes
4. assess ventricular dilatation and neurosurgical sequelae in the two treatment groups by clinical neuroimaging.

Methods

Design

This was a long-term follow-up study of a multicentre randomised controlled trial set in neonatal intensive care units in Bristol (UK), Katowice (Poland), Glasgow (UK) and Bergen (Norway).

Participants

The children, now aged 10 years, had been randomised to the DRIFT trial as preterm infants and all had suffered a severe degree of PHVD. A small feasibility study preceded the follow-up study, in which all assessments were tested for suitability in the children. The families and children assisted in designing the follow-up study to suit their needs and requirements.

Sample size

In total, 77 children were randomised to the DRIFT trial during 2003–6, of whom 69 survived until age 2 years. Based on a similar effect size documented with severe cognitive disability at age 2 years, a two-group continuity corrected chi-squared test with a 5% two-sided significance level had 80% power

to detect the difference in severe cognitive disability between a control group proportion of 59% and an odds ratio (OR) of 0.17 (i.e. an intervention proportion of 19.7%) when the sample size in each group is 28. With 60 infants (30 in each group), the power was 97% (with an alpha of 5%) to detect a mean cognitive difference of one standard deviation (SD) (commonly 15 points) between the DRIFT and standard treatment groups. It was anticipated that 45 UK children would be assessed in Bristol and 15 Polish children in Katowice, assuming a 90% follow-up rate. Those from Bergen and Glasgow would be sought if numbers were proving difficult to achieve.

Primary outcome

The primary outcome was cognitive disability at school age, expressed as a cognitive quotient (CQ). The British Ability Scales version three was used for children with a developmental age of ≥ 3 years. For children below this threshold, the Bayley Scales of Infant and Toddler Development version three was administered. The final scores were in the format of a cognitive developmental quotient (0 to 100+).

Secondary outcomes

- Cerebral visual function: parent-reported visual ability and parent-completed cerebral visual impairment (CVI) questionnaire.
- Sensorimotor disability: children were assessed using the Movement Assessment Battery for Children-2 (Movement ABC). Severity and numbers with cerebral palsy (CP) were also compared.
- Emotional/behavioural function: parent-completed Strengths and Difficulties Questionnaire (SDQ).
- Parent-reported education outcomes.
- Neurosurgical sequelae on structural brain magnetic resonance imaging (MRI).
- Preference-based measures of HRQoL: parents completed two generic measures of their child's HRQoL at 10-year follow-up, using the Health Utilities Index – 3 (HUI3) and the EuroQol-5 Dimensions, five-level version (EQ-5D-5L).
- Costs of initial hospitalisation and treatment during the neonatal period (including emergency transportation, periods of intensive care and readmissions based on hospital data in the Bristol cohort).
- Costs of subsequent health care in childhood (based on hospital data from the Bristol cohort).
- Health and social care costs and impact on family at 10-year follow-up (based on parent recall).
- Decision analysis model: a simple decision analytical model to estimate the cost-effectiveness of DRIFT compared with standard care from birth to age 18 years. The primary perspective was that of NHS and Personal Social Services in accordance with National Institute for Health and Care Excellence (NICE) guidance. In secondary analysis, we broaden the perspective to include education costs.

Research ethics

Ethics approval was granted by the NHS Health Research Authority (14/SW/1078).

Results

Between September 2015 and April 2016 families were contacted and asked to take part in the 10-year follow-up study. In two patients (two in DRIFT, zero in standard treatment), we were unable to find a contact address or number. This left 67 patients where the survival status was known. Of these, there were two deaths in the DRIFT arm and two deaths in the standard treatment arm, one patient declined (in the standard treatment arm) and 10 gave no response, leaving 52 available for assessment: 28 in the DRIFT arm and 24 in the standard treatment arm.

Among the 52 children available for follow-up assessments at 10 years, there were imbalances of gender and birthweight favouring the standard treatment group. There were 22 males in the DRIFT arm (79%) whereas the standard treatment arm had a lower proportion of males (63%). Birthweight was much higher in the standard treatment arm (mean 1322 g) than in the DRIFT arm (1102 g). We prespecified in the analysis plan that any baseline characteristics that differed by more than 10%/0.5 SDs would be adjusted for in a sensitivity analysis.

Cognitive disability

Given the larger than expected attrition rate, precision was lower than hoped and was exacerbated further by large SDs for the cognitive ability quotient. Despite this, results are in parallel with those at 2 years, with crude estimates giving weak evidence that the DRIFT intervention increases cognitive ability at 10 years ($p = 0.096$). After adjusting for the prespecified covariates of gender, birthweight and grade of IVH, this evidence was strengthened and indicated that children who were in the DRIFT arm of the trial, on average, had a CQ score of 23.5 points higher than those who received standard treatment ($p = 0.009$). This translates into a developmental cognitive advantage of 2.5 years.

Sensitivity analysis for primary outcome

The binary outcome, alive without severe cognitive disability, gave very similar results to the continuous CQ outcome [unadjusted OR 3.6, 95% confidence interval (CI) 1.2 to 11.0; $p = 0.026$ and adjusted OR 10.0, 95% CI 2.1 to 46.7; $p = 0.004$]. Both the unadjusted and adjusted model gave strong evidence to suggest that DRIFT had a positive impact on children's cognitive outcomes at 10 years. The number needed to treat was three.

Vision

Overall, the results show that those in the DRIFT arm were almost four times more likely to have a 'good' visual outcome than the standard treatment arm (adjusted OR 3.73); however, the *p*-value shows only very weak evidence to support this ($p = 0.136$). No difference was found in CVI mean score (–0.12, 95% CI –0.47 to 0.24; $p = 0.502$).

Sensorimotor disability

There was no difference in mean Movement ABC scores (–1.0, 95% CI –16.8 to 14.8; $p = 0.896$). Children in the DRIFT arm were 1.1 times more likely to have CP than those in the standard treatment arm (OR 1.10, 95% CI 0.36 to 3.35; $p = 0.862$). After adjustment for gender, birthweight and grade of IVH, this changed to 63% lower odds of CP in the DRIFT group (OR 0.37, 95% CI 0.07 to 2.00; $p = 0.249$); this is largely due to those in the DRIFT having less favourable baseline characteristics. Although the percentage of children with CP was higher in the DRIFT arm than in the standard treatment arm (61% vs. 58%, respectively), those in the DRIFT arm were less likely to have CP categorised as severe. After adjustment, those in the DRIFT arm were 32% more likely to be ambulant than those in the standard treatment arm (1.32, 95% CI 0.24 to 7.25; $p = 0.751$). However, given the large CI and *p*-value, the evidence to support this finding was not strong; the result could have simply happened by chance.

Emotional/behavioural function

There was no difference in mean SDQ score ($p = 0.584$).

Neuroimaging

There were no major differences relating to residual neurosurgical conditions needing referral. Residual catheter tracks were more often seen in the standard treatment group and in association with ventricular reservoirs.

Education outcomes

After adjustment, those in the DRIFT arm had lower odds (0.27) of special school attendance in the last 12 months than those in the standard treatment arm ($p = 0.059$).

Harms

Despite the excess secondary haemorrhages in the DRIFT group, the primary outcomes were better and the secondary outcomes were no worse than in the standard treatment group. It does not appear that secondary haemorrhages that occurred during the DRIFT procedure had a long-term detrimental effect. High-resolution structural brain MRI at 10 years showed no evidence of damage associated with insertion of the DRIFT irrigation catheters. There was no difference in ongoing neurosurgical problems between the treatment arms at age 10 years.

Cost of initial hospitalisation

Participants allocated to DRIFT had irrigation therapy for an average of 5.2 days at an estimated cost of £1513 per participant. Some of this initial cost of DRIFT was offset by the fact that fewer patients had reservoir procedures during the neonatal stay. The total mean costs of the neonatal stay were higher in patients who had DRIFT, but the CI was wide and included zero (unadjusted mean difference £6556, 95% CI –£11,161 to £24,273). The finding was sensitive to adjustment for covariates, particularly birthweight. After adjustment for birthweight, gender and IVH grade, estimated mean costs of neonatal care were lower in patients who had DRIFT although CIs were still wide and included zero (adjusted mean difference –£3056, 95% CI –£19,449 to £13,335).

Postnatal hospital admissions and total NHS secondary care costs

Participants allocated to DRIFT spent an average of 19.4 additional days in hospital up to age 2 years and an average of 26.6 additional days in hospital between age 2 years and 31 March 2016. Participants allocated to standard care spent fewer additional days in hospital (8.8 days at age 0–2 years; 18.5 days at age 2 years upwards). The unadjusted total costs of hospital care after the initial neonatal stay were higher in participants allocated to DRIFT (unadjusted mean difference £3413, 95% CI £12,408 to £19,234). This finding was very sensitive to adjustment for covariates, particularly gender and birthweight. After adjustment, the estimated mean cost among participants allocated to DRIFT was lower (adjusted mean difference –£9739, 95% CI –£27,558 to £8080).

Use of ambulatory health and social care at ten-year follow-up

There was little evidence of a difference in emergency and outpatient care in the last 12 months at 10-year follow-up. Participants in both arms of the trial reported an average of just over 0.4 visits to the emergency department and just over 2.8 outpatient clinic visits. The adjusted mean difference in costs was marginally higher in participants allocated to DRIFT (adjusted mean difference £2, 95% CI –£264 to £267). The costs of other ambulatory care during the last 6 months were higher in participants randomised to standard care (adjusted mean difference –£108, 95% CI –£596 to £380) but the CI was wide.

Family income, expenses and child's educational needs

Overall, a similar proportion of parents/carers were employed at the 10-year follow-up. However, a lower proportion of households of participants who received DRIFT had benefits as their main source of income (adjusted OR 0.23, 95% CI 0.04 to 1.22), although the CI included 1. A higher percentage of parents of participants in the standard treatment arm reported that their child attended a special unit or special school (adjusted OR 0.13, 95% CI 0.02 to 0.82). Owing to the high cost of special schooling, this is potentially economically important; the adjusted mean difference in estimated annual school costs was –£5321, 95% CI –£9772 to –£870.

Health-related quality of life

In adjusted analyses, both the EQ-5D-5L and HUI3 scores of HRQoL tended to be higher in survivors who were allocated to DRIFT than in those who were allocated to standard care. However, the CIs around the adjusted mean differences in EQ-5D-5L score (0.06, 95% CI –0.11 to 0.22) and HUI3 score (0.13, 95% CI –0.09 to 0.35) included zero.

Decision analytical model

DRIFT has the potential to be a cost-effective intervention at current NICE thresholds. Exploratory analysis using a simulation model to interpolate and extrapolate costs and outcomes to age 18 years indicated that the additional benefit [8.96 quality-adjusted life-years (QALYs) vs. 8.33 QALYs] in the DRIFT arm justifies the higher NHS and social service costs (£112,341 vs. £102,611). The incremental cost-effectiveness ratio (£15,621) was below the NICE thresholds of £20,000 to £30,000 per QALY and the incremental net monetary benefit (£2711) was positive. When education costs are included or using costs and utility scores adjusting for gender, IVH grade and birthweight, DRIFT has the potential to both save money and improve outcomes for children.

Conclusions

Implications for health care

The school-age follow-up of the DRIFT trial strengthens the evidence of benefit found at 2 years and adds further evidence of safety of the intervention. We can conclude that DRIFT improves cognitive function when taking into account birthweight, IVH grade and gender. The cost of the intervention is moderate; DRIFT has the potential to be cost-effective. In some scenarios, DRIFT may save money and improve outcomes owing to the possible reduction in the need for special education.

Recommendations for research

The role of any NHS implementation of DRIFT, ideally in a few specialised tertiary centres, delivered through the existing neonatal operational delivery networks, will need to be studied prospectively in a multicentre trial. As well as measures of cognition and functional measures, the data from the 10-year outcomes indicate that any future studies should continue to collect data on vision, motor skills and education, given the trends seen in the secondary outcomes that the study was not powered to address.

A larger proportion of infants with PHVD is now extremely immature. Further refinements in DRIFT may need to be studied in this very immature group of patients.

For infants with parenchymal infarction in addition to PHVD, there is scope to supplement DRIFT with novel interventions to promote brain tissue repair in the future.

Trial registration

This trial is registered as ISRCTN80286058.

Funding

Funding for the 10-year follow-up study was provided by the Health Technology Assessment programme of the National Institute for Health Research. The DRIFT trial and 2-year follow-up was funded by Cerebra and the James and Grace Anderson Trust.

Chapter 1 Introduction

Scientific background

Haemorrhage into the ventricles of the brain is one of the most serious complications of preterm birth, despite improvements in the survival of preterm infants. Large intraventricular haemorrhage (IVH) carries a high risk of neurological disability and, by causing a progressive obliterative arachnoiditis at the basal cisterns and the outlet foramina of the fourth ventricle, disturbs the flow and absorption of cerebrospinal fluid (CSF).[1] This leads to post-haemorrhagic ventricular dilatation (PHVD).

Severe IVH with PHVD is a neurological complication seen in preterm infants, with significant neurodisability in survivors. Infants most at risk (those born at < 32 weeks of gestation and with a birthweight of < 1500 g) have high rates of grade 3 or 4 IVH (around 6%), estimating approximately 800–900 new cases of grade 3 and 4 IVH annually in the UK.[2] Preterm birth rates are rising[3] and survival rates of extremely preterm infants continue to improve;[4] therefore, it can be projected that the number of infants affected by PHVD will increase in the future.

In the US National Institute of Child Health and Development Neonatal Network, one-third of infants with birthweights of < 1000 g develop IVH, and, of those, about 10% require implantation of a ventriculoperitoneal (VP) shunt for PHVD.[5] In a European study,[6] 29% of all preterm infants with severe IVH required implantation of a VP shunt.

Of children with PHVD, 40% will develop cerebral palsy (CP) and approximately 25% will have multiple disabilities. The National Institute of Child Health and Development study,[5] the largest of its kind, studied > 1000 preterm infants at 18–22 months corrected age with severe grade IVH (grade 3 and 4), of whom almost 25% had PHVD (defined as requirement for a VP shunt). They demonstrated significant risk of cognitive impairment with PHVD with a median Mental Development Index (MDI) score 20 points lower in children with PHVD than in those with severe grade IVH without PHVD. The median MDI in children with grade 3 IVH and PHVD was 61 points, and in those with grade 4 IVH and PHVD it was 50 points. Overall, 68% of children with severe grade IVH and PHVD had moderate cognitive impairment [MDI below two standard deviations (SDs)] and 41% had severe cognitive impairment (MDI below three SDs). Furthermore, 70% of children with PHVD had CP and 30% had visual impairment. The presence of a haemorrhagic parenchymal infarction, in addition to PHVD, increased the risk of CP to between 80% and 90%.[5] A logistic regression analysis[7] of factors affecting school performance at 14 years of age in a cohort of 278 preterm infants showed that peri- or intra-ventricular haemorrhage was the primary risk factor for special education. Intraventricular blood and ventricular expansion have adverse effects on the immature periventricular white matter by a variety of mechanisms including physical distortion, raised intracranial pressure,[8] free radical generation facilitated by free iron[9] and inflammation.[10]

The prevalence of visual defects is higher among prematurely born children than children born at term[11] and, in particular, the spectrum of vision problems known collectively as 'cerebral visual impairment' (CVI) is a recognised complication of preterm brain injury, particularly if involving the periventricular white matter.[12] Visual functions correlate with neurodevelopmental outcome and brain volume in preterm infants.[13] Severe CVI is the leading cause for children being registered as blind in the UK[14] and the developed world and may additionally be associated with ocular, optic nerve or refractive problems that cause further impairment. Less severe CVI can damage visual skills and have an important effect on school performance and tasks of everyday life.[15] Clinical assessment of CVI is difficult before the age of 5 years; however, a recent study[16] found evidence of CVI in 89% of children with known central nervous system damage.

Every preterm infant with severe CP or severe cognitive or visual impairment will require lifelong parental and social care. The cost to society resulting from the complications of prematurity is significant. Based on

2003 US figures,[17] the estimated lifetime costs per infant with CP, severe cognitive impairment or blindness is £614,000, £675,000 and £400,000, respectively. Data from the UK EPICure study[18] indicate that, by 11 years of age, the annual health and social service costs of children with serious neurodevelopmental disability are almost double those of children without disability (£1225 vs. £695, respectively). This adds a significant additional economic burden on the NHS and social care. This estimate excludes the substantial economic burden on parents/carers, special educational services and other public funds. A recent confidential inquiry into premature death in adults with learning disabilities in England highlighted the complex lifelong health and social care needs of individuals with learning disabilities. On average, each person with learning difficulties had five additional medical conditions and received seven prescription medications; 64% of individuals lived in residential care homes, the majority with 24-hour paid-nursing care.[19]

Reducing the rate of VP shunt insertion has been an important long-term objective in the management of IVH and PHVD. The large amount of blood and protein in the CSF combined with the small size and instability of the patient makes early VP shunt surgery impossible. Shunt implantation at the generally accepted weight threshold of 2 kg, usually around term age, is still associated with a higher infection and malfunction rate.[20,21] Unfortunately, several interventions have failed to reduce the need for shunt insertion, and no intervention has reduced neurodisability rates as a result of PHVD. Repeated lumbar punctures (LPs) are often ineffective at allowing removal of enough CSF. Direct ventricular puncture through the anterior fontanelle leads to needle track damage through the brain parenchyma. Repeated LPs or ventricular taps do not reduce the risk that a shunt will eventually be required; they have no effect on neuromotor impairment and are associated with a significant risk of ventriculitis (at 7% in the Ventriculomegaly Trial[22–24]).

In an effort to control PHVD by reducing CSF production, the International PHVD Drug Trial Group[25] investigated the effects of acetazolamide and frusemide in a randomised trial in 1998. Not only did these drugs not lead to an improvement in neuromotor development or CSF diversion requirements, but the data monitoring committee stopped the trial because of worse outcome in the treated group.

In practice, once two LPs or one ventricular tap have been necessary to control the ventricular dilatation, insertion of a ventricular reservoir is preferred. A reservoir provides an easy and safe route for repeated aspiration of ventricular CSF, with low infection rates.[6,26] Insertion requires an anaesthetic in a neurosurgical theatre and can be safely performed in babies weighing < 800 g. This is a temporary measure and allows repeated drainage of CSF until the need for permanent CSF diversion can be established through VP shunt insertion. The most commonly encountered risks after reservoir insertion are infection and malfunction.[6,26]

The timing of insertion of a ventricular reservoir remains controversial. In a retrospective study,[27] early insertion, before crossing the 97th + 4 mm ventricular index line, was associated with lower rates of VP shunt insertion. The Early vs Late Ventricular Intervention Study (ELVIS; ISRCTN43171322)[28] randomised between the two treatment thresholds, with death or shunt dependence and disability at 2 years being the main treatment outcomes. The trial has ended but results are as yet not published. Endoscopic lavage is a new neurosurgical intervention used for PHVD in which the ventricles are washed out under direct vision using a small endoscope. A small feasibility study[29] using historical controls seemed promising in terms of safety and reducing the need for VP shunt insertion. Long-term outcomes are not known and the research group concluded that this intervention needs to be tested objectively in a randomised controlled trial (RCT).

In summary, no medical or surgical intervention for PHVD has objectively demonstrated either a reduction in the need for a permanent VP shunt or a reduction in death or neurodisability. Current practice in the UK consists of repeated LPs followed by insertion of a ventricular reservoir to enable regular tapping to reduce pressure. The complications are a combined infection and device failure rate exceeding 10%, as highlighted above.

Drainage, irrigation and fibrinolytic therapy (DRIFT)[30–32] is a surgical approach that was developed because of the unsatisfactory results of other treatments. The objectives are to reduce pressure and distortion early and to remove proinflammatory cytokines and free iron from within the ventricles. The procedure involves

insertion of right frontal and left occipital ventricular catheters under anaesthesia. Tissue plasminogen activator (TPA), a fibrinolytic, is injected intraventricularly at a dose that is insufficient to produce a systemic effect and this is left for approximately 8 hours. Under continuous intracranial pressure monitoring, the ventricles are irrigated by artificial CSF through the frontal catheter. The occipital ventricular catheter is simultaneously connected to a sterile closed ventricular drainage system and the height of the drainage reservoir adjusted to increase or decrease drainage to maintain an intracranial pressure below 7 mmHg and a net loss of 60–100 ml of CSF per day. The drainage fluid initially looks like cola but gradually clears, at which point irrigation is stopped and the catheters removed. This commonly takes 72 hours but can take up to 7 days.

After initial feasibility testing showed that DRIFT was technically possible and promising,[30] the DRIFT randomised trial started recruiting in 2003. Babies were elegible for the study if they were born preterm, they had had IVH and their cerebral ventricles had expanded over predetermined limits. With parental consent, 77 babies were randomised in Bristol, Katowice (Poland), Glasgow or Bergen (Norway) to either DRIFT or standard therapy, which consisted of non-surgical conventional management (LPs to control excessive expansion and pressure symptoms). If repeated LPs were needed, a ventricular reservoir was surgically inserted to facilitate tapping CSF.

There were no differences in the short-term outcomes: need for VP shunt or death at 6 months.[33] At 2 years post term, severe disability or death was significantly reduced in the DRIFT group.[32] There was an important decrease in severe cognitive disability (Bayley MDI three SDs below the mean) from 59% to 31% [adjusted odds ratio (OR) 0.17, 95% confidence interval (CI) 0.05 to 0.57] and the difference in median MDI was > 18 points. Sensorimotor disability remained substantial in both treatment groups at 2 years: overall, 48% were unable to walk, 20% were unable to communicate and 9% had no useful vision. Severe sensorimotor disability was less common in the DRIFT group but without reaching statistical significance.

Although short-term neurodevelopmental measures are essential in the initial management of perinatal interventions, longer-term measures provide far greater validity in assessing long-term functioning (and the medical, societal and financial implications of these). Therefore, the main objective of the follow-up study of the DRIFT trial was to assess if the cognitive advantage seen at 2 years with DRIFT continued through to school age. Secondary objectives were to assess the long-term visual and motor function, emotional and behavioural difficulties, brain structure and quality of life (QoL) as well as the cost-effectiveness of DRIFT.

Chapter 2 Trial design and methods

Trial design

The DRIFT study was originally conducted in 2003–6 as a multicentre RCT that recruited premature infants with PHVD. Infants were randomised to receive standard treatment or surgical DRIFT. Now, 10 years on, the children have been followed up to investigate the difference in cognitive ability at school age between the two groups.

The school-age follow-up of the DRIFT trial was designed in partnership with the children and parents who attended a small feasibility study in Bristol. Families gave their input into the methods for initial contact, parent and participant literature, feedback on the study assessments and the timing of the assessments so as to not distract from school attendance. These families gave valuable advice on how to make the assessment day engaging for the children. Mr Steven Walker-Cox and his son, who have lived experience of prematurity and DRIFT, helped to write the letter of invitation, study materials and information leaflets and consent/assent forms for parents and children. They also contributed to the research ethics application.

Research ethics approval (14/SW/1078) was granted by the National Research Ethics Service Committee South West-Central Bristol prior to commencing the school-age follow-up. The University of Bristol acted as sponsor.

Participants

Children previously enrolled in, and randomised to, the DRIFT trial between 2003 and 2006 were from Bristol, Katowice, Glasgow or Bergen.

Children were eligible for the DRIFT trial if they matched all of the following criteria:

- IVH documented on ultrasonography.
- Age of no more than 28 days.
- Progressive dilatation of both lateral ventricles with each side:
  - ventricular width 4 mm over the 97th centile (a)

  OR
  - anterior horn diagonal width 4 mm (1 mm over 97th centile) (b)
  - thalamo-occipital distance 26 mm (1 mm over 97th centile) (c)
  - third ventricle width 3 mm (1 mm over 97th centile) (d)

  OR
  - measurements above (a) or (b–d) on one side combined with obvious midline shift indicating a pressure effect.

Exclusion criteria were:

- prothrombin time of > 20 seconds
  OR
- accelerated partial thromboplastin time of > 50 seconds or a platelets count of < 50,000/μl.

Interventions

DRIFT was developed as a surgical approach for reducing iron and proinflammatory cytokines from CSF and reducing pressure and distortion early. The procedure involves insertion of right frontal and left occipital ventricular catheters under anaesthesia. TPA, a fibrinolytic, is injected intraventricularly at a dose that is insufficient to produce a systemic effect and this is left for approximately 8 hours. Under continuous intracranial pressure monitoring, the ventricles are irrigated by artificial CSF through the frontal catheter. The occipital ventricular catheter is simultaneously connected to a sterile closed ventricular drainage system and the height of the drainage reservoir adjusted to increase or decrease drainage to maintain an intracranial pressure below 7 mmHg and a net loss of 60–100 ml of CSF per day. The drainage fluid initially looks like cola but gradually clears, at which point irrigation is stopped and the catheters removed. This commonly takes 72 hours but can take up to 7 days.

Standard treatment consisted of up to two LPs to drain CSF followed by insertion of a ventricular reservoir with regular tapping of CSF to reduce ventricular distension to within the specified dimensions.

Primary outcome

Cognitive disability at school age

Cognitive assessments were undertaken by child psychologists. The British Ability Scales version three (BAS III)[34] (see *Appendices 1* and *2*) was used for children with a developmental age of ≥ 3 years. For children who did not meet this threshold, the Bayley Scales of Infant and Toddler Development (BSID III)[35] (see *Appendix 3*) was administered. The final scores were in the format of a cognitive developmental quotient (0 to 100+). The primary analysis was based on the cognitive scores of surviving children, although a sensitivity analysis that included children who died (as a result of their disability) was also carried out, in which the cognitive development quotient for these children could reasonably be assumed to be zero.

Secondary outcomes

Cerebral visual function

For the main visual outcomes, we used parent-reported data as they were available for the majority and could be compared with the 2-year outcomes. Parents were asked whether their child had vision that was of 'No concerns', 'Normal with Correction' or 'Useful but not fully correctable' or was 'Blind or perceives light only'. A binary outcome was created that split these into a good visual outcome (no concerns/normal with correction) or a poor visual outcome (useful but not fully correctable/blind or perceives light only). A 23-question assessment of CVI was also carried out by the vision specialists[36] (see *Appendix 4*). A mean score was created from all available questions and analysed between the groups. In the case of those who attended assessments in Bristol, vision specialists directly assessed a range of visual functions including visual acuity, visual field, eye movements and vision processing skills.

Sensorimotor disability

Assessments of motor function and disability were made by a paediatric physiotherapist. Children were assessed using the Movement Assessment Battery for Children-2 (Movement ABC)[37] (see *Appendix 5*). As well as this assessment, the number and severity of CP were also compared between the two groups.

Emotional/behavioural function

Parents were asked to fill out the Strengths and Difficulties Questionnaire (SDQ)[38] (see *Appendix 6*), which assesses how their child behaves in various circumstances; their final score classifies them as having 'normal behaviour' or 'abnormal behaviour'.

Methods

Sample size

In total, 77 children (54 from Bristol and 20 from Poland, two in Glasgow and one in Bergen) were randomised to the DRIFT trial during 2003–6, of whom, 69 survived until the age of 2 years. Based on a similar effect size documented with severe cognitive disability at age 2 years, a two-group continuity-corrected chi-squared test with a 5% two-sided significance level would have 80% power to detect the difference in severe cognitive disability between a control group proportion of 59% and OR of 0.17 (i.e. an intervention proportion of 19.7%) when the sample size in each group is 28. With 60 infants (30 in each group), we would have 97% power (with an alpha of 5%) to detect a mean cognitive difference of one SD (commonly 15 points) between the DRIFT and control groups.

It was anticipated that 45 UK children would be assessed in Bristol and 15 Polish children in Katowice, assuming a 90% follow-up rate. Those from Bergen and Glasgow would be sought if numbers were proving difficult to obtain.

Randomisation

A computer-generated randomisation scheme was used to assign infants to treatment groups in a 1 : 1 ratio.[31] Given that the trial was taking place in four different centres, the randomisation process was stratified by centre in blocks of eight, 10 or 12. Each infant was allocated to treatment using sequentially numbered, doubled-up envelopes that each contained either a 'DRIFT' or 'standard treatment' card.

Envelope preparation and random number allocation were carried out using StatsDirect software (StatsDirect, Altrincham, UK) by a research assistant not involved in enrolment or treatment. Patients were enrolled by one of the neonatologist investigators, all of whom, at that stage, were blind to treatment allocation. When the informed consent process was completed and signed, the next trial envelope was opened and treatment allocation confirmed.

Blinding

Owing to the nature of the intervention delivery, once the envelope had been opened, it was not possible to blind practices/parents to their allocation to either DRIFT or standard treatment. At 10 years, all investigators (child psychologists, visual specialists, etc.) were blinded to treatment allocation and grade of IVH as these were not apparent. All analysts (statisticians/health economists) were blinded to treatment allocation as far as possible. Given that the results were published at 2 years in favour of the DRIFT arm, it could be argued that the statisticians could have easily assumed which group was which. However, they continued the analysis with groups A and B, and only the senior statistician was shown the allocation in order for a draft abstract to be written (February 2016). As soon as the deaths were analysed in April 2016, the statistician felt that she could no longer be classed as 'blinded'.

Statistical methods

The main statistical analyses were prespecified using a statistical analysis plan (SAP) and the health economics using a health economics analysis plan. The final version of the SAP was accepted and agreed on 6 April 2016. Although a short interim analysis was completed in February 2016 on children assessed up to that point, no major changes were made to the SAP after this time. Given the large differences seen between the two groups at 2 years, it was difficult for the statistician to remain blinded. Final data analysis started on 6 April 2016 and finished in October 2016. Stata® 14.1 (StataCorp LP, College Station, TX, USA) was used for all statistical and health economic analyses in this trial. Binary outcomes were presented as *n* (%) while continuous outcomes were presented as mean (SD)/median [interquartile range (IQR)], as appropriate. For secondary and subgroup analyses, emphasis was placed more on descriptive statistics than on *p*-values. An informal Bonferroni technique was applied when interpreting *p*-values; alpha divided by four secondary and seven subgroup analyses: $0.05 \div 11 = 0.0045$. The *p*-values for exploratory outcomes were interpreted with extreme caution.

Primary analysis

Cognitive assessments were undertaken by child psychologists. The BAS III (see *Appendices 1* and *2*) was used for children with a developmental age of ≥ 3 years. For children who did not meet this threshold, the BSID III (see *Appendix 3*) was administered. The final scores were in the format of a cognitive developmental quotient; therefore, a continuous variable. The primary analysis was conducted using the intention-to-treat (ITT) principle using linear regression. The DRIFT team had determined, a priori, the variables that they believed might confound the final result. Compatible with the previous investigation at 2 years, adjustments were made for grade of IVH, birthweight and gender.

Null hypothesis: the average score for cognitive disability is the same for both groups.

Alternative hypothesis: the average score for cognitive disability is different between the groups.

As stated in the protocol, we were also interested in the proportion of children alive and without severe cognitive disability (BAS III score of < 3 SDs for age) at 10 years compared with those with severe disability or who had died owing to disability. To avoid splitting the 5% alpha between two primary outcomes, it was added as a sensitivity analysis.

The primary hypothesis was that neurosurgical DRIFT would reduce severe cognitive impairement in children assessed at school age. This was measured using cognitive ability tests. At 10 years children were assessed using the BSID III (for those anticipated to be performing at below the 3 years level), BAS III early years (for those anticipated to be performing between the 3 years and 7 years levels) or BAS III school age scoring system (for the remainder).

A quotient score was generated in the following way:

1. Cognitive and language developmental age-equivalent (DAE) scores yielded from the BSID III were collected and averaged to produce an overall DAE.
2. DAE scores from the BAS III early years assessment were averaged across the core scales (verbal comprehension, picture similarities, naming vocabulary, pattern construction, matrices and copying) to produce an overall DAE. DAE scores of 'less than 3' were given an age of 2 years and 11 months.
3. DAE scores from the BAS III school age assessment were averaged across the core scales (recognition of designs, word definitions, pattern construction, matrices, verbal similarities and quantitative reasoning) to produce an overall DAE. DAE scores of 'less than 5' were given an age of 4 years and 11 months.
4. All DAE scores were then divided by the child's actual age and then multiplied by 100 to achieve the child's 'Cognitive Quotient Score'.

Secondary analysis

Visual assessment

Visual assessments consisted of parent-reported outcomes and assessments carried out by vision specialists. The main visual outcome was parent reported, as used at 2 years. For the main visual outcomes, parents were asked whether their child had vision that was of 'No concerns', 'Normal with correction' or 'Useful but not fully correctable' or was 'Blind or perceives light only'. A binary outcome was created that split these into a good visual outcome (no concerns/normal with correction) or a poor visual outcome (useful but not fully correctable/blind or perceives light only). Differences in the proportions of visually impaired children between the groups were assessed using logistic regression.

A 23-question assessment of CVI was also administered by the vision specialists.[36] An average score was derived from the answers to all available questions and analysed between the groups. This was based on applicability of the questions as in some cases (such as blindness) the questions were not appropriate.

After analysis of the data, it was decided that the child who was blind should not have been given this visual assessment and, therefore, was removed from the analysis. Originally, we had prespecified that we would use a linear regression model to compare CVI scores. However, on inspection of the data, it became clear that the data were negatively skewed (with 22% of children scoring the maximum score of 5; see *Results*). Therefore, both a comparison of means and a non-parametric test were carried out to assess if interpretation was similar.

Motor function and disability

Assessments of motor function and disability were made by a paediatric physiotherapist. Children were assessed using the Movement ABC[37] (see *Appendix 5*). Scores were then classified according to test recommendations as mild (green), moderate (amber) or severe (red). These was analysed using ordinal logistic regression. Children who could not complete the task owing to CP were automatically placed in the severe category (as prespecified in the analysis plan).

Cerebral palsy

The number of diagnosed cases of CP was also compared between the two groups using logistic regression. At 10 years, children were either diagnosed with CP or not diagnosed with CP. Severity of CP was classified using the Gross Motor Function Classification System (GMFCS):[39]

- Level 1 – children walk at home, school, outdoors and in the community and can climb stairs without the use of a railing.
- Level 2 – children walk in most settings and climb stairs holding onto a railing.
- Level 3 – children walk using a hand-held mobility device in most outdoor settings.
- Level 4 – children use methods of mobility that require physical assistance or powered mobility in most settings.
- Level 5 – children are transported in a manual wheelchair in all settings.

Children without CP or with CP level 1 or 2 were classified as ambulant.

Emotional/behavioural function

Parents were asked to fill out the SDQ[38] (see *Appendix 6*), which assesses how children behave in various circumstances, with the final score being used to classify a child's behaviour as 'normal or 'abnormal. Differences in the overall score between the two groups were assessed using linear regression. Differences between subscores were assessed using the Mann–Whitney *U*-test.

Sensitivity analysis

Several sensitivity analyses were conducted to test robustness of the results from the statistical analyses and, in some cases, increase understanding of the relationship between the dependent and independent variables. All were performed in the same way as the primary analysis. All sensitivity analyses were prespecified before final analysis began. At 2 years, a binary outcome was used; therefore, the team decided to duplicate this at 10 years (removing the sensorimotor element). Accounting for deaths in a trial that focuses on neurodevelopmental outcomes is a hotly debated topic.[40] The team felt that various methods should be included as sensitivity analyses to ensure that results were consistent. Death was included in four different ways. Initially, only the three deaths post 2-year follow-up were included and given a score of 0. The team felt confident that this was appropriate given that these deaths could directly be linked to the child's disability. These three deaths were also included in a binary outcome that combined them with those who had severe disability and an ordinal outcome of five categories where death was considered the worst outcome. Last, all deaths were included in the binary outcome (including the eight deaths before 2 years). Although this outcome reflects that used at 2 years it includes deaths which were unrelated to cognitive ability. Cause of death before 2 years was difficult to determine and many were linked to neonatal complications.

Sensitivity analyses included:

- cognitive ability quotient (using BSID III/BAS III age-equivalent scores), including deaths as 0
- proportion alive and without severe cognitive disability at 10 years versus severely disabled/died owing to disability
- grading of disability (mild/moderate/severe/dead) as an ordinal outcome
- imputation of missing data at 10 years (details below)
- cognitive ability quotient for the Bristol cohort only.

Similarly to Biering *et al.*,[41] we chose to carry out five different imputation models, summarised in *Table 1*, that made different assumptions about the data, particularly death.

Initially, baseline variables were assessed to determine if they were predictive of missingness in the primary outcome using logistic regression. We then established, using linear regression, if they were appropriate predictors of the primary model. Any baseline variables associated with the primary outcome of interest or its missingness were added to the imputation model to inform the imputation process.

Using Stata 14's 'mi impute chained' function, we created 40 imputations and used predictive mean matching, as regression produced inappropriate imputations. A random seed of 65,898 was chosen for all models. All children lost to follow-up were assumed to be alive, as imputing a death indicator variable proved impossible.

Subgroup analysis

Subgroups were used to test whether or not the effects of the DRIFT intervention were more pronounced in certain subgroups of children. Although underpowered, tests of interaction between the dichotomised variables and treatment therapy were carried out to test whether or not treatment effect differed between subgroups. These interaction terms were added to the primary analysis model. All subgroup analyses were prespecified in the analysis plan apart from maternal education.

Subgroup analyses included:

- gestation (≥ 28 weeks vs. < 28 weeks)
- grade of IVH (grade 3 vs. 4)
- age of randomisation (day 1–20 vs. ≥ 21 days)
- unilateral versus bilateral dilatation on ultrasonography at randomisation
- gender
- pre- and post-enhanced vigilance in 2006
- maternal education (post hoc).

TABLE 1 Multiple imputation assumptions

Assumption	Deaths		Lost to follow-up
	Pre 2 years of age	Post 2 years of age	
1	✗	CQ = missing, NI	CQ = missing, NI
2	✗	CQ = missing, death = 1	CQ = missing, death = 0
3	✗	CQ = 0, NI	CQ = missing, NI
4	✗	CQ = 0, death = 1	CQ = missing, death = 0
5	✗	✗	CQ = missing, NI

CQ, cognitive quotient; death, indicator variable of death; NI, no indicator variable for death; ✗, removed from the imputation model.

Exploratory analyses

Although added before final analysis in April 2016, these were not prespecified in the trial protocol; therefore, they are only exploratory analyses and should be interpreted with this in mind.

- Educational outcomes:
  - mainstream schooling versus special school
  - special educational needs (SEN) support, yes/no
  - Key Stage 1 (KS1) scores
  - Key Stage 2 (KS2) scores
  - neurosurgical interventions after the neonatal period.
- Proportion with reservoirs.
- Shunt, yes/no (as assessed at 6 months).
- Death, yes/no (as assessed at 6 months and 2 years).

Neuroimaging

At 10 years, children assessed in Bristol who consented, and had no contraindications, to magnetic resonance imaging (MRI) were eligible for structural brain MRI.

Structural MRI scans were acquired on a 3 tesla Siemens Skyra scanner (Erlangen, Germany) with the use of a 32-channel radiofrequency head coil using 3D full volume T1-weighted inversion recovery gradient echo. Magnetisation-Prepared Rapid Gradient-Echo sequence (MP-RAGE) was also acquired in the sagittal plane, comprising 192 slices; repetition time: 1900 ms; time of echo: 2.2 ms; 0.9 mm isotropic voxel; matrix: 128 × 128. T2 Turbo Spin Echo Axial plane time to acquisition: 2:53; voxel size: 0.4 × 0.4 × 3.0 mm; 40 slices.

Participants were scanned after parental consent, participant assent and a safety check. They were excluded if contraindications to MRI were identified or if travel to Bristol was not possible.

Scans were assessed blinded to treatment allocation in one sitting by a team of three neonatal specialists with neuroimaging interests (ASC, AW, KL) and two neurosurgeons (IP, KA). Each scan was classified by consensus as follows:

- residual catheter tracts visible (frontal or occipital)
- parenchymal lesions
- ventricular reservoir in situ
- VP shunt in situ
- evidence of possible active hydrocephalus (dilated ventricles)
- residual clinical condition requiring neurosurgical referral.

Chapter 3 Trial results

Participant flow

Figure 1 shows the layout of the trial and the different levels of drop-out and analysis. At 2 years' follow-up there had been eight deaths but no loss to follow-up. At 10 years' follow-up, four more deaths had occurred as well as two children who could not be traced, one who declined to participate in the follow-up and 10 non-responders.

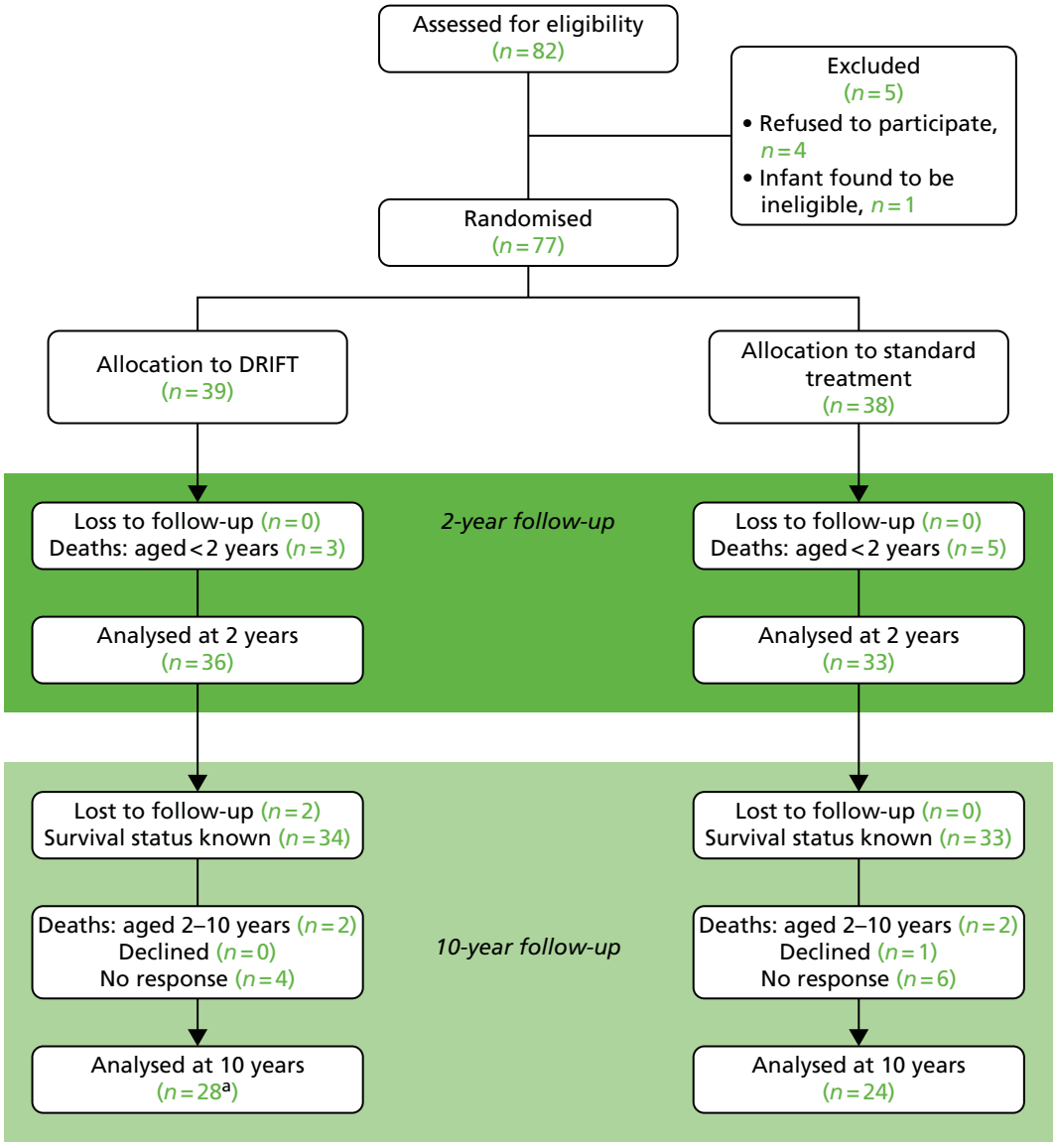


FIGURE 1 DRIFT CONSORT diagram. CONSORT, Consolidated Standards of Reporting Trials. a, One child did not complete a cognitive text and, therefore, could not be included in the primary analysis.

Recruitment

Originally, when the trial began, 77 babies were recruited to either receive DRIFT or standard treatment ($n = 39$ and $n = 38$, respectively). During this period, the trial was temporarily stopped by the Data Monitoring Committee, which was concerned by the rate of secondary haemorrhages; however, the trial was allowed to continue with increased vigilance. After a further 6 months, an a priori interim analysis was performed and the trial was closed owing to the low chance of seeing a significant result in the primary outcome – reduction in shunt surgery/death.[31] These children were then followed up and underwent numerous tests at approximately age 2 years.[32] Overall conclusions were that the 6-month time point was too soon after randomisation to be able to evaluate the intervention, while the 2-year follow-up showed promising results in favour of the DRIFT intervention. At 2 years' follow-up there had been three deaths in the intervention arm and five in the standard treatment arm. The intervention appeared to reduce severe cognitive/sensorimotor ability or death (adjusted OR 0.25, 95% CI 0.08 to 0.82). At this time point, all of the parents consented to take part, giving a sample size of 77 (including the eight deaths).

Approximately 8 years later (between September 2015 and April 2016), the parents were then contacted and asked to take part in the 10-year follow-up study. Unfortunately, trial investigators were unable to find a contact address or telephone number for two patients (in the DRIFT arm). This left 67 patients whose survival status was known. Of these, two patients in the DRIFT arm and two patients in the standard treatment arm died, one patient declined to participate (in the standard treatment arm) and 10 gave no response, leaving 52 available for assessment (see *Figure 1*). The death certificates confirmed that two deaths were due to the patient's disability; in the other two cases, death certificates (one per arm) were not available, so the cause of death was assumed to be disability based on these participants' low scores at 2-year follow-up.

For the primary outcome, we obtained a cognitive score for 51 children: 27 in the DRIFT arm and 24 in the control arm. The distribution of patients across centre and gender can be seen in *Figure 2*. In the sensitivity analysis (substituting scores of 0 for those who died post 2 years), we had 29 and 26 for DRIFT and standard treatment, respectively.

Baseline data

There were 77 patients who were randomised to the DRIFT trial; baseline comparisons are shown in *Table 2*. The team prespecified in the analysis plan that any baseline characteristics that differed by > 10%/0.5 SDs would be adjusted for in a sensitivity analysis. Only gender showed an imbalance of this magnitude at

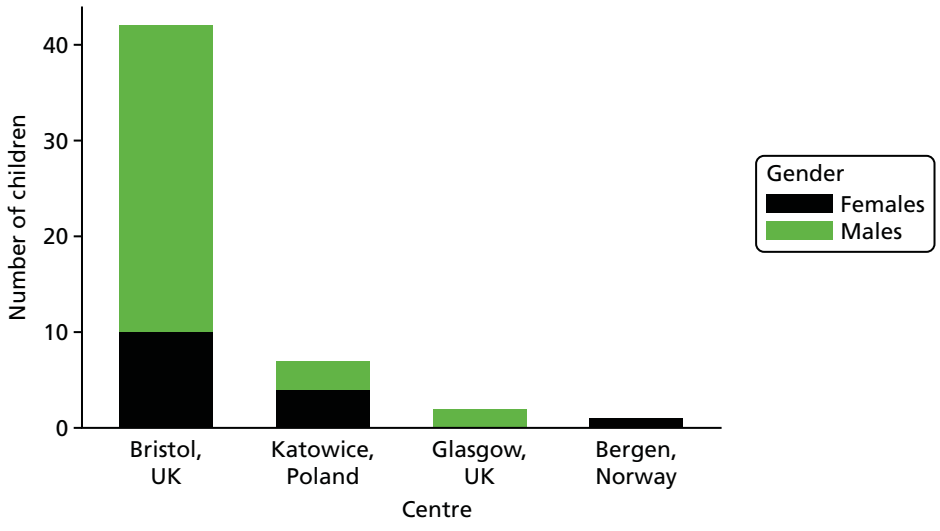


FIGURE 2 Assessment of children at 10 years, by centre.

TABLE 2 Baseline characteristics

Characteristic	DRIFT		Standard treatment	
	N	Mean (SD) or *n* (%)	*N*	Mean (SD) or *n* (%)
Total number of participants	39		38	
Centre				
Bristol, UK	39	27 (69)	38	27 (71)
Katowice, Poland		10 (26)		10 (26)
Glasgow, UK		1 (3)		1 (3)
Bergen, Norway		1 (3)		0 (0)
Sociodemographics at birth				
Age at randomisation (days)	39	19.18 (4.73)	38	18.47 (4.95)
Gender: male[a]	39	29 (74)	38	24 (63)
Median IMD 2015[b] (IQR)	22	23.50 (29.00)	23	20.00 (24.00)
Clinical characteristics at birth				
Birthweight (g)	39	1104.08 (346.23)	38	1251.21 (468.34)
Gestation (weeks)	39	27.69 (2.64)	38	28.21 (2.89)
Grade of IVH: 4	39	20 (51)	38	19 (50)
Maternal age at birth (years)	17	28.24 (6.70)	19	27.47 (6.06)

IMD, Index of Multiple Deprivation.
a Difference of 10%/0.5 SDs or higher between the groups.
b English IMD 2015 scores, UK Data Service Census Support, http://geoconvert.mimas.ac.uk/ (accessed 5 September 2016). Higher scores indicate higher levels of deprivation. IMD based on the children's home postcode at birth for those residing in England only.

baseline; therefore, this sensitivity analysis was removed (given that this was already a prespecified covariate). Birthweight showed moderate imbalance, approximately 0.36 SDs.

Among the 52 children available for follow-up assessments at 10 years, there were imbalances in gender and birthweight (*Table 3*). There were 22 males in the DRIFT arm (79%), whereas the standard treatment arm had a lower proportion of males (63%). Birthweight was much higher in the standard treatment arm (mean 1322 g) than in the DRIFT arm (1102 g). After including the three deaths (used in the sensitivity analysis of the primary analysis), this reduced the imbalance in gender to 9% and all other balances/imbalances remained.

Among those assessed at 10 years, secondary haemorrhages were experienced by 29% of the DRIFT arm compared with 13% of the standard treatment arm. On average, at 10 years, mothers in the DRIFT arm had a higher education level than those in the standard treatment arm (*Table 4*). Unfortunately, this was not recorded at baseline, so we cannot be sure where this sits within the causal pathway (i.e. whether this is a confounding factor or determined as a result of the intervention).

TABLE 3 Baseline characteristics of those assessed at 10 years, by trial arm

	Trial arm			
	DRIFT		Standard treatment	
Characteristic	*N*	Mean (SD) or *n* (%)	*N*	Mean (SD) or *n* (%)
Total number of participants	28		24	
Centre				
Bristol, UK	28	23 (82)	24	19 (79)
Katowice, Poland		3 (11)		4 (17)
Glasgow, UK		1 (4)		1 (4)
Bergen, Norway		1 (4)		0 (0)
Sociodemographics at birth				
Age at randomisation (days)	28	18.68 (5.00)	24	19.17 (4.53)
Gender: male[a]	28	22 (79)	24	15 (63)
Median IMD 2015[b] (IQR)	18	23.50 (30.00)	18	25.50 (14.00)
Clinical characteristics at birth				
Birthweight (g)[a]	28	1101.89 (335.54)	24	1322.46 (534.68)
Gestation (weeks)	28	27.64 (2.56)	24	28.50 (3.05)
Grade of IVH: 4	28	14 (50)	24	11 (46)
Maternal age at birth (years)	14	28.50 (6.99)	12	28.17 (6.32)

IMD, Index of Multiple Deprivation.
a Difference of 10%/0.5 SDs or higher between the groups.
b English IMD 2015 scores, UK Data Service Census Support, http://geoconvert.mimas.ac.uk/ (accessed 5 September 2016). Higher scores indicate higher levels of deprivation. IMD based on the children's home postcode at birth for those residing in England only.

In order to determine whether or not those lost to follow-up/died differed from those used in the final analysis, baseline characteristics were compared (*Table 5*). Comparing baseline characteristics between those in our sample with those who have died and those who have either declined or have been lost to follow-up shows us the representativeness of our sample.

Overall, a greater proportion of infants from Poland were lost to follow-up than in the other centres, largely because we are unable to trace patient records in Poland. This is because in Poland, in contrast to the UK, there is no system of single personal numbers that allows patients to be traced. Overall, of those who were lost to follow-up, 57% required a shunt while only 37% of those in our sample had a shunt. There were fewer reservoirs among those lost to follow-up than our sample. Those who did not survive were characteristically more vulnerable and, on average, had lower birthweights and shorter gestation periods and were more likely to have a grade 4 IVH. However, surprisingly, the deprivation index was lower for those who died, suggesting that they were less deprived than those who survived. Given the small sample sizes for IMD, this is most likely a chance finding ($p = 0.048$).

TABLE 4 Characteristics at 2 and 10 years, by trial arm

	Trial arm			
	DRIFT		Standard treatment	
Characteristic	*N*	Mean (SD) or *n* (%)	*N*	Mean (SD) or *n* (%)
Measures at 2 years				
Experienced second IVH[a]	28	8 (29)	24	3 (13)
Shunt	28	11 (39)	24	8 (33)
Reservoir[a]	28	13 (46)	24	19 (79)
Infection	28	0 (0)	24	1 (4)
Measures at 10 years				
Age at 10-year assessment (years)	28	10.56 (1.07)	24	10.76 (1.06)
Weight (kg)	28	35.41 (10.05)	23	34.73 (10.51)
Height (cm)	28	139.09 (12.22)	23	142.26 (11.34)
Head circumference (cm)	28	52.88 (2.53)	23	52.00 (3.43)
MRI performed at 10 years	28	15 (54)	24	12 (50)
Median IMD at 10 years[b] (IQR)	18	23.50 (30.00)	16	25.50 (24.00)
Maternal education[a]				
Left school at age 16 years	28	10 (36)	23	11 (48)
Further education		6 (21)		5 (22)
University degree		12 (43)		7 (30)

IMD, Index of Multiple Deprivation.
a Difference of 10%/0.5 SDs or higher between the groups.
b English IMD 2015 scores, UK Data Service Census Support, http://geoconvert.mimas.ac.uk/ (accessed 5 September 2016). Higher scores indicate higher levels of deprivation. IMD based on the children's home postcode at birth for those residing in England only.

Numbers analysed

Contamination was not a problem in this trial as the intervention was given shortly after birth and could not be requested by the control arm. When DRIFT was followed by persistent enlargement of ventricles and excessive head growth (2 mm/day), management continued with LPs and ventricular reservoir.[31]

Among the original recruits (77 babies), there were three deaths in the DRIFT arm and five in the standard treatment arm by 2 years. We are unable to determine the survival status of two children at 10 years. Deaths and losses to follow-up were all relatively balanced between the group (chi-squared test: $p \geq 0.261$) (*Table 6*); therefore, the further analysis of infants lost to follow-up was not performed. The two children for whom we could not establish survival status were explored in a sensitivity analysis by including them in a best- and a worst-case scenario.

The numbers analysed for each outcome were also relatively balanced between the groups, especially for the cognitive outcomes (chi-squared test, where the lowest *p*-value seen was 0.197) (*Table 7*).

TABLE 5 Baseline characteristics at 10 years: followed up, died or lost to follow-up

Characteristic	Sample at 10 years		Deaths		Uncontactable/declined	
	N[a]	Mean (SD) or *n* (%)	*N*[a]	Mean (SD) or *n* (%)	*N*[a]	Mean (SD) or *n* (%)
Total number of participants	52		12		13	
Centre[a,b]						
Bristol, UK	52	42 (81)	12	6 (50)	13	6 (46)
Katowice, Poland		7 (13)		6 (40)		7 (54)
Glasgow, UK		2 (4)		0 (0)		0 (0)
Bergen, Norway		1 (2)		0 (0)		0 (0)
Sociodemographics at birth						
Age at randomisation (days)	52	18.90 (4.75)	12	18.25 (5.75)	13	19.08 (4.57)
Gender: male	52	37 (71)	12	8 (67)	13	8 (62)
Median IMD 2015[a,c] (IQR)	36	25.00 (25.50)	5	9.00 (9.00)	4	21.5 (15.5)
Clinical characteristics at birth						
Birthweight (g)[a]	52	1203.69 (448.17)	12	961.92 (151.53)	13	1266.92 (397.32)
Gestation (weeks)[a]	52	28.04 (2.80)	12	26.67 (2.35)	13	28.77 (2.71)
Experienced second IVH	52	11 (21)	12	2 (17)	13	3 (23)
Shunt[b]	52	19 (37)	12	5 (42)	13	7 (54)
Reservoir[b]	52	32 (62)	12	8 (67)	13	6 (46)
Infection	52	1 (2)	12	0 (0)	13	0 (0)
Grade of IVH: 4[a]	52	25 (48)	12	8 (67)	13	6 (46)

IMD, Index of Multiple Deprivation.
a Difference of 10%/0.5 SDs or higher between the 10-year sample and those children who died.
b Difference of 10%/0.5 SDs or higher between the 10-year sample and those lost to follow-up.
c English IMD 2015 scores, UK Data Service Census Support, http://geoconvert.mimas.ac.uk/ (accessed 5 September 2016). Higher scores indicate higher levels of deprivation. IMD based on the children's home postcode at birth for those residing in England only.

TABLE 6 Losses to follow-up, by trial arm

Losses	Trial arm, *n*/*N* (%)		*p*-value[a]
	DRIFT	Standard treatment	
Loss to follow-up			
Deaths at 2 years of age	3/39 (8)	5/38 (13)	0.432
Complete loss to follow-up[b]	2/36 (6)	0/33 (0)	–
Of those with known survival status			
Deaths (post 2 years of age) as a result of disability	2/34 (6)	2/33 (6)	0.975
Deaths (post 2 years of age) not as a result of disability	0/34 (0)	0/33 (0)	–
Declined participation	1/34 (3)	1/33 (3)	0.982
Non-responders	3/34 (9)	6/33 (18)	0.261
Attended 10-year follow-up	28/34 (82)	24/33 (73)	0.345

a Chi-squared test.
b These children were untraceable; therefore, we are unsure of their survival status.

TABLE 7 Denominators for assessment, by trial arm

Assessment	Trial arm, *n* (%)		
	DRIFT	Standard treatment	*p*-value[a]
Completion			
BAS III school age score	21 (78)	13 (54)	
Full completion	21 (100)	12 (92)	0.197
Items missing (score created)	0 (0)	1 (8)	
BAS III early year scores	4 (15)	5 (21)	
Full completion	3 (75)	5 (100)	0.236
Items missing (score created)	1 (25)	0 (0)	
BSID III scores	2 (7)	6 (25)	
Full completion	2 (100)	3 (50)	0.206
Items missing	0 (0)	3 (50)	
Visual assessment (parent)	27 (96)	24 (100)	0.350
Visual assessment (CVI)	28 (100)	21 (88)	0.054
Full completion	22 (79)	14 (67)	0.350
Items missing (score created)	6 (21)	7 (33)	
Movement ABC	17 (61)	13 (54)	0.634
CP status	28 (100)	24 (100)	–
SDQ	28 (100)	22 (92)	0.119
Full completion	26 (93)	22 (100)	0.201
Items missing (score created)	2 (7)	0 (0)	

a Chi-squared test.

Outcomes

Primary outcomes

The primary hypothesis was that DRIFT would reduce severe cognitive disability in children assessed at school age. The histogram in *Figure 3* shows the distribution of cognitive quotient (CQ) scores (range 2.07–130.60 points). The graph shows a relatively normal distribution, albeit slightly bimodal. The box plot in *Figure 4* shows the distribution of CQ scores for each group, by trial arm. The results show that those in the DRIFT arm had a median CQ score of 72.3 points, whereas those in the standard treatment arm had a median CQ score of 46.7 points. The maximum CQ score was 130.6 points (achieved in the DRIFT arm), which means that one child had a cognitive ability age 30% higher than his actual age. The highest score achieved in the standard treatment arm was 107.2 points. There were two quotients of < 30 points in the DRIFT arm, compared with seven in the standard treatment arm.

The histogram in *Figure 5* shows the distribution of CQ scores (giving those who died a score of 0 points) (range 0.00–130.60 points). The graph shows a relatively normal distribution, slightly skewed by the scores of 0 points. The box plot in *Figure 6* shows the scores by trial arm (giving those who died a score of 0 points). The results show that those in the DRIFT arm had a median CQ score of 72.0 points whereas those in the standard treatment arm had a median CQ score of 44.6 points.

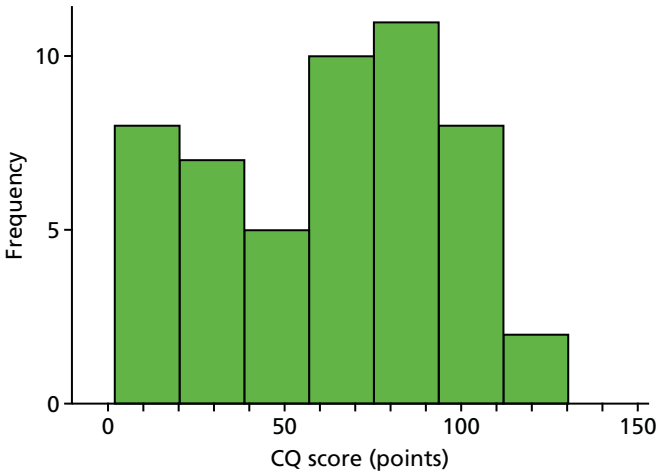


FIGURE 3 Histogram of CQ scores, excluding deaths.

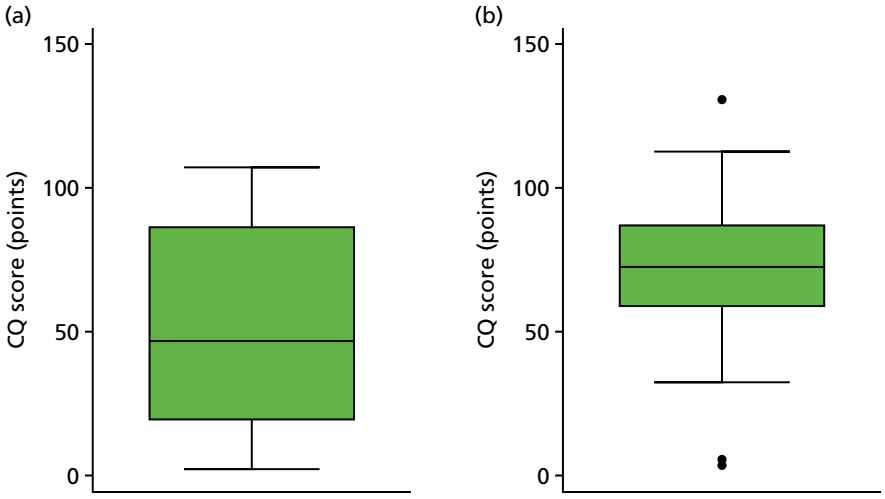


FIGURE 4 Cognitive quotient scores, by trial arm. (a) Standard treatment; and (b) DRIFT.

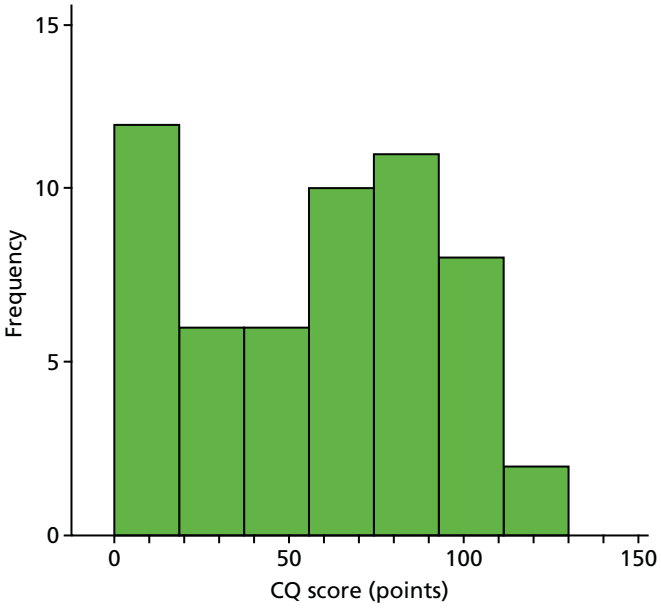


FIGURE 5 Sensitivity analysis: histogram of CQ scores, including deaths as a score of 0 points.

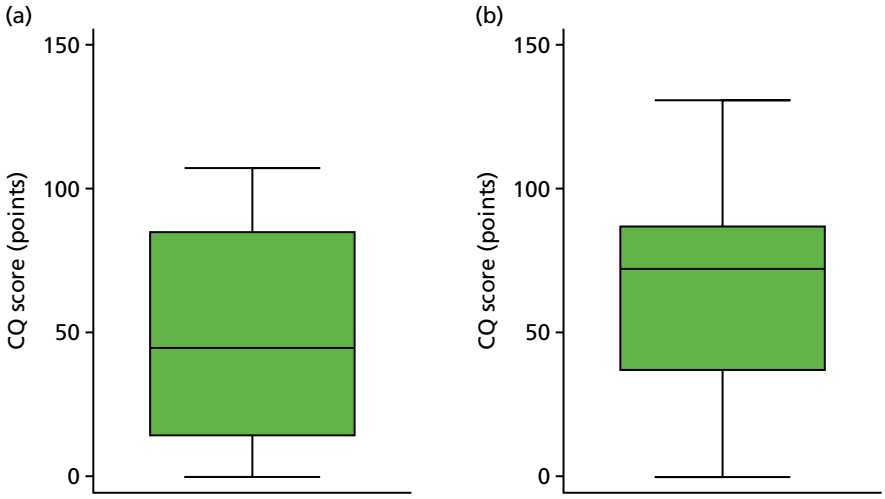


FIGURE 6 Sensitivity analysis: CQ scores, by trial arm. (a) Standard treatment; and (b) DRIFT.

Table 8 shows the results for the primary analysis, both including and excluding deaths. Given the larger than expected attrition/death rate, precision was lower than hoped and was exacerbated further by large SDs for the cognitive ability quotient. Despite this, results are in parallel with those at 2 years, with crude estimates giving very weak evidence that the DRIFT intervention increases cognitive ability at 10 years ($p = 0.096$). After adjusting for gender, birthweight and grade of IVH, this evidence was strengthened and indicated that children who were in the DRIFT arm of the trial had, on average, a CQ score of 23.47 points higher than those who received standard treatment ($p = 0.009$). This translates into a developmental cognitive advantage of 2.5 years.

Given the look of the histogram (see *Figure 3*), we felt that it was important to explore the regression assumptions to ensure that we had used the right model for our data. Looking at the mean and median of our overall data, it was clear that they were similar: median 68.71 (IQR 54.28), mean 61.96 (SD 33.44). The skewness and kurtosis were –0.21 and 2.08, respectively; therefore, we were satisfied that the distribution was fairly symmetrical but slightly platykurtic (flat). The relationship between CQ score and birthweight was fairly linear (linear regression *p*-value = 0.015; *Figure 7*). After running the adjusted model, the residuals are approximately normally distributed (*Figures 8* and *9*). There is also no evidence to suggest that there is an increasing variance over the values of the linear predictor (*Figure 10*). Therefore, the team felt confident that a linear regression model was appropriate.

TABLE 8 Primary outcome: CQ score, by trial arm[a]

Outcome	Trial arm, mean (SD)		Difference in means[a] (95% CI)	*p*-value[a]	Adjusted difference in means[b] (95% CI)	*p*-value[b]
	DRIFT	Standard treatment				
CQ score (points)	69.33 (30.06)	53.68 (35.70)	15.65 (–2.86 to 34.16)	0.096	23.47 (6.23 to 40.71)	0.009
CQ score (points)[c]	64.55 (34.04)	49.55 (37.22)	15.00 (–4.28 to 34.27)	0.125	22.33 (4.77 to 39.89)	0.014

a Linear regression, crude estimates.
b Linear regression, adjusted for gender, birthweight and grade of IVH (prespecified adjustment).
c Giving children who have died post 2 years of age a score of 0 points (sensitivity analysis).

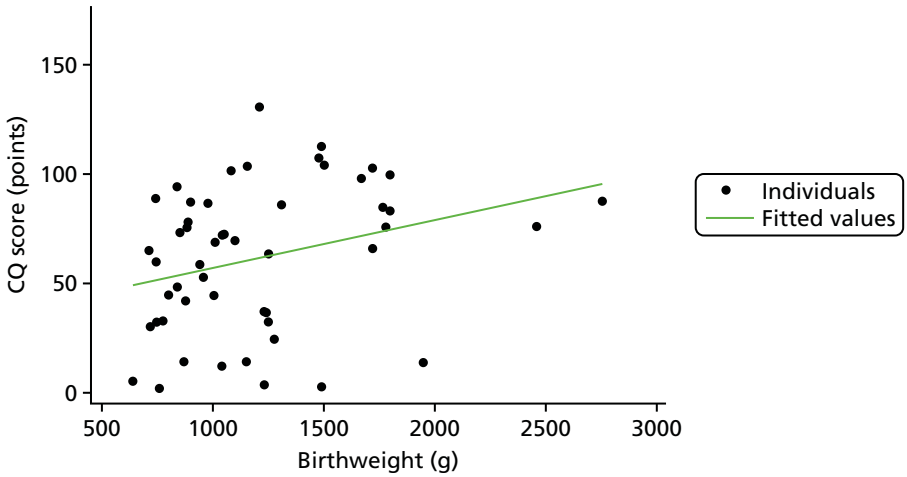


FIGURE 7 Regression diagnostic: relationship between CQ score and birthweight.

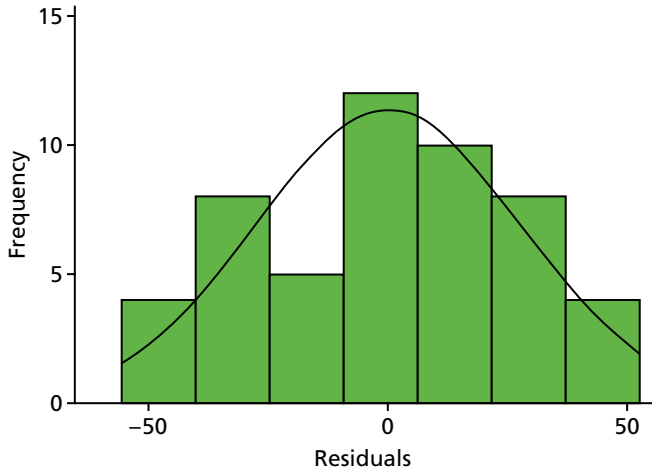


FIGURE 8 Regression diagnostic: plotted histogram of residuals from the primary regression model.

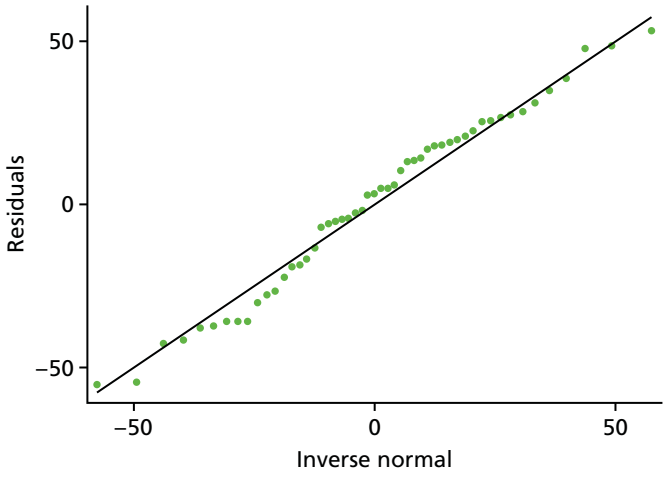


FIGURE 9 Regression diagnostics: residuals vs. normal distribution, estimated from the primary analysis model.

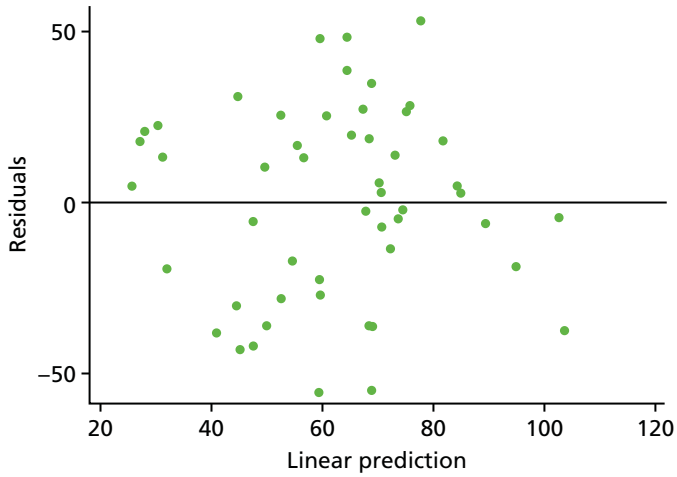


FIGURE 10 Regression diagnostics: residuals vs. predicted values, estimated from the primary analysis model.

Assessing prespecified covariants

Covariates were prespecified in the SAP and included birthweight, IVH grade and gender. These were the same covariates used in the 2-year follow-up study in which these variables had previously been shown to be imbalanced at 6-month follow-up. *Table 9* shows how each of the covariates were individually related to the cognitive ability quotient. The results show that, for each additional gram of birthweight, CQ score at 10 years increased by 0.02 points (or 20 points per 1 kg). Those with grade 3 IVH had CQ scores that were, on average, 24.37 points higher than those with grade 4 IVH. Girls had CQ scores that were, on average, 13.96 points higher than those of boys. *Figures 11* and *12* illustrate the spread of CQ scores across gender and IVH grade.

Figure 13 shows the relationship between cognitive outcome and birthweight (g). Although the fitted line appears to be curved, the outlying values of birthweight may be suggesting more curvature than there actually is. A simple likelihood ratio test comparing the model with and without a quadratic term gives a *p*-value of 0.346, suggesting that the null hypothesis of a linear relationship is not rejected.

It is also important to establish which of these three covariates are strengthening the relationship between arm and CQ score. *Table 10* shows the regression coefficients after adjustment for each covariate on its own.

All three covariates strengthen the relationship between cognitive score and trial arm. Birthweight has proved to be the strongest adjustment, offering a strong difference between the groups with and without deaths included as zero. Gender and IVH grade both strengthen the difference between the groups, but to a smaller degree than birthweight.

TABLE 9 Cognitive quotient score, adjusting for each covariate independently, without trial arm

Covariate	Difference in mean cognitive ability[a] (95% CI)	*p*-value[a]
Cognitive ability at 10 years		
Birthweight	0.02 (0.00 to 0.04)	0.035
IVH[b]	–24.37 (–42.09 to –6.66)	0.008
Gender[c]	–13.96 (–34.88 to 6.96)	0.186

a Linear regression.
b Coded 0 for grade 3 IVH and 1 for grade 4 IVH.
c Coded 0 for females and 1 for males.

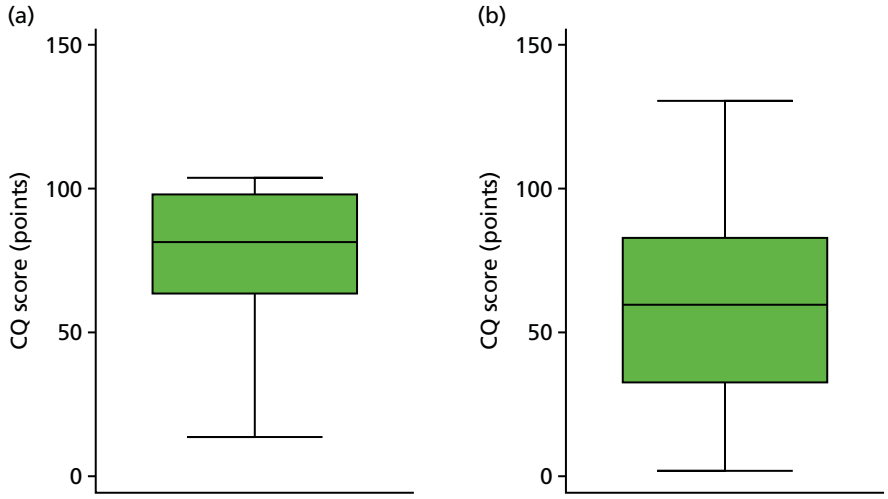


FIGURE 11 Relationship between cognitive outcome and gender. (a) Female; and (b) male.

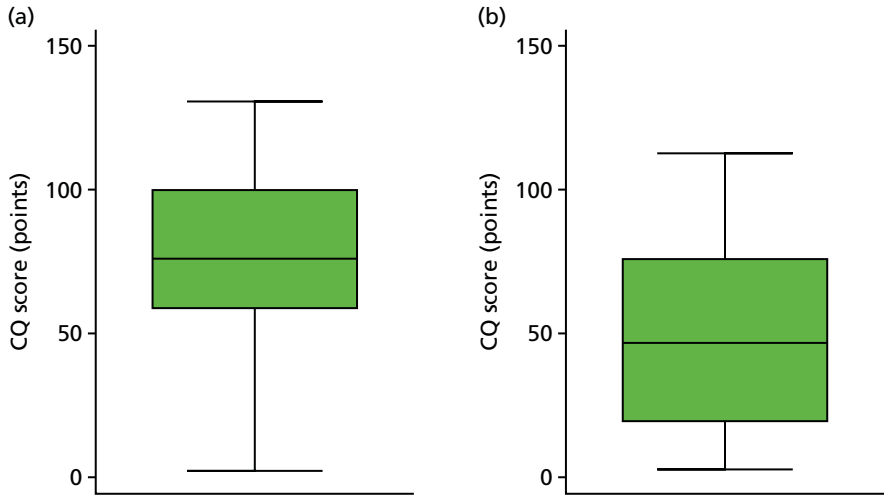


FIGURE 12 Relationship between cognitive outcome and grade of IVH. (a) Grade 3 IVH; and (b) grade 4 IVH.

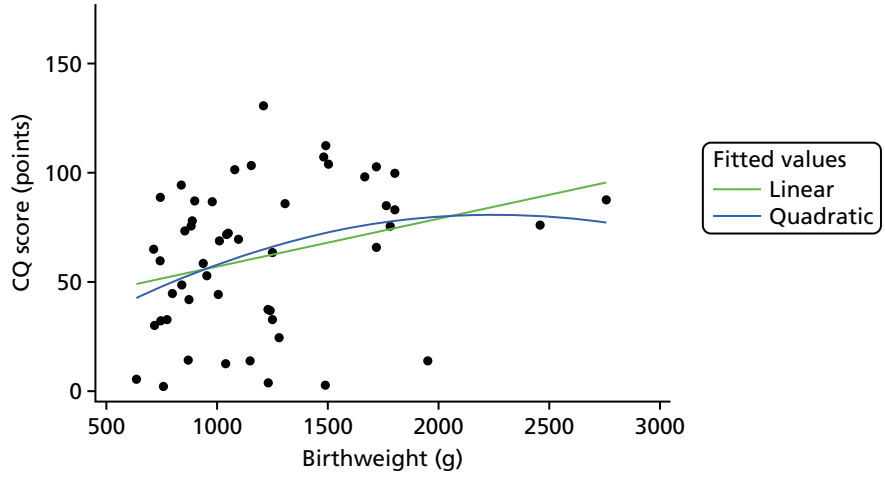


FIGURE 13 Relationship between CQ score and birthweight with a linear and quadratic fit line.

TABLE 10 Cognitive quotient score by trial arm, adjusting for each covariate independently

CQ adjusted for	Difference in means[a] (95% CI)	*p*-value[a]	Difference in means[b] (95% CI)	*p*-value[b]
Birthweight	21.79 (3.90 to 39.67)	0.018	21.90 (3.68 to 40.12)	0.019
IVH grade	16.22 (–1.06 to 33.50)	0.065	16.56 (–1.29 to 34.41)	0.068
Gender	19.16 (0.63 to 37.70)	0.043	16.04 (–3.39 to 35.57)	0.104

a Linear regression.
b Linear regression, giving children who have died post 2 years of age a score of 0 (sensitivity analysis).

Although gender shows a very weak relationship with cognitive ability at 10 years, and only a small adjustment when added as a covariate, it was imbalanced between the arms using the 10%/0.5 SDs rule. IVH and birthweight are both predictors of cognitive score, so adjustment is appropriate. Overall, these three covariates are appropriate for this analysis when taking into account both their relationship with the outcome and distribution across arms.

Secondary outcomes

Visual

Figure 14 shows the four categories of sight by arm. The two lightest shades (solid outer line) make up a positive visual outcome and the two darkest shades (dashed outer line) make up a negative visual outcome.

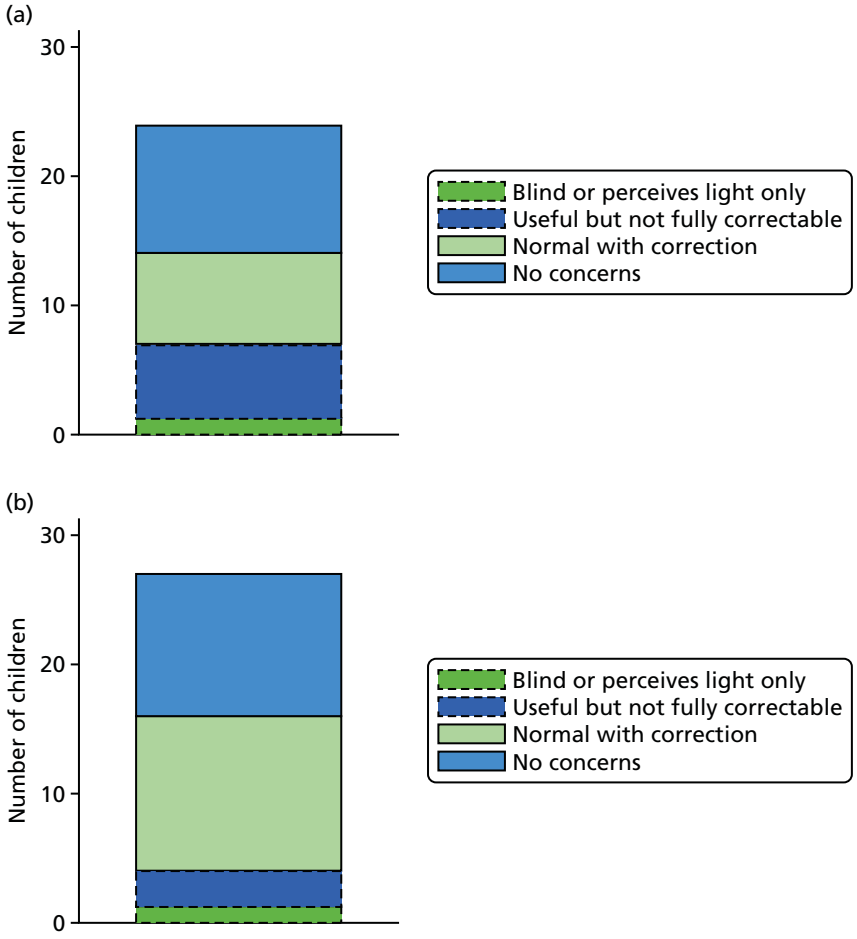


FIGURE 14 Visual outcome, by trial arm. (a) Standard treatment; and (b) DRIFT.

There appears to be a larger proportion of 'good' visual outcomes in the DRIFT arm than in the standard treatment arm. A logistic regression model below shows this difference in greater detail.

As well as this binary outcome, a 23-question visual assessment task (see *Appendix 4*) was also filled out by vision specialists, who asked the parents various questions relating to CVI. Each question was scored 0–5, with higher scores indicating better cerebral vision.

Table 11 shows the results from the visual questions. For the binary visual outcome, this was answered for 27 and 24 children in the DRIFT arm and standard treatment arm, respectively. Overall, the results show that those in the DRIFT arm were almost four times more likely to have a 'good' visual outcome than those in the standard treatment arm (adjusted OR 3.73); however, the *p*-value provides only very weak evidence to support this (*p*-value of 0.136). We realised after analysing the result that one child had been given the CVI questionnaire even though he was blind. As a sensitivity analysis, we re-ran the analysis removing this child as their result was considered inappropriate. The result remained unchanged. The mean score for CVI is very slightly lower in the DRIFT arm; however, this result is consistent with chance (*p*-value of 0.502). The Mann–Whitney *U*-test (a suitable non-parametric comparator to the regression model) gave a very similar result, with even weaker evidence of a difference (*p*-value of 0.618). The team felt that it was safe to conclude that there is little evidence that the intervention had an effect on parent-reported CVI. *Figures 15* and *16* show the distribution of CVI scores including and excluding the blind child, respectively.

Sensorimotor

It was prespecified that any child for whom the Movement ABC classification score was missing and who was diagnosed with CP would automatically be placed in the severe category. The results of which are presented in *Figure 17*.

Overall, the percentage of children with 'severe' sensorimotor scores was higher in the DRIFT group: 83% vs. 74% (see *Table 11*). On closer inspection, it became clear that, although many more children were slipping into this category, within this category, scores were higher in the DRIFT group [DRIFT, mean 38.69 (SD 11.76); standard treatment, mean 29.69 (SD 11.21) for the 'severe' category]. *Figure 18* shows the distribution of the sensorimotor scores, by group.

TABLE 11 Secondary outcome: parent-reported visual assessment results

Outcome	Trial arm, *n* (%)/mean (SD)		Difference[a] (95% CI)	*p*-value[a]	Adjusted difference[b] (95% CI)	*p*-value[b]
	DRIFT	Standard treatment				
Visual function (parent reported)						
Good vision	23 (85%)	17 (71%)	2.37 (0.60 to 9.40)[c]	0.221[c]	3.73 (0.66 to 21.14)[c]	0.136[c]
CVI mean score	4.50 (0.70)	4.65 (0.38)	–0.15 (–0.49 to 0.19)[d]	0.379[d]	–0.12 (–0.47 to 0.24)[d]	0.502[d]
CVI median score	4.76 (0.67)	4.78 (0.48)		0.618[e]		
CVI mean score[f]	4.59 (0.55)	4.65 (0.38)	–0.07 (–0.35 to 0.22)[d]	0.640[d]	–0.04 (–0.33 to 0.26)[d]	0.793[d]

a Crude estimates.
b Adjusted for age, gender, birthweight and grade of IVH (prespecified adjustment).
c Logistic regression.
d Linear regression.
e Wilcoxon (Mann–Whitney) test.
f After removing any CVI scores for 'blind' children as these were considered inappropriate.

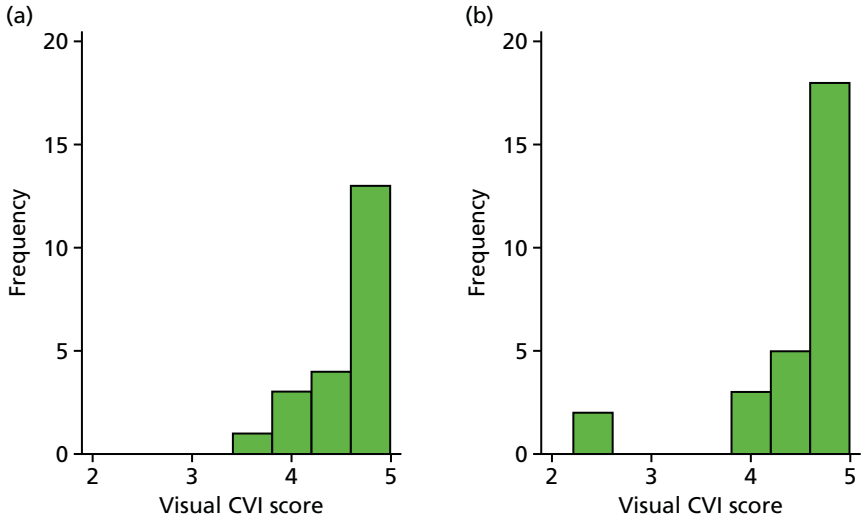


FIGURE 15 Histogram of visual results, by trial arm. (a) Standard treatment; and (b) DRIFT.

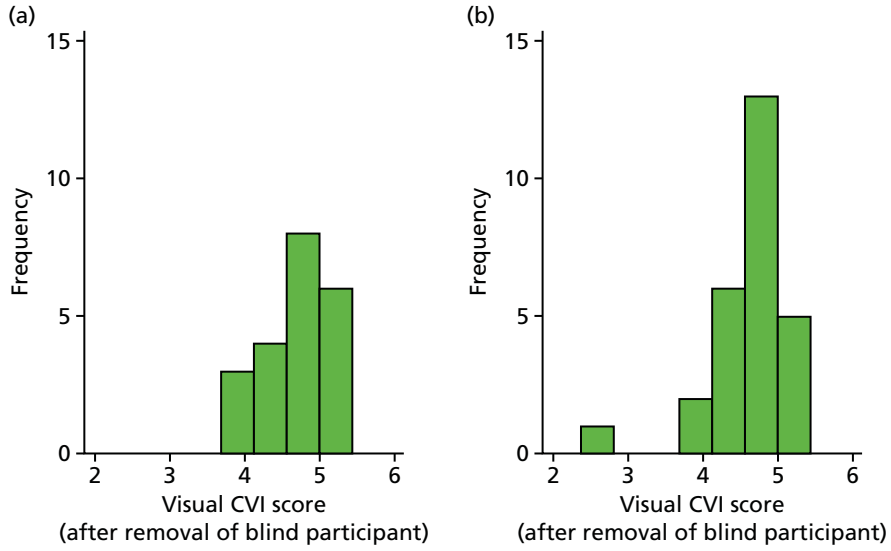


FIGURE 16 Histogram of visual results after removal of blind child, by trial arm. (a) Standard treatment; and (b) DRIFT.

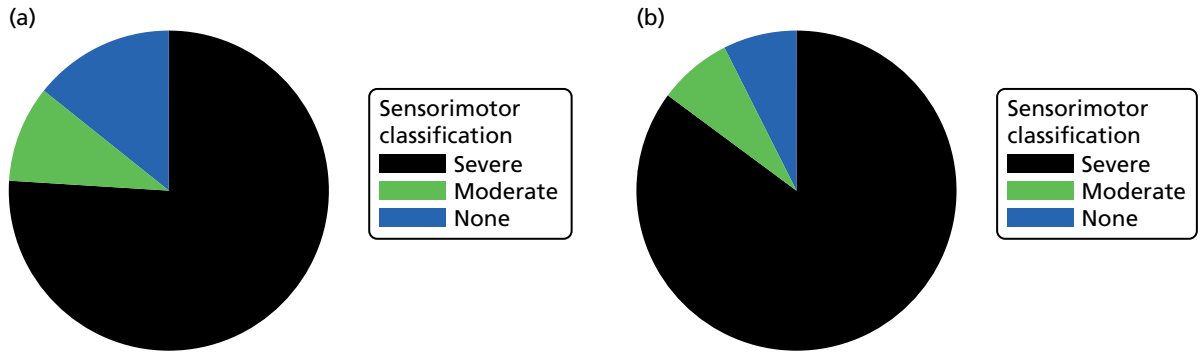


FIGURE 17 Sensorimotor ABC classification, by trial arm. (a) Standard treatment; and (b) DRIFT.

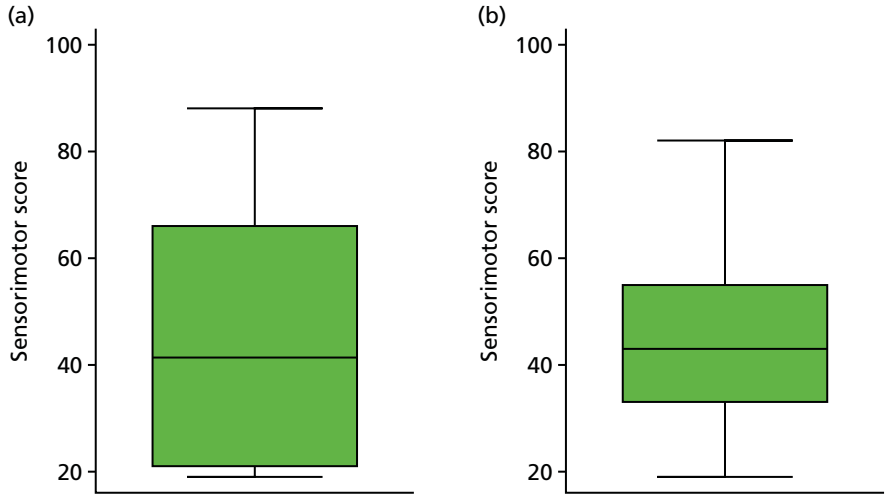


FIGURE 18 Sensorimotor score, by trial arm. (a) Standard treatment; and (b) DRIFT.

Closer inspection of the results showed that the assumption that children diagnosed with CP would score < 55 was appropriate as the average sensorimotor score among children in this category who completed the test was 28.92 (maximum 45.00) and the average sensorimotor score among children without a CP diagnosis was 59.74 (maximum 88.00). Nevertheless, we conducted the same test, using only those who had carried out the test, and achieved a score to see whether or not the assumption made an impact on our findings. Reassuringly, this gave a very similar result.

It was also thought (post hoc) that a dichotomised outcome would also allow us to feel confident with the conclusions drawn; therefore, this was carried out in the same way as the original (prespecified analysis) but dichotomising on a score of < 55 or ≥ 55 (severe vs. moderate/mild disability). All of these analyses are presented in *Table 12*.

Reassuringly, all of the models gave a similar result, with the conclusion that, although small positive effects were seen in the DRIFT arm, after adjustment, these results were consistent with chance. Adjustment did appear to change the conclusion from negative to positive for the DRIFT intervention. Looking at each of the covariates individually, as with the primary outcome, birthweight caused the largest shift in treatment effect. When using the continuous measure, we did achieve weak evidence to suggest that those in the DRIFT intervention had better sensorimotor scores than the standard treatment arm; however, this was not prespecified as an outcome and has a very low sample size. Therefore, there were no strong differences seen between the arms when looking at motor ability.

The number of children diagnosed with CP was also prespecified as a secondary outcome. Children in the DRIFT arm were 1.1 times more likely than those in the standard treatment arm to have CP (*Table 13*). After adjustment for gender, birthweight and grade of IVH, this changed to a 63% lower odds of CP in the DRIFT group; we know this is largely because those in the DRIFT group had less favourable baseline characteristics. Looking at each of the covariates individually, as with the primary outcome, birthweight caused the largest shift in treatment effect. Although the DRIFT arm included a higher percentage of children with CP than the standard treatment arm (61% vs. 58%, respectively), it appeared that children in the DRIFT arm were less likely to have CP categorised as severe. After adjustment, those in the DRIFT arm were 80% more likely to be ambulant than those in the standard treatment arm. However, given the large CI and *p*-value, there was not strong enough evidence to support this and it could have simply happened by chance. As with the Movement ABC scoring, the results provided no substantial evidence to suggest a difference between the groups.

TABLE 12 Secondary outcome: severity of sensorimotor disability

Outcome	*N* (D : S)	DRIFT, *n* (%)/ mean (SD)	Standard treatment, *n* (%)/mean (SD)	Difference[a] (95% CI)	*p*-value[a]	Adjusted difference[b] (95% CI)	*p*-value[b]
Sensorimotor disability[c,d]							
None/green (3)		2 (7%)	3 (14%)				
Moderate/amber (2)	27 : 21	2 (7%)	2 (10%)	0.55 (0.13 to 2.34)[e]	0.416[e]	3.66 (0.33 to 40.34)[e]	0.290[e]
Severe/red (1)		23 (85%)	16 (76%)				
Sensorimotor disability[c,f]							
None/green (3)		2 (12%)	3 (23%)				
Moderate/amber (2)	17 : 13	2 (12%)	2 (15%)	0.48 (0.10 to 2.29)[e]	0.359[e]	2.45 (0.23 to 26.66)[e]	0.461[e]
Severe/red (1)		13 (76%)	8 (62%)				
Sensorimotor disability[c,g]							
Severe/red vs. rest	27 : 21	23 (85%)	16 (76%)	1.80 (0.42 to 7.75)[h]	0.432	0.19 (0.012 to 3.29)[h]	0.257
Continuous score[i]	17 : 13	45.94 (17.40)	46.96 (24.87)	–1.02 (–16.82 to 14.78)[j]	0.896	11.29 (–1.87 to 24.46)[j]	0.089

a Crude estimates.
b Adjusted for age, gender, birthweight and grade of IVH.
c Movement ABC-2 'Traffic Light' System: green zone is described as 'no movement difficulty detected'; amber zone as 'suggests child is "at risk" of having a movement difficulty, monitoring required' and red zone as 'denotes a significant movement difficulty'.
d A priori: ordinal logistic regression – if a child had a missing score for the Movement ABC classification and was diagnosed with CP, they were automatically put into the severe category (prespecified outcome).
e Ordinal logistic regression.
f Post hoc check 1: ordinal logistic regression, using only those with a motor component score.
g Post hoc check 2: logistic regression, using a dichotomised outcome by combining moderate/mild motor disability.
h Logistic regression.
i Post hoc check 3: linear regression, using the continuous sensorimotor score.
j Linear regression.

Emotional/behavioural difficulties

To assess emotional and behavioural difficulties, parents were asked to fill in the SDQ (see *Appendix 6*). The results are shown in *Table 14*. The subscales were almost all skewed to the left, indicating more 'normal' behaviour; therefore, subscales were assessed using the Mann–Whitney non-parametric *U*-test. However, the total score did approximately follow a normal distribution and, therefore, was assessed using linear regression (as prespecified in the analysis plan).

Higher values of the SDQ total score indicate more 'abnormal' behaviour and there was no difference between the two groups (adjusted mean difference 2.01, 95% CI –2.78 to 6.81; $p = 0.401$). Although the 'Conduct Problems' subscale showed more favourable results in the standard treatment arm ($p = 0.033$), this is, given the large number of tests carried out here for a single secondary outcome, most likely a 'chance finding'. *Figure 19* shows how the total score varied between groups.

Magnetic resonance imaging findings

There were no major differences relating to residual neurosurgical conditions needing referral; results are presented by arm in *Table 15*.

Residual catheter tracks were more often seen in the standard treatment group and in association with ventricular reservoirs.

TABLE 13 Secondary outcome: CP diagnosis at 10 years

Outcome	DRIFT, *n* (%)	Standard treatment, *n* (%)	Difference[a] (95% CI)	*p*-value[a]	Difference[b] (95% CI)	*p*-value[b]
CP						
Without CP	11 (39)	10 (42)				
With CP	17 (61)	14 (58)	1.10 (0.36 to 3.35)	0.862	0.37 (0.07 to 2.00)	0.249
CP level						
1	7 (41)	5 (36)				
2	4 (24)	3 (21)				
3	2 (12)	0 (0)				
4	0 (0)	2 (14)				
5	4 (24)	4 (29)				
Ambulatory status						
Ambulant (level 1–2)[c]	11 (65)	8 (57)	1.38 (0.32 to 5.88)	0.667	1.32 (0.24 to 7.25)	0.751
Non-ambulant (level 3–5)[c]	6 (35)	6 (43)				

a Crude estimates.
b Adjusted for age, gender, birthweight and grade of IVH.
c For children with CP, they were categorised as ambulant (GMFCS level 1–2) or non-ambulant (GMFCS level 3–5).

TABLE 14 Secondary outcome: SDQ results

Outcome	*N* (D : S)	DRIFT, mean (SD)[a]	Standard treatment, mean (SD)[a]	Difference (95% CI)[b]	*p*-value[b]	Difference (95% CI)[c]	*p*-value[c]
Emotional/behavioural difficulties (as predefined means)							
Emotional symptoms[d]	28 : 22	3.32 (2.88)	2.59 (2.11)		0.502		
Conduct problems[e]	28 : 22	2.68 (1.93)	1.55 (1.44)		0.033		
Hyperactivity/ inattention[f]	28 : 22	5.54 (3.18)	6.14 (2.85)		0.555		
Peer relationships[g]	28 : 22	3.36 (2.63)	3.09 (2.29)		0.760		
Pro-social behaviour[h]	28 : 22	7.11 (2.63)	6.95 (2.28)		0.567		
Impact score[i]	28 : 22	2.46 (2.53)	2.23 (2.94)		0.530		
SDQ total score[j]	28 : 22	14.89 (8.48)	13.36 (6.59)	1.53 (–2.89 to 5.94)	0.490	2.01 (–2.78 to 6.81)	0.401

a The subscores had a skewed distribution; however, means (SDs) were used to make the direction clear as medians (IQR) were often the same for both groups.
b Crude estimates, Mann–Whitney *U*-test for the skewed subscores and linear regression for the overall score.
c Adjusted for age, gender, birthweight and grade of IVH (for parametric tests only).
d On a scale of 0–10, where 0–3 corresponds to normal behaviour, 4 is borderline and 5–10 is abnormal.
e On a scale of 0–10, where 0–2 corresponds to normal behaviour, 3 is borderline and 4–10 is abnormal.
f On a scale of 0–10, where 0–5 corresponds to normal behaviour, 6 is borderline and 7–10 is abnormal.
g On a scale of 0–10, where 0–2 corresponds to normal behaviour, 3 is borderline and 4–10 is abnormal.
h On a scale of 0–10, where 6–10 corresponds to normal behaviour, 5 is borderline and 0–4 is abnormal.
i On a scale of 0–10, where 0 is normal, 1 is borderline and 2–10 is abnormal.
j A summation of the emotional symptoms, conduct problems, hyperactivity and peer relationships. On a scale of 0–40, where 0–13 corresponds to normal behaviour, 14–16 is borderline and 17–40 is abnormal.

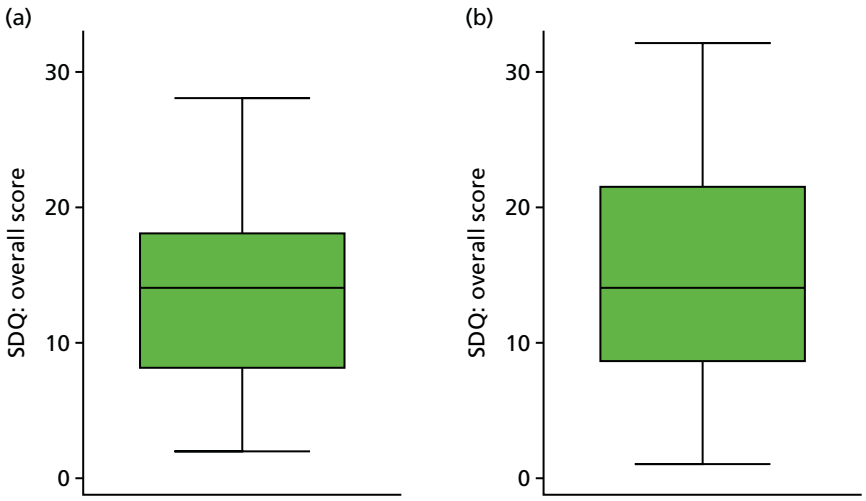


FIGURE 19 Strengths and Difficulties Questionnaire score, by trial arm. (a) Standard treatment; and (b) DRIFT.

TABLE 15 Magnetic resonance imaging findings

Scan findings	Trial arm, *n* (%)	
	DRIFT (*N* = 16)	Standard treatment (*N* = 12)
Residual catheter tract	3 (19)	4 (33)
Parenchymal lesion	7 (44)	5 (42)
Reservoir	9 (56)	9 (75)
VP shunt	7 (44)	4 (33)
Possible active hydrocephalus	2 (13)	1 (8)
Residual condition	2 (13)	2 (17)

Sensitivity analyses

Various different techniques were used to address the primary analysis; these are of an exploratory nature. Reassuringly, all of the analyses gave compatible results (presented in *Table 16*). As well as adjustments to the way the outcome was measured, we also looked into adjustments for factors that may influence overall effect. First, we adjusted for centre (post hoc) and established that this made no difference to the conclusion. Compared with the original crude model, adjusting for centre weakened the average difference from 15.65 to 14.55 CQ points. The proportion of Polish children was higher in the standard treatment arm than in the DRIFT arm and, consequently, the proportion of children from Bristol was higher in the DRIFT arm. On average, CQ scores were higher in Bristol children than in Polish children (mean difference 14.40, 95% CI 14.08 to 42.88), which may explain the weakened effect after adjustment for centre. The binary outcome gave very similar results to the continuous CQ outcome. Both the unadjusted and adjusted models provided strong evidence to suggest that DRIFT had a positive impact on children's cognitive outcomes at 10 years. Using the figure estimated, we calculated a number needed to treat (NNT) of three using the following calculation: 1/(14/26 – 8/29). Including all deaths (pre and post 2 years of age) as a negative outcome gave the same result. The ordinal outcome, unsurprisingly, offered a similar result to our primary analyses. In total, 50% of children in the standard treatment arm had severe cognitive disability (> 3 SDs below the population mean), compared with 21% of children in the DRIFT arm. Those in the DRIFT arm were at 3.63 times more likely to be in a higher category (better outcome) than those in the standard treatment arm, after adjustment for covariates. This method allowed us to differentiate between deaths and grades of disability by increasing the number of categories.

Adjustment for maternal education was a decision made after data analysis had begun. Unfortunately, maternal age and education were not collected at baseline; however, maternal education (left school at

TABLE 16 Sensitivity analysis for the primary outcome

Sensitivity analysis	*N* (D : S)	Trial arm, *n* (%)/mean (SD)		Difference (95% CI)[a]	*p*-value[a]	Adjusted difference (95% CI)[b]	*p*-value[b]
		DRIFT	Standard treatment				
Original primary analysis							
CQ	27 : 24	69.33 (30.06)	53.68 (35.70)	15.65 (–2.86 to 34.16)[d]	0.096	23.47 (6.23 to 40.71)[d]	0.009
CQ[c]	29 : 26	64.55 (34.04)	49.55 (37.22)	15.00 (–4.28 to 34.27)[d]	0.125	22.33 (4.77 to 39.89)[d]	0.014
Continuous measure of cognitive ability							
CQ (Bristol cohort only)	23 : 19	71.76 (27.42)	57.83 (34.78)	13.93 (–5.46 to 33.33)[d]	0.154	24.88 (6.82 to 42.94)[d]	0.008
CQ (Bristol cohort only)[c]	24 : 20	68.77 (30.56)	54.94 (36.24)	13.84 (–6.48 to 34.15)[d]	0.177	23.27 (4.65 to 41.88)[d]	0.016
Binary measure[e]							
Alive and without severe cognitive disability (post 2 years)	29 : 26	21 (72%)	11 (42%)	3.58 (1.16 to 11.04)[f]	0.026	9.96 (2.12 to 46.67)[f]	0.004
Alive and without severe cognitive disability (including all 12 deaths)	32 : 31	21 (66%)	11 (35%)	3.47 (1.23 to 9.78)[f]	0.019	7.69 (1.96 to 30.11)[f]	0.003
Cognitive disability category							
1. Dead	29 : 26	2 (7%)	2 (8%)				
2. Severe		6 (21%)	13 (50%)				
3. Moderate		7 (24%)	2 (8%)	2.04 (0.77 to 5.42)[g]	0.151	3.63 (1.21 to 10.90)[g]	0.022
4. Mild		8 (28%)	4 (15%)				
5. No cognitive disability		6 (21%)	5 (19%)				

Sensitivity analysis	*N* (D : S)	Trial arm, *n* (%)/mean (SD)		Difference (95% CI)[a]	*p*-value[a]	Adjusted difference (95% CI)[b]	*p*-value[b]
		DRIFT	Standard treatment				
Additional adjustments for the original primary analysis (CQ score)							
Adjusted for centre	27 : 24	69.33 (30.06)	53.68 (35.70)	13.76 (–4.45 to 31.92)[d]	0.135	22.00 (5.69 to 38.30)[d,h]	0.009
Adjusted for centre (Bristol vs. others)	27 : 24	69.33 (30.06)	53.68 (35.70)	14.55 (–3.78 to 32.87)[d]	0.117	23.19 (6.35 to 40.04)[d,h]	0.008
Adjusted for maternal education[i]	27 : 23	69.33 (30.06)	55.90 (34.77)	11.50 (–6.86 to 29.87)[d]	0.214	20.08 (2.96 to 37.21)[d,h]	0.023
Adjusted for baseline imbalance[j]	27 : 24	69.33 (30.06)	53.68 (35.70)	24.58 (6.69 to 42.46)[d]	0.008		

D, DRIFT; S, standard treatment.
a Crude estimates.
b Adjusted for gender, birthweight and grade of IVH, unless otherwise stated.
c Giving children who have died post 2 years a score of 0.
d Linear regression.
e Alive and well vs. severely disabled/died owing to disability.
f Logistic regression.
g Ordinal logistic regression.
h Additional adjustment for gender, birthweight and IVH.
i Using maternal education measured at 10 years (missing for one standard treatment patient), making the assumption that maternal education collected at 10 years has not changed since baseline.
j Factors imbalanced at baseline between those analysed at 10 years were gender and birthweight (additional adjustment for IVH was not carried out as it would duplicate the primary analysis).

16 years of age, further education or university degree) was measured at 10 years. The team felt that, although imprecise, this was an adequate estimate of maternal education at baseline. The mean CQ score for infants born to 'university degree' mothers was 75.49 points, for infants born to 'further education' mothers was 50.84 points and for infants bon to mothers who 'left school at 16' was 57.84 points ($p = 0.094$). The proportion of mothers with a university degree was higher in the DRIFT arm than in the control arm (43% vs. 30%). Therefore, adjustment resulted in a weakened effect estimate. However, it should be pointed out that adjustment for birthweight, IVH, gender and maternal education still produced compatible results to the primary analysis ($p = 0.023$).

An adjustment for imbalances at baseline was prespecified in the analysis plan and defined as any difference of ≥ 10%/0.5 SDs between the groups. Referring back to *Table 2*, the variables classed as imbalanced were gender and birthweight. Adjustment for only these two factors resulted in strong evidence that the DRIFT intervention improves cognitive outcome at 10 years.

Best- and worst-case scenarios

Unfortunately, two patients could not be followed up at 10 years and their survival status was, therefore, unknown. Given the number of deaths post 2 years of age, it is unlikely that these patients would have died; however, it is important to understand the significance of these patients by calculating the extremes. Therefore, for a best-case scenario, the two children in the DRIFT arm were presumed to be alive and well (with the median score of their group), and vice versa for the worst-case scenario; the two children in the DRIFT arm were presumed, dead (with a score of 0). Results from these analyses are presented in *Table 17*. These assumptions are very extreme and the results, as expected, show that the best-case scenario strengthens the unadjusted treatment effect, whereas the worst-case scenario weakens it to produce a treatment difference consistent with chance (but still in favour of DRIFT).

Multiple imputation

In order to carry out a multiple imputation model, we must first assess whether or not the data are missing at random (MAR). In *Figures 20* and *21*, the red markers highlight the gestations/birthweights when the CQ score is missing. When checking cognitive scores across gestation and birthweight levels, it appears that missing data are evenly spread across these variables. At first glance, the MAR assumption appears to be valid across these two variables.

TABLE 17 Best- and worst-case scenarios

Sensitivity analysis	*N* (D : S)	Trial arm, mean (SD)		Difference (95% CI)[a]	*p*-value[a]	Adjusted difference (95% CI)[b]	*p*-value[b]
		DRIFT	Standard treatment				
Different scenarios for the two patients with unknown survival status							
Best-case scenario[c]	31 : 26	65.04 (32.94)	49.55 (37.22)	15.49 (–3.13 to 34.12)	0.101	20.67 (3.68 to 37.65)	0.018
Worst-case scenario[d]	31 : 26	60.38 (36.62)	49.55 (37.22)	10.83 (–8.83 to 30.50)	0.274	15.28 (–3.72 to 34.29)	0.113

D, DRIFT; S, standard treatment.
a Crude estimates.
b Adjusted for gender, birthweight and grade of IVH.
c Assuming the two children in the DRIFT arm were all alive and well (with the median score for their group) at 10 years.
d Assuming the two children in the DRIFT arm died.
Giving children who have died post 2 years of age a score of 0.

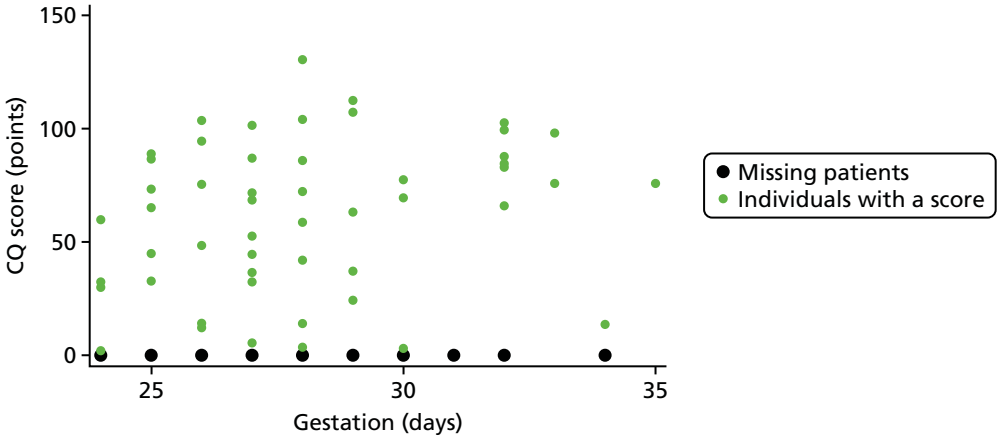


FIGURE 20 Testing the MAR assumption: gestation.

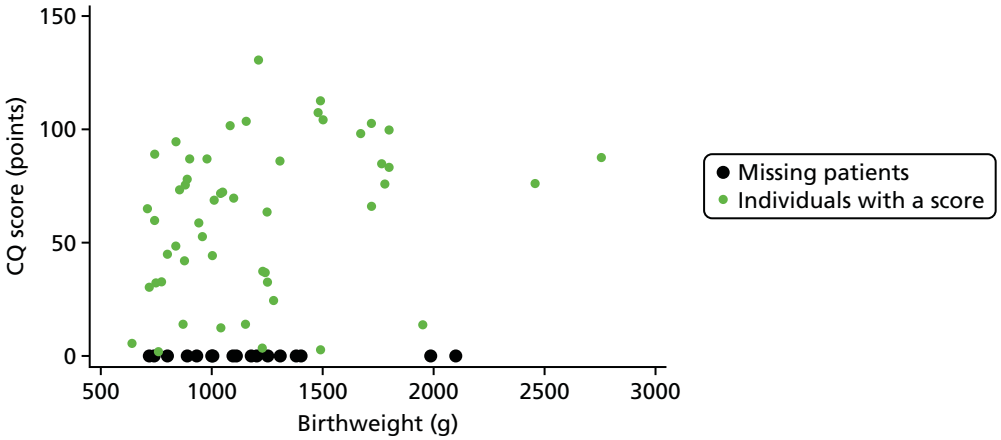


FIGURE 21 Testing the MAR assumption: birthweight.

Logistic regression was used to test which baseline characteristics and follow-up data points were predictive of missing CQ at 10 years. There were several variables that were predictive of missing CQ [centre, receiving a shunt at 2 years, mental development quotient (DQ) at 2 years and disability level at 2 years]. There were also several variables that were useful predictors of CQ [trial arm, age at entry (days), birthweight, gestation, IVH grade, the following measures at 2 years: mental DQ, motor DQ, gait, sitting, hand, speech, vision, disability, and the following measures at 10 years: vision, seizures, shunts, cerebral palsy, sensorimotor, hyperactivity, peer relationships and prosocial].

To examine the relationship between mental and motor DQs at 2 years and CQ at 10 years, we created a scatterplot. The scatterplots in *Figures 22* and *23* show how well a straight line fits each of these relationships.

Assumptions for each multiple imputation model are described in *Table 1*. Each assumption has its own strengths and weaknesses, each treating death due to disability in a different way. The results are presented in *Table 18*.

To ignore death completely would result in a stronger result ($p = 0.005$ vs. $p = 0.009$). To impute for those deaths would offer a slightly stronger result ($p \leq 0.008$ vs. $p = 0.009$), whereas to give them a hypothesised value of zero slightly weakens the result ($p \geq 0.015$ vs. $p = 0.009$). However, in the five models developed here, the results remain compatible with the main analysis and with each other.

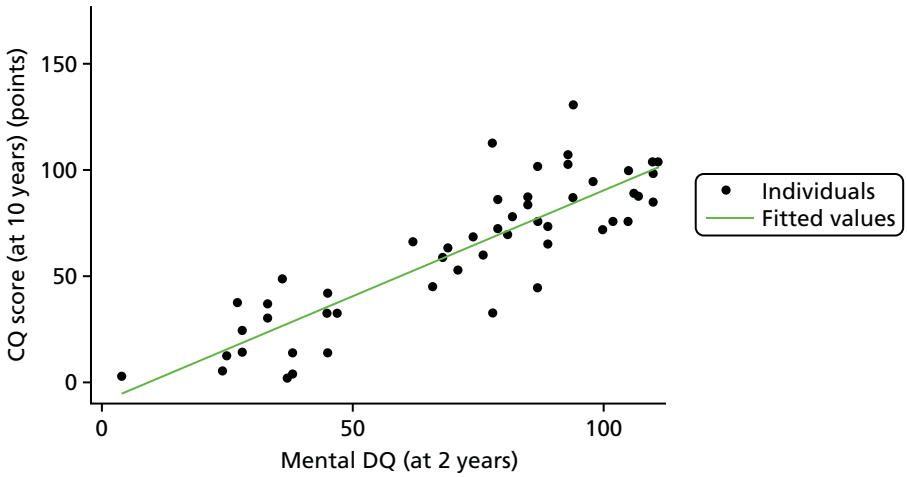


FIGURE 22 Relationship between 2-year and 10-year scores: mental DQ.

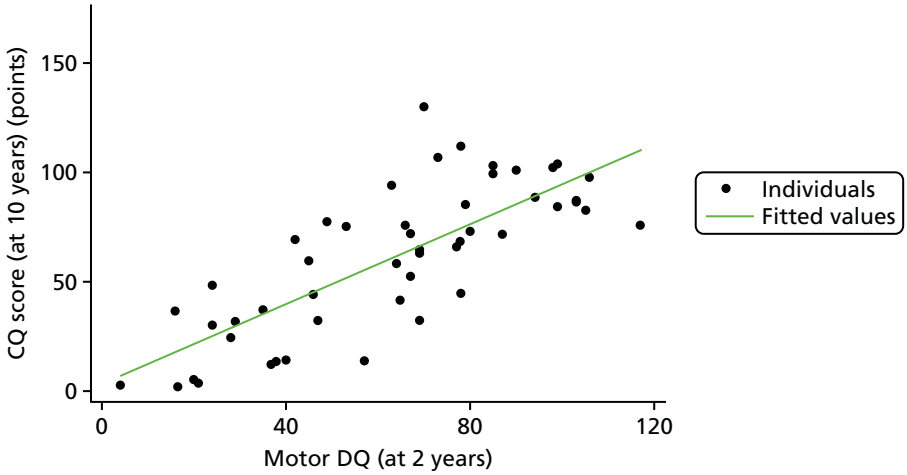


FIGURE 23 Relationship between 2-year and 10-year scores: motor DQ.

TABLE 18 Multiple imputation

Sensitivity analysis	*N* (D : S)	Trial arm, mean (SE)		Difference (95% CI)[a]	*p*-value[a]	Adjusted difference (95% CI)[b]	*p*-value[b]
		DRIFT	Standard treatment				
Imputation of cognitive DQ							
Assumption 1	36 : 33	65.24 (5.63)	50.81 (6.23)	14.43 (–2.10 to 30.96)	0.086	21.17 (5.66 to 36.68)	0.008
Assumption 2	36 : 33	65.42 (5.45)	50.87 (6.31)	14.54 (–1.98 to 31.07)	0.083	21.42 (6.21 to 36.64)	0.007
Assumption 3	36 : 33	62.95 (5.80)	49.41 (6.44)	13.55 (–3.84 to 30.93)	0.124	20.53 (4.49 to 36.56)	0.013
Assumption 4	36 : 33	62.80 (5.91)	49.58 (6.47)	13.22 (–4.49 to 30.93)	0.140	20.08 (3.79 to 36.38)	0.017
Assumption 5	34 : 31	66.85 (5.42)	53.70 (6.43)	13.14 (–3.67 to 29.96)	0.123	20.47 (4.62 to 36.31)	0.012

D, DRIFT; S, standard treatment.
a Crude estimates using a linear regression model.
b Linear regression model, adjusted for gender, birthweight and grade of IVH.
Using 'mi impute chained' with the following predictors: at randomisation – arm, gender, age at entry (days), centre, birthweight and gestation; at 2 years – shunt, seizure, speech, vision, hearing, disability, sitting, hand, mental DQ and motor DQ; at 10 years – vision, hearing, seizures, shunts, oxygen and CP (and indicator of death, where appropriate).

Subgroup analysis

Subgroup analyses were almost all selected a priori and explored using formal tests of interaction; maternal education was the only post hoc subgroup analysis. Given the small sample size in this study, these analyses were heavily underpowered, resulting in the risk of false-negative results. With this in mind, focus was concentrated more on the estimates and CIs than on the *p*-values. Subgroup analyses results are presented in *Table 19*.

The interaction effect mean differences can be interpreted as the effect of DRIFT compared with standard treatment in one subgroup relative to the effect in the other subgroup. Overall, no obvious differences were seen in the subgroups. Of all of these analyses, the only one that may warrant further consideration is gestation. After adjustment for birthweight and gender, the difference between the arms appeared to be greater for those with gestation ≥ 28 weeks (18.85 points) than for those with gestation < 28 days. This offers very weak evidence to suggest that DRIFT may be more effective for those with higher gestation; however, the unadjusted results were in the opposite direction. The unadjusted results for grade of IVH suggested that the DRIFT may be more effective in those with grade 4 IVH; however, adjustment weakened this result. Scores appeared to be much higher for those who were cared for with increased vigilance after the stopping period; however, the interaction is difficult to interpret because of the small samples provided in the cross-tabulation. As stated previously, the small sample sizes in each subgroup mean that these analyses are heavily underpowered and should be interpreted with caution.

Exploratory analysis

The team collected some additional information on the children's educational level at KS1 and KS2 (provided children had reached this level) from each of the children's named teacher. The expected level for children at KS1and KS2 is 2b and 4b, respectively. Those who were scored using the *P* levelling were below the level of the tests. Whether or not the child received SEN support was also recorded on the Client Service Receipt Inventory (CSRI),[42] along with data on speech and language therapy (SLT) attendance in the past 6 months and special school attendance in the past 12 months. All results are in *Table 20*.

As the sample sizes for educational levels are small (27 children with KS1 scores and 17 children with KS2 scores), more emphasis should be put on descriptives than on *p*-values. *Table 20* shows that children in the DRIFT arm were more likely than those in the standard treatment arm to score level 1 or above (71% vs. 46% at KS1; 63% vs. 44% at KS2). However, the number of children scoring above average was similar in each group (21% vs. 23% for KS1; 25% vs. 22% for KS2). Using the SEN data, it appears that the percentage of children receiving SEN support is similar in each arm, but slightly higher in the DRIFT arm.

More data were available for SLT and special school attendance. After adjustment, those in the DRIFT arm had lower odds (0.27) of special school attendance in the last 12 months than those in the standard treatment arm ($p = 0.059$). They also had lower odds (0.30) of attending SLT sessions in the previous 6 months ($p = 0.079$). These outcomes were not prespecified in the trial protocol and sample sizes here were very small; therefore, there should be no overinterpretation of the results.

Binary visual outcomes

This is an exploratory analysis with low numbers and multiple testing. Thus, any significance attached to findings (or lack of them) is not so important but, rather, the pattern of observations help to generate hypotheses for future study. *Table 21* gives the baseline characteristics of those who were followed up at 10 years.

Table 22 shows the visual results, as assessed by the blinded ophthalmologist. All negative outcomes are coded as '1' and positive outcomes as '0'. When comparing DRIFT with standard treatment, almost all differences are consistent with chance. Initially, before adjustment, the DRIFT arm had more horizontal

TABLE 19 Subgroup analysis for cognitive disability scores

Subgroup	*N* (D : S)	Subgroup specific, mean (SD)		Interaction[a] (95% CI)	*p*-value[a]	Interaction[b] (95% CI)	*p*-value[b]
		DRIFT	Standard treatment				
Gestation (weeks)							
< 28	15 : 10	62.51 (28.42)	42.86 (30.68)	3.20 (–33.72 to 40.11)	0.862	–18.85 (–54.67 to 16.98)	0.295
≥ 28	12 : 14	77.85 (31.06)	61.40 (38.07)				
Grade of IVH							
3	14 : 13	75.60 (26.07)	71.10 (36.34)	24.99 (–9.24 to 59.23)	0.149	15.65 (–19.80 to 51.11)	0.379
4	13 : 11	62.58 (33.56)	33.09 (22.05)				
Age (days)[c]							
< 21	15 : 15	75.22 (30.11)	60.71 (37.34)	–5.48 (–42.92 to 31.96)	0.770	–5.02 (–38.26 to 28.23)	0.762
≥ 21	12 : 9	61.96 (29.58)	41.96 (31.27)				
Dilation[d]							
Unilateral	4 : 4	64.00 (23.67)	35.37 (16.91)	15.72 (–35.36 to 66.81)	0.539	6.80 (–39.98 to 53.58)	0.771
Bilateral	23 : 20	70.25 (31.39)	57.34 (37.59)				
Gender							
Male	22 : 15	65.33 (31.60)	47.57 (34.38)	–5.27 (–47.71 to 37.17)	0.804	–3.60 (–42.06 to 34.85)	0.851
Female	5 : 9	86.90 (12.50)	63.86 (37.55)				

Subgroup	*N* (D : S)	Subgroup specific, mean (SD)		Interaction[a] (95% CI)	*p*-value[a]	Interaction[b] (95% CI)	*p*-value[b]
		DRIFT	Standard treatment				
Vigilance[e]							
Pre-enhanced	22 : 23	67.65 (33.09)	51.49 (34.82)	–43.45 (–117.82 to 30.92)	0.246	–15.24 (–84.51 to 54.03)	0.660
Post-enhanced	5 : 1	76.69 (6.40)	103.98 (0.00)				
Maternal education[f]							
Low[f]	10 : 11	64.16 (37.76)	52.10 (31.79)	–0.93 (–38.99 to 37.12)	0.961	–15.84 (–50.78 to 19.10)	0.366
High	17 : 12	72.37 (25.29)	59.38 (38.37)				

D, DRIFT; S, standard treatment.

a Crude estimates from the interaction term in the linear regression model.

b Estimates from the interaction term in the logistic regression model, adjusted for gender, birthweight and IVH, as appropriate.

c Age at randomisation.

d Dilation on ultrasonogarphy at randomisation.

e In 2006, the trial was temporarily stopped as the committee members were concerned about the large number of secondary haemorrhages in the DRIFT arm. After this time, seven more patients were recruited during an 'enhanced vigilance' period.

f Maternal education was collected at 10 years and, therefore, only classes as an indicator of education at baseline (to be viewed with caution). This was classed as 'low' if the mother left school at age 16 years and 'high' if the mother carried on with further education post age 16 years and/or went to university.

TABLE 20 Educational outcomes, by trial arm

Outcome	Trial arm, *n* (%)		OR[a] (95% CI)	*p*-value[a]	OR[b] (95% CI)	*p*-value[b]
	DRIFT	Standard treatment				
KS1 scores[c]						
Level 1 or above	10 (71)	6 (46)	2.92 (0.59 to 14.33)	0.187	7.37 (0.82 to 66.10)	0.074
Level P1–P8	4 (29)	7 (54)				
Unknown	25 (64)	25 (76)				
Level ≥ 2b[d]	3 (21)	3 (23)	0.91 (0.15 to 5.58)	0.918	1.24 (0.16 to 9.74)	0.840
KS2 scores[c]						
Level 1 or above	5 (63)	4 (44)	2.08 (0.30 to 14.55)	0.459	[e]	[e]
Level P1–P8	3 (38)	5 (56)				
Unknown	20 (51)	21 (55)				
Too young to assess	11 (28)	8 (21)				
Level ≥ 4b[f]	2 (25)	2 (22)	1.17 (0.12 to 10.99)	0.893	1.52 (0.13 to 18.31)	0.743
SEN support						
Yes	11 (65)	10 (56)	1.47 (0.38 to 5.72)	0.581	0.88 (0.14 to 5.39)	0.888
No	6 (35)	8 (44)				
Special school attendance in the past 12 months						
Yes	8 (29)	11 (48)	0.44 (0.14 to 1.39)	0.161	0.27 (0.07 to 1.05)	0.059
No	20 (71)	12 (52)				
SLT in last 6 months?						
Yes	9 (35)	11 (61)	0.34 (0.10 to 1.17)	0.087	0.30 (0.08 to 1.15)	0.079
No	17 (65)	7 (39)				

a Crude adjustments using logistic regression.
b Logistic regression model, adjusted for grade of IVH, gender and birthweight.
c The children analysed at KS1 do not necessarily match the children analysed at KS2.
d Level 2b is the average level of achievement at KS1.
e There were too few observations to carry out a test.
f Level 4b is the average level of achievement at KS2.

TABLE 21 Baseline characteristics for those having visual assessments at 10 years

Characteristic	*N* (D : S)	Trial arm, *n* (%)		N (3 : 4)	IVH grade, *n* (%)	
		DRIFT	Standard treatment		3	4
Gender	28 : 24	22 (79%)	15 (63%)	27 : 25	19 (70%)	18 (72%)
Mean birthweight (g)	28 : 24	1102 (336)	1322 (535)	27 : 25	1333 (537)	1064 (274)
IVH grade 4	28 : 24	14 (50%)	11 (46%)	–	–	–

D, DRIFT; S, standard treatment.

TABLE 22 Binary visual outcomes

Outcome	*N* (D : S)	Trial arm, *n* (%)		OR (95% CI)[a]	*p*-value[a]
		DRIFT	Standard treatment		
Possible/definite field loss	17 : 13	2 (12)	3 (23)	0.44 (0.06 to 3.16)	0.417[b]
Nystagmus	19 : 16	4 (21)	2 (13)	1.87 (0.29 to 11.84)	0.508[b]
Could not do rectangles (open or closed) vs. could	18 : 14	3 (17)	5 (36)	0.36 (0.07 to 1.88)	0.226[b]
Could not do postbox vs. could	19 : 14	1 (5)	1 (7)	0.72 (0.04 to 12.64)	0.824[b]
Poor binocular, left or right vision (all > 0)	20 : 16	5 (25)	7 (44)	0.43 (0.10 to 1.76)	0.240[b]
Strabismus	19 : 16	12 (63)	10 (63)	1.03 (0.26 to 4.07)	0.968[b]
Horizontal pursuit < 5	19 : 14	14 (74)	5 (36)	5.04 (1.13 to 22.50)	0.034[b]
Vertical pursuit < 5	19 : 13	12 (63)	5 (38)	2.74 (0.64 to 11.75)	0.174[b]
Horizontal saccade < 5	19 : 14	14 (74)	5 (36)	5.04 (1.13 to 22.50)	0.034[b]
Vertical saccade < 5	19 : 13	11 (58)	5 (38)	2.20 (0.52 to 9.30)	0.284[b]
Contour score of > 1	16 : 12	4 (25)	4 (33)	0.67 (0.13 to 3.47)	0.630[b]
		IVH grade, *n* (%)			
	N (3 : 4)	3	4	OR (95% CI)[a]	*p*-value[a]
Possible/definite field loss	16 : 14	0 (0)	5 (36)	–	0.014[c]
Nystagmus	19 : 16	1 (5)	5 (31)	8.18 (0.84 to 79.54)	0.070[b]
Could not do rectangles (open or closed) vs. could	18 : 14	3 (17)	5 (36)	2.78 (0.53 to 14.50)	0.226[b]
Could not do postbox vs. could	18 : 15	0 (0)	2 (13)	–	0.199[c]
Poor binocular, left or right vision (all > 0)	21 : 15	2 (10)	10 (67)	19.00 (3.11 to 116.1)	0.001[b]
Strabismus	20 : 15	9 (45)	13 (87)	7.94 (1.41 to 44.80)	0.019[b]
Horizontal pursuit < 5	18 : 15	7 (39)	12 (80)	6.29 (1.29 to 30.54)	0.023[b]
Vertical pursuit < 5	17 : 15	5 (29)	12 (80)	9.60 (1.86 to 49.48)	0.007[b]
Horizontal saccade < 5	18 : 15	7 (39)	12 (80)	6.29 (1.29 to 30.54)	0.023[b]
Vertical saccade < 5	17 : 15	4 (24)	12 (80)	13.00 (2.40 to 70.46)	0.003[b]
Contour score of > 1	17 : 11	2 (12)	6 (55)	9.00 (1.35 to 59.78)	0.023[b]

D, DRIFT; S, standard treatment.
a Crude adjustments.
b Logistic regression.
c Fisher's exact test was used when the independent variable was a perfect predictor.

pursuit scores that were < 5. However, after adjustment for IVH, gender and birthweight, this difference was reduced. This is unsurprising given that those in the DRIFT arm were suffering from less favourable baseline characteristics. The percentage of children who could not do rectangles was smaller in the DRIFT arm than in the standard treatment arm, a difference that was after adjustment for baseline factors. It may be that this, given the large number of tests, is a chance finding. These should be considered as exploratory and 'hypothesis generating'.

When comparing grade 3 IVH with grade 4, visual outcomes at 10 years were very different. For almost all binary outcomes, there was evidence to suggest that the odds of negative outcomes were higher for those in grade 4 than for those in grade 3. We observed no difference in the number of children who could not do rectangles. All other visual outcomes were substantially worse for the grade 4 children than for the grade 3 children.

Owing to perfect prediction with the covariates, most of the models could not be adjusted. 'Poor binocular, right or left vision' was defined as '1' for those who scored > 0 (and answered at least one) for all of the following: binocular – distance acuity single optotype; binocular – distance acuity; right – distance acuity single optotype; right – distance acuity croweded optotype; left – distance acuity single optotype; and left – distance acuity crowded optotype. If participants scored ≥ 0 for any of those questions, the score was defined as '0'. For 'poor binocular, right or left vision', gender and IVH were both perfect predictors. In total, 0% of girls with grade 3 IVH and 33% of girls with grade 4 IVH had 'poor binocular, right or left vision', compared with 13% of boys with grade 3 IVH and 75% of boys with grade 4 IVH.

Possible or definite field loss was defined as '1' if the binocular visual field was variably or definitely reduced and as '0' if normal. Of those with grade 4 IVH, 36% had field loss, compared with 0% of those with grade 3 IVH; therefore, this adjustment was not made for the arm comparison and the chi-squared test was used for the comparison between IVH grades. Nystagmus classed as 'None', 'in PP' or 'at extremes of gaze'. Only those with nystagmus 'in PP' were defined as '1' for nystagmus. Gender was a perfect predictor for this as 22% of boys had nystagmus, compared with 0% of girls; therefore, the adjusted model could not be performed. 'Could not do rectangles (open or closed)' was defined as '1' if the child 'could not do' either the open or closed rectangle and defined as '0' if they were 'normal' or only had 'some problems' for both. 'Could not do postbox' was defined in the same way. IVH was a perfect predictor for this, as 13% of those with grade 4 IVH could not do the postbox, compared with 0% of those with grade 3 IVH. Strabismus was defined as '1' for those who did not achieve 'normal' for the cover test unaided at 33 cm and as '0' for those who achieved 'normal'.

Neonatal outcomes at 2 years

Information collected from both groups at 2 years was compared and is presented in *Table 23*. As reported previously, there were eight deaths before the 2-year time point: five in the standard treatment arm and three in the DRIFT arm. There were 16 (41%) VP shunts in the DRIFT arm and 15 (39%) in the standard treatment arm.[32] Reservoirs were required by 28 (74%) children in the standard treatment arm and by 18 (46%) in the DRIFT arm (adjusted OR 0.27, 95% CI 0.10 to 0.76).

Harms

Despite the excess secondary haemorrhages in the DRIFT group, the primary outcomes were better and the secondary outcomes no worse than in the standard treatment group. It does not appear that secondary haemorrhages that occurred during the DRIFT procedure had a long-term detrimental effect.

Visual field defects were also no more frequent in the DRIFT group despite insertion of the occipital irrigation catheters.

High-resolution structural brain MRI at 10 years showed no evidence of damage associated with insertion of the DRIFT irrigation catheters. A larger proportion of the standard treatment group required ventricular reservoirs, and more residual frontal tracts associated with reservoirs were seen in the standard treatment group. There was no difference in ongoing neurosurgical problems between the treatment arms at age 10 years.

TABLE 23 Outcomes at 2 years

Outcome	Trial arm, *n* (%)		Difference[a] (95% CI)	*p*-value[a]	Difference[b] (95% CI)	*p*-value[b]
	DRIFT	Standard treatment				
Death						
Yes	3 (8)	5 (13)	0.55 (0.12 to 2.48)	0.437	0.45 (0.10 to 2.14)	0.317
No	36 (92)	33 (87)				
Shunt[c]						
Yes	16 (41)	15 (39)	1.07 (0.43 to 2.65)	0.890	0.99 (0.38 to 2.61)	0.982
No	23 (59)	23 (61)				
Reservoir						
Yes	18 (46)	28 (74)	0.31 (0.12 to 0.80)	0.015	0.27 (0.10 to 0.76)	0.013
No	21 (54)	10 (26)				

a Crude estimates using logistic regression.
b Logistic regression model, adjusted for birthweight, gender and grade of IVH.
c One patient in the DRIFT arm required a shunt after 2 years. Adding this child did not affect the result.

Chapter 4 Economic analysis of the costs and outcomes of the DRIFT intervention

Introduction

The National Institute for Health and Care Excellence (NICE) currently recommends that DRIFT should not be used routinely in the NHS, but only in the context of research.[43] If DRIFT were to be used in routine NHS care, it is important to consider whether or not the upfront costs of the procedure are justified by improvements in patient outcomes and/or reduced costs of care later in life. The original RCT and early follow-up at 2 years of age did not collect detailed resource use data or include an economic evaluation. Therefore, we cannot conduct a comprehensive economic evaluation comparing the cumulative costs of care, survival and QoL following DRIFT and standard care. In our primary economic evaluation, we conducted a cost–consequence analysis[44] assessing whether DRIFT increases or reduces NHS secondary care resource use since birth, and providing a more detailed snapshot of health care, social care and educational costs, productivity losses and QoL among survivors at follow-up after 10 years. If DRIFT improves outcomes or reduces costs of care at school age, it is likely to become more cost-effective over the future lifetime of survivors. In exploratory analyses, we use a decision analytical model to extrapolate costs and outcomes to age 18 years.

A total of 54 of the 77 children in the DRIFT study were recruited in Bristol and are the focus of the economic evaluation (*Figure 24*). We excluded children recruited at other centres in other countries owing to the expense and logistical difficulty of tracking down hospital notes and linking to routine hospital data in the years since birth.

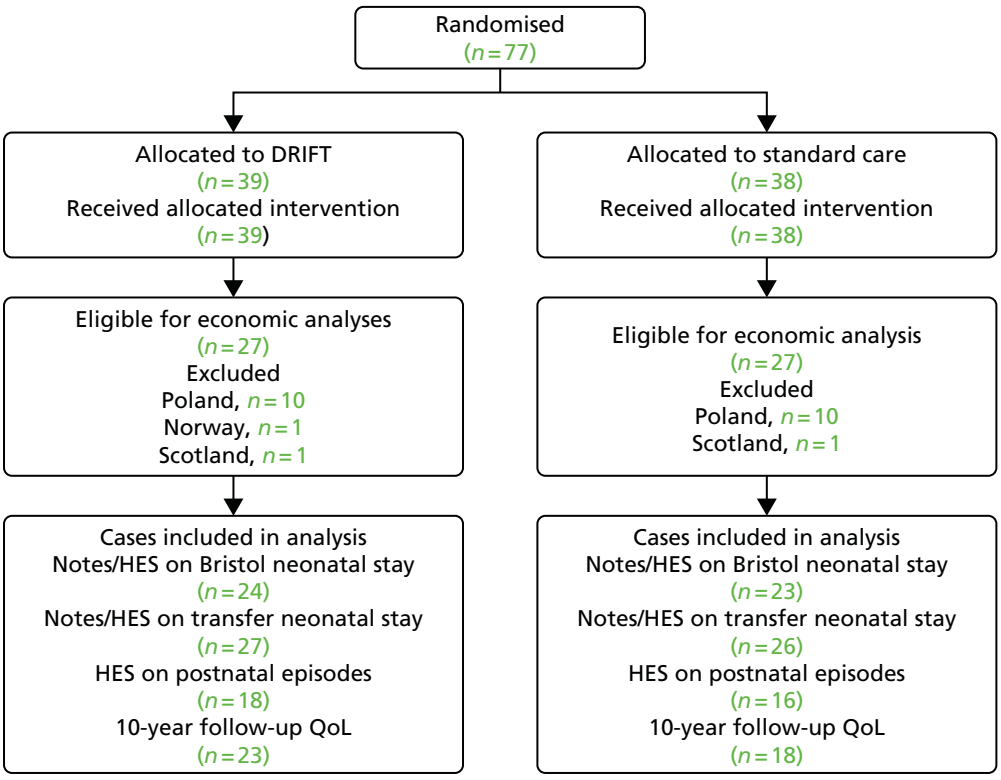


FIGURE 24 The CONSORT diagram for the cost–consequence analysis. CONSORT, Consolidated Standards of Reporting Trials; HES, Hospital Episode Statistics.

Methods

Resource use, data collection and valuation

The DRIFT procedure

Every infant randomised to the DRIFT group received the DRIFT procedure. A full description of the DRIFT procedure has been published previously.[30,45] Given that DRIFT is not currently recommended by NICE outside research and is not in widespread use, there is no national tariff for this procedure. Therefore, we used microcosting to estimate the likely cost to NHS hospitals if DRIFT were routinely provided at neonatal intensive care units (NICUs) with neurosurgical support. The use of some resources will depend on decisions about where and how to provide the DRIFT procedure. For example, some infants with PHVD would need to be transported to a NICU with neurosurgical support in order to receive the DRIFT procedure. However, the same would be true for infants being considered for other neurosurgical procedures (e.g. reservoir or shunt) as part of standard care. In the RCT, a high proportion of participants were transferred to Bristol from other hospitals. However, if in routine use at NICUs nationwide, fewer babies would need to be transferred over shorter distances. In our analysis, we included a cost of transport (£1101 one way[46]) for every participant who was transferred from Bristol to an outlying hospital after receiving DRIFT or standard care.

The resources used for the DRIFT procedure and their costs are summarised in *Table 24*. The amount of neurosurgical time will depend on whether the neurosurgical team is on-site or needs to travel from a nearby site. We assumed that neurosurgical support is on-site and, therefore, neurosurgical time to prepare for, and perform, the procedure would be approximately 2 hours, which includes time to remove the catheters after DRIFT has concluded. We assumed that, if used in routine practice, the procedure would be performed on the NICU under intravenous anaesthesia, as happened in the RCT.

The DRIFT procedure itself utilises disposable equipment and a small number of reusable items [i.e. pressure transducer, IVAC™ pump (Carefusion, Basingstoke, UK)]. Because the reusable equipment is relatively inexpensive (e.g. IVAC pump ≈£100) and can be reused a large number of times, we assumed that the proportionate capital cost for each baby is effectively £0. DRIFT uses infusions of artificial CSF, fibrinolytic and antibacterial drugs to irrigate the ventricles. DRIFT also requires daily microbiology screening of CSF and frequent monitoring by nursing staff of fluid infusion and drainage, necessitating one-to-one nursing. Some of these resources (e.g. artificial CSF) are 'variable' costs, in that the amount used increases as the number of days of DRIFT increases. Others (e.g. Alteplase) are 'fixed' costs, used only once. DRIFT was typically conducted for up to 5 days, but could be continued for longer. The number of days of the DRIFT procedure was extracted from the original trial records and hospital notes where available. Where unavailable, the days were imputed based on the mean number of days of DRIFT in all participants who received it. The need for frequent nurse monitoring during DRIFT may increase days on the intensive care unit (ICU) rather than the high-dependency unit (HDU) or special care unit (SCU). However, many babies would need to be on the ICU for other care needs. We extracted information on ICU, HDU and SCU days post randomisation (see *Initial Bristol neonatal stay*). For the DRIFT RCT, nurses received some training on the DRIFT procedure; we assumed that, if DRIFT were used routinely, this training would be part of general professional development and would have negligible incremental costs.

Standard care

Standard care required no intervention unless there was excessive head enlargement or clinical suspicion of raised intracranial pressure. The standard intervention, if required, was LP, removing 10 ml/kg CSF. Additional LPs depended on recurrence of these clinical signs. Children in the standard care arm received between zero and five LPs, with an average of two procedures per baby. However, LPs were also undertaken for children in the intervention arm both before and after receiving DRIFT. As LP is a cheap and minimally invasive routine procedure performed by neonatologists on the NICU, we assumed that its costs were bundled in with NICU day costs.

TABLE 24 Cost of the DRIFT procedure

Resource	Resource type	Units	Cost per unit (£)	Cost source	Total cost (£)
Surgeon time					
Neurosurgeon time (hours)	Fixed	2	138	*Unit Costs of Health and Social Care 2015*[47]	276
Anaesthesia					
IV morphine (ampoule)	Fixed	1	0.99	*British National Formulary*[48]	0.99
IV pancuronium (ampoule)	Fixed	1	5.00	*British National Formulary*[48]	5.00
Non-reusable equipment					
Circuit	Fixed	1	66.15	Hospital[a]	66.15
Collection bags	Fixed	3	13.29	Hospital[a]	39.87
Cannula	Fixed	2	51.01	Hospital[a]	102.02
Three-way taps	Fixed	2	0.47	Hospital[a]	0.94
Syringes	Fixed	6	0.06	Hospital[a]	0.36
Cavilon sticks (3M, Bracknell, UK)	Fixed	2	1.29	Hospital[a]	2.54
Tegaderm dressings (3M, Bracknell, UK)	Fixed	2	0.22	Hospital[a]	0.44
Giving set	Fixed	3	5.93	Hospital[a]	17.79
Orange needles	Fixed	2	0.01	Hospital[a]	0.02
Mersilk sutures (Ethicon, Bridgewater, NJ, USA)	Fixed	4	1.52	Hospital[a]	6.08
Steristrips (3M, Bracknell, UK)	Fixed	2	0.22	Hospital[a]	0.44
Umbilical cutdown pack	Fixed	1	21.90	Hospital[a]	21.90
Surgeon's gloves	Fixed	1	1.10	Hospital[a]	1.10
Arterial line	Variable p.d.	1	12.92	Hospital[a]	12.92
Infusions					
Alteplase (20 mg vial) (Actilyse; Boehringer Ingelheim Int., Ingelheim, Germany)	Fixed	1	45.00	*British National Formulary*[48]	45.00
Artificial CSF part 1 (500 ml)	Fixed	1	53.11	Hospital[a]	53.11
Artificial CSF part 2 (5 mg)	Fixed	1	9.20	Hospital[a]	9.20
Gentamicin (5 mg) (Genticin; Roche, Basel, Switzerland)	Fixed	1	5.40	Hospital[a]	5.40
Vancomycin (10 mg) (Vancocin; Flynn Pharma Ltd, Dublin, Ireland)	Fixed	1	7.70	Hospital[a]	7.70
Artificial CSF part 1 (500 ml)	Variable p.d.	2	53.11	Hospital[a]	106.22
Artificial CSF part 2 (5 mg)	Variable p.d.	2	9.20	Hospital[a]	18.40
Gentamicin (5 mg)	Variable p.d.	2	5.40	Hospital[a]	10.80
Vancomycin (10 mg)	Variable p.d.	2	7.70	Hospital[a]	15.40
Screening					
CSF MCS test	Variable	1	8.00	Hospital[a]	8.00
Total fixed cost (£)					662.09
Total variable cost per day (£)					163.74

MCS, microbiology culture and sensitivity; p.d., per diem.
a North Bristol NHS Trust.

Subsequent neurosurgical procedures to manage post-haemorrhagic ventricular dilatation during the initial neonatal stay

In standard care, if LPs failed to drain enough CSF to normalise head growth, a ventricular reservoir was indicated. If DRIFT was followed by persistent enlargement of ventricles and excessive head growth despite LPs, a ventricular reservoir was also used. If an infant in either group required repeated reservoir taps to control head growth, a VP shunt was indicated.[32] We used the discharge summary and letter to record the number of babies who had reservoir or VP shunt procedures during the initial neonatal stay after randomisation. We used a microcosting approach (*Tables 25* and *26*) to estimate the cost of these procedures during the initial neonatal stay.

Initial Bristol neonatal stay

Information on the initial neonatal stay in Bristol was extracted from two complementary sources: (1) the hospital notes and (2) linked Hospital Episode Statistics (HES) provided by NHS Digital (under data sharing agreement DARS-NIC-30560-W4V1T-v0.5; Copyright 2016, reused with the permission of The Health & Social Care Information Centre. All rights reserved). We excluded neonatal days at Bristol or outlying hospitals prior to randomisation.

TABLE 25 Cost of reservoir

Resource	Units	Cost per unit (£)	Source	Total cost (£)
Neurosurgeon A (hours)	1.5	138.00	*Unit Costs of Health and Social Care 2015*[47]	207.00
Neurosurgeon B (hours)	1.5	138.00	*Unit Costs of Health and Social Care 2015*[47]	207.00
Theatre time (minutes)	180	5.00	Hospital[a]	900.00
Radio-opaque proximal catheter	1	127.20	Hospital[a]	127.20
Reservoir (10 mm)	1	223.20	Hospital[a]	223.20
Total cost (£)				1664.40

a North Bristol NHS Trust.

TABLE 26 Cost of shunt

Resource	Units	Cost per unit (£)	Source	Total cost (£)
Neurosurgeon A (hours)	2	138.00	*Unit Costs of Health and Social Care 2015*[47]	276.00
Neurosurgeon B (hours)	2	138.00	*Unit Costs of Health and Social Care 2015*[47]	276.00
Theatre time (minutes)	150	5.00	Hospital[a]	750.00
Medium pressure valve (3.5 cm × 1.8 cm)	1	506.40	Hospital[a]	506.40
Radio-opaque distal catheter	1	454.80	Hospital[a]	454.80
Radio-opaque proximal catheter	1	127.20	Hospital[a]	127.20
Disposable catheter passer	1	49.20	Hospital[a]	49.20
Total cost (£)				2439.60

a North Bristol NHS Trust.

From Bristol hospital discharge summaries and letters, we extracted the number of days the participant spent on the ICU, HDU or SCU before discharge home or transfer to an outlying hospital. Hospital services in Bristol have been reconfigured in the years since the original RCT and a proportion of discharge summaries and letters were untraceable (10 out of 54, 18.5%). In most cases, the medical notes detailed the breakdown of stay by ICU/HDU/SCU days but, in some cases, only overall length of NICU stay was available without a more detailed breakdown. We used data linkage to HES[49] data to provide a more complete picture of the Bristol neonatal stay. HES records care provided to all NHS and privately funded patients treated in English NHS hospitals. For participants ($n = 42$) recruited at the Bristol site who survived and whose parents consented to data linkage, we sent identifiers (date of birth, NHS number, gender and postcode at birth) to NHS Digital which matched and extracted data on every episode of hospital care. This included the date of the admission, clinical details (e.g. diagnoses, procedures), length of stay and the hospital providing the care. HES data did not provide a breakdown of NICU stay by ICU/HDU/SCU days.

After exclusion of pre-randomisation and duplicate episodes, HES data identified 696 episodes of care provided by 37 different hospitals between birth and 31 March 2016, including at least one episode for all 42 participants for whom linkage was attempted. For the vast majority of episodes (> 99%), there was an exact match on NHS number, date of birth and gender, indicating that specificity is likely to be high. In 37 out of 42 (88.1%) participants, HES data identified the initial Bristol neonatal stay, indicating that a minority of episodes of care were not identified in HES.

In total, 47 out of 54 (87%) participants had data on the initial Bristol neonatal stay: in both HES and hospital notes ($n = 34$), in hospital notes alone ($n = 10$) or in HES alone ($n = 3$). We used *NHS Reference Costs 2014 to 2015*[46] (*Table 27*) to cost NICU care. When details were available from the hospital notes, we used specific costs for each day of ICU, HDU and SCU care. We calculated the mean proportion of all Bristol NICU days spent on each unit type and used this to impute a weighted daily NICU cost for those patients for whom only overall NICU length of stay was recorded.

Post-Bristol (transfer) neonatal stay

A high proportion of babies (36 out of 54, 66.7%) were transferred from Bristol to outlying hospitals for ongoing care after the initial neonatal stay. We requested details of these transfer episodes, including a breakdown by ICU/HDU/SCU days, from consultants working at these hospitals; however, in some cases (6 out of 36, 16.7%), no details of the transfer episode could be identified. Again, we used linked HES data to provide a more complete picture of the transfer neonatal stay for the participants whose parents consented to data linkage. In total, 35 out of 36 (97%) participants had data on the transfer neonatal stay: in both HES and hospital notes ($n = 15$), in hospital notes alone ($n = 15$) or in HES alone ($n = 5$). Piecing together multiple sources of data for Bristol NICU and transfer stays required some judgement, for example if the discharge date on a discharge summary and HES disagreed. These conflicts were generally minor and judgements were made while the analyst was blind to randomised allocation.

We used *NHS Reference Costs 2014 to 2015*[46] (see *Table 27*) to cost transfer NICU care. When details were available from the hospital notes, we used specific costs for each day of ICU, HDU and SCU care. Preliminary analysis of these notes indicated that a high proportion of transfer NICU care was provided on

TABLE 27 Unit costs of initial hospital stay

Resource	Unit cost (£)	Source
ICU day	1176.47	*NHS Reference Costs 2014 to 2015*[46]
HDU day	847.15	*NHS Reference Costs 2014 to 2015*[46]
SCU day	532.95	*NHS Reference Costs 2014 to 2015*[46]
Neonatal critical care transportation (one way)	1101.00	*NHS Reference Costs 2014 to 2015*[46]

the SCU. Therefore, if details were not available, we multiplied the total NICU days by the cost of a SCU day. Some babies were transferred to more than one hospital before discharge. For each of these subsequent transfers, a cost for neonatal critical care transportation was applied.

NHS secondary care post initial neonatal stay

We used linked HES data to identify NHS inpatient and day case care post initial neonatal stay until 31 March 2016 for participants ($n = 42$) recruited at the Bristol site who survived and whose parents consented to 10-year follow-up and data linkage. However, six of these participants lived in Wales and an additional two were known to have emigrated from England soon after birth. Therefore, these analyses are restricted to the remaining 34 participants. As recruitment took place over a range of years (2003–6), the duration of follow-up varied by participant from 9.3 years to 13.2 years with a mean of 11.3 years. Mean duration of follow-up was similar between trial arms (DRIFT 11.2 years, standard care 11.3 years).

Resource use at long-term follow-up

Data on resource use at long-term follow-up were provided by parents completing a questionnaire based on the CSRI,[42] with the assistance of a member of the research team if required. Questions related to parent(s)' productivity, child's education including SEN, child's outpatient and emergency department (ED) care in the last 12 months, child's primary health and social care use and additional expenses incurred because of the child's health in the last 6 months. *NHS Reference Costs 2014 to 2015*[46] and *Unit Costs of Health and Social Care 2015*[47] (*Table 28*) were used, where available, to value health and social care.

Parents were asked whether or not their child had attended a special school or special unit in the last 12 months. If the answer was in the affirmative, they were asked to specify whether it was a special unit within a mainstream school or a special school. If the child attended a special school, parents were asked if it was a day school or boarding school and if it was government or privately funded. All parents were

TABLE 28 Unit costs of health and social care at 10-year follow-up

Resource	Unit cost (£)	Source
Outpatient visit	114.50	*NHS Reference Costs 2014 to 2015*[46]
ED visit	131.92	*NHS Reference Costs 2014 to 2015*[46]
School nurse	53.70	*NHS Reference Costs 2014 to 2015*[46]
Health visitor	51.21	*NHS Reference Costs 2014 to 2015*[46]
Dentist	142.57	*NHS Reference Costs 2014 to 2015*[46]
GP	44.00	*Unit Costs of Health and Social Care 2015*[47]
Paediatrician	174.00	*Unit Costs of Health and Social Care 2015*[47]
Optician	30.00	The College of Optometrists[50]
Child development centre	46.23	Romeo *et al.*[51]
Speech therapist	92.50	*NHS Reference Costs 2014 to 2015*[46]
Hearing specialist	76.58	*NHS Reference Costs 2014 to 2015*[46]
Family/individual counselling	90.56	*NHS Reference Costs 2014 to 2015*[46]
Home help/care worker (1 hour)	24.00	*Unit Costs of Health and Social Care 2015*[47]
Day centre care (8 hours)	136.00	*Unit Costs of Health and Social Care 2015*[47]
Social worker (1 hour)	55.00	*Unit Costs of Health and Social Care 2015*[47]
After-school club	6.00	Assumption

GP, general practitioner.

also asked whether their child had been given a statement of SEN. We wrote to the schools of all children recruited in Bristol and consenting to school-age follow-up asking for details on additional funding received for education and health care plan or statement of SEN. Of 38 schools that responded, 24 confirmed that they received additional funding. In many cases, the value of funding was not reported; where reported, it ranged from £5280 to £36,000 with a median of £21,026. Owing to the large number of missing data on schooling costs, we used published unit costs to differentiate the costs of schooling.

Mainstream schools teaching children with SEN have a notional SEN budget and are expected to meet the cost of additional support for pupils with SEN up to £6000 per pupil per year.[52] This represents a notional average of SEN costs, with some children requiring more or less support. Official Department for Education (DfE)[53] statistics differentiate between the average total expenditure per pupil per year in a local authority maintained mainstream secondary school (£6125 in 2014/15) and special schools (£23,078). DfE figures do not distinguish between special schools and special units in mainstream schools. However, it is likely that the complexity of needs and, therefore, expenditure per pupil is, on average, lower in special units. We assumed that the cost of special unit education in a mainstream school was an average of the additional costs of SEN and special school education (*Table 29*). We used these DfE figures in our primary analysis.

The cost of special schooling varies considerably depending on the individual needs of the pupil. For example, a report commissioned by the National Association of Independent Schools and Non-Maintained Special Schools (NASS)[55] estimated that the total annual cost of special education and care varied by 78%, from £93,711 for day-only education up to £167,268 for 52-week boarding school. The reason for the large discrepancy between DfE estimates and NASS estimates appears to be that the latter includes therapy costs, family disability living allowance, equipment, short breaks, travel and facilities costs that are excluded from the DfE figures. After inclusion of these costs, the NASS report concluded that costs at independent special schools are similar to those at equivalent local authority maintained special schools. Therefore, in sensitivity analyses (see *Table 29*), we use NASS figures to estimate the unit costs of education and educational care in special schools.

Parents/carers were asked if they were currently employed and, if so, how many hours they worked on average per week. They were also asked to provide the same information for their partner, if applicable. We used this to estimate the household hours worked per week. We used *Annual Survey of Hours and Earnings: 2015 Provisional Results*[54] mean gross pay per hour to estimate household weekly income from employment (see *Table 29*). This will not detect any impact of child health on the type of employment that the parents/carers are willing and able to take up. Therefore, we asked a supplemental question about whether the main source of household was from earned income or benefits to get a better overview of household income.

TABLE 29 Unit costs of productivity and schooling at 10-year follow-up

Resource	Unit cost (£)	Source
Productivity (hourly gross pay)	15.70	*Annual Survey of Hours and Earnings: 2015 Provisional Results*[54]
Mainstream school (per year)	6125	DfE[53]
Additional cost of SEN education in mainstream school (per year)	6000	Education Funding Agency[52]
Special unit, mainstream school (per year)	17,601	Assumption
Special school (per year)	23,078	DfE[53]
	93,711[a]	Clifford and Theobold[55]

a Figure used in sensitivity analysis.

Although HES data sets on outpatient (from 2003) and ED (from 2007) care are available, we focused exclusively on HES data on day case and inpatient care. Outpatient and ED data sets were designated as 'experimental' statistics at the start of the DRIFT follow-up period. Our decision was based on the high cost of acquiring linked data and probable lower cost and impact on our conclusions of NHS outpatient and ED care. A snapshot of emergency, outpatient and other community care not captured by HES was elicited from parents at the 10-year follow-up (see *Resource use at long-term follow-up*).

Hospital stays with multiple episodes of care were concatenated and the dominant Healthcare Resource Group (HRG) code was used to estimate the cost of care. HRGs, which group clinically similar admissions requiring similar levels of resources, are the basis of hospital reimbursements for care provided. As care occurred over a period of > 10 years, several different versions of HRG codes were recorded. We applied the most recent available *NHS Reference Costs 2014 to 2015*[46] for the HRG and, where necessary, inflated the cost to 2014/15 values assuming 2.5% inflation per annum.[56] Based on data on admission type and length of stay, each admission was classified as a 'day case', 'elective long stay', 'non elective short stay' or 'non elective long stay', as reference costs vary by admission type. In some cases, hospitals receive additional reimbursements if patients spend an unexpectedly long time in hospital [excess bed-day (EBD)]. For each HRG, we calculated the trim point (the days after which EBD payments apply) and, for patients whose hospital stay exceeded the trim point, estimated the EBD cost based on the national EBD reference cost for that HRG. All costs of care occurring after the first year of life were discounted at 3.5% per annum in line with NICE guidance.[57]

Health-related quality of life

In addition to the cognitive, functional and other outcomes described in earlier sections of this report, parents were also asked to complete two generic measures of their child's health-related QoL (HRQoL) at the 10-year follow-up. The measures, the Health Utilities Index – 3 (HUI3)[58,59] and the EuroQol-5 Dimensions (EQ-5D), five-level version (EQ-5D-5L),[60] are preference-based measures, which produce a single 'utility' score anchored at best possible HRQoL (score 1) and HRQoL equivalent to death (score 0). We selected the HUI3 [covering eight attributes: (1) vision, (2) hearing, (3) speech, (4) emotion, (5) pain, (6) ambulation, (7) dexterity and (8) cognition] to allow direct comparison with previous work in neurodevelopmental disability in childhood.[61] We selected the EQ-5D-5L [covering five attributes: (1) mobility, (2) self-care, (3) usual activities, (4) pain/discomfort and (5) anxiety/depression] as it is commonly used in the UK by NICE[57] to judge the cost-effectiveness of new medical technologies. A youth version of the EQ-5D [EuroQol-5 Dimensions – Youth (EQ-5D-Y)] is now available with modified age-appropriate language for self-completion by children aged 7–12 years. However, owing to the prevalence of cognitive impairment and CP, parents completed the questionnaire on behalf of their child. As the EQ-5D-5L measures five levels on each attribute (rather than three levels in the EQ-5D-Y), we chose the EQ-5D-5L as potentially more sensitive to differences in participants' QoL. The EQ-5D also includes a visual analogue scale (VAS) that parents can use to rate their child's health today on a scale from 0 (worst health imaginable) to 100 (best health imaginable).

The UK adult value set[62] was employed to estimate the EQ-5D-5L utility score; for the HUI3 scores, the multiattribute health status classification system was used.[59] Both the HUI3 and the EQ-5D-5L are designed to be used prior to randomisation, and at repeated intervals post randomisation, to calculate quality-adjusted life-years (QALYs). In the UK, NICE favours QALYs when comparing the cost-effectiveness of different medical technologies within, and between, different patient populations. In our primary analysis, we report the mean score among survivors completing the questionnaire at the 10-year follow-up. In sensitivity analyses, we also present mean scores after including, with a score of zero, participants known to have died before 10-year follow-up.

Decision analysis model methods

We developed a simple decision analytical model to estimate the cost-effectiveness (cost per QALY) of DRIFT compared with standard care from birth to age 18 years. The primary perspective was that of NHS and Personal Social Services in accordance with NICE guidance.[57] In secondary analysis, we broaden the

perspective to include education costs. We initially planned a discrete health state Markov model stratifying children by the degree of disability and survival (none, mild, moderate, severe, dead), such as that outlined by Petrou and Khan.[18] In such a model, each health state would be assigned a cost representing the costs of care and a utility score representing the impact on the individual's HRQoL. However, the small sample size and infrequent follow-up in the DRIFT study meant that we could not reliably estimate transition probabilities between health states of a discrete health state model.

Therefore, we developed a simple two-state (alive or dead) Markov cohort decision model with a 1-year cycle length based on parameters derived directly from DRIFT trial data among participants recruited at Bristol. To estimate costs and health benefits, we assumed that transitions between health states occur halfway through each cycle (i.e. a half-cycle correction). We used a 3.5% annual discount rate for both costs and QALYs. The following model parameters, stratified by trial arm, were derived from DRIFT trial data: (1) cost of DRIFT (microcosting), (2) cost of remainder of NICU stay (hospital notes and HES data), (3) cost of postnatal inpatient care from age 0 to 2 years (HES data), (4) cost of postnatal inpatient care from age 2 to 10 years (HES data), (5) mortality from age 0 to 2 years (trial follow-up), (6) mortality from age 2 to 10 years (trial follow-up), (7) EQ-5D-5L index scores at 10-year follow-up (parent report), (8) 12-month cost of ambulatory hospital care at 10-year follow-up (parent report), and (9) 6-month cost of primary and community care at 10-year follow-up (parent report). We assigned probability distributions to all these stochastic parameters. Where cost and mortality rate parameters span > 1 year (e.g. postnatal inpatient cost or mortality from age 2 to 10 years) or < 1 year (e.g. primary and community care), we annualised and made the simplifying assumption that these costs and rates were constant across the years. We provide more detail on these parameters and their probability distributions in *Results*.

Inevitably, a model based on scant data requires a number of large assumptions. We made the following key assumptions in the model: (1) mortality between age 10 and 18 years is zero in both arms of the trial; (2) the EQ-5D-5L scores observed among survivors at 10-year follow-up are representative of scores among survivors at all ages; (3) the costs of education observed among survivors at 10-year follow-up are representative of these costs among all school ages (4–18 years) survivors; (4) the ambulatory- and community care costs observed among survivors at 10-year follow-up are representative of costs among survivors at all ages; and (5) we used the results of the unadjusted analyses comparing costs and outcomes between DRIFT and standard care. We tested the sensitivity of our model findings to some of these key assumptions.

Analysis

Primary economic evaluation: within-trial cost–consequence analysis

A cost–consequence analysis was conducted to compare the costs and effects of DRIFT with standard care. The analysis included neonatal stay costs, NHS secondary care costs up to 31 March 2016 and a snapshot of broader NHS costs, social care costs, educational costs, family expenses and productivity losses at 10 years' follow-up. Participants were analysed according to the treatment group to which they were randomised (i.e. an ITT approach) and we report on all available cases for each analysis.

Mean resource use and mean costs per patient were estimated in both trial arms. Regressions using ordinary least squares (OLS) and a general linear model (GLM) using a gamma family and log-link were employed to obtain the differences in mean costs between DRIFT and standard care arms. Gamma log-link GLM is commonly used to analyse small samples of skewed cost data (i.e. high-cost outliers).[63] As the point estimates and CIs were similar between OLS and GLM models, we present only the OLS results. Logistic regression was used to evaluate differences in binary outcomes. In line with the primary outcome analysis, results are presented 'unadjusted' and 'adjusted' for the baseline covariates: grade of IVH, birthweight and gender. Stata 14.1 was used for all health economic analyses.

Sensitivity analyses

We conducted the following sensitivity analyses to test the robustness of our primary analyses to different assumptions:

- using NASS costs of special education, which include health and social care needs while at special school, to provide a less conservative estimate than the DfE figures
- imputing EQ-5D-5L scores of zero for participants known to have died before 10-year follow-up
- no discounting of costs that occur in the years after birth.

Secondary economic evaluation: decision analysis model

The results of the model were summarised using incremental cost-effectiveness ratios (ICERs), the incremental net monetary benefit (INMB) statistic and cost-effectiveness acceptability curves.[64] We used the model to judge the probability that DRIFT is cost-effective at age 18 years against the NICE threshold of £20,000 to £30,000 per QALY. In our analyses, we used the lower figure ($\lambda = £20,000$) in calculating INMB. We conducted a probabilistic sensitivity analysis to quantify uncertainty about the cost-effectiveness of DRIFT. Monte Carlo simulation was used to repeatedly draw a randomly selected estimate of each model parameter from its estimated distribution. We used a conventional number of iterations ($n = 10,000$) to empirically estimate the uncertainty surrounding the mean INMBs calculated from the model. The model was built in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) and programmed in Visual Basic for Applications® (Microsoft Corporation, Redmond, WA, USA) to run the simulation.

Because of the number of large assumptions required to estimate cost per QALY, we consider the results of the decision analysis model to be exploratory. We used deterministic sensitivity analyses to test the impact of the following key assumptions on the findings of the model.

- estimating costs and outcomes at age 10 years rather than age 18 years
- using costs and utility scores adjusted for baseline covariate, gender, IVH grade and birthweight
- including educational costs
- using HUI3 rather than EQ-5D-5L utility scores.

Results

Participants included in the economic analyses

The numbers of participants with data available for each analysis differ (see *Figure 24*). Participants included in the analysis of initial neonatal costs were relatively similar to all participants recruited in Bristol in terms of birthweight, IVH grade 4 and gender (*Table 30*). However, participants included in the post-neonatal hospital cost and 10-year follow-up analyses tended to have a higher birthweight and were less likely to have grade 4 IVH. This is unsurprising given that these participants are survivors. The imbalance between DRIFT and standard care participants in terms of birthweight and gender widened slightly among participants included in the post-neonatal hospital cost and 10-year follow-up analyses.

Initial hospitalisation

Participants allocated to DRIFT had irrigation therapy for an average of 5.2 days at an estimated cost of £1513 per participant (*Table 31*). Some of this initial cost of DRIFT was offset by the fact that fewer patients had reservoir procedures during the neonatal stay. Participants allocated to DRIFT tended to spend fewer days in the Bristol NICU (mean 29.8 days) than participants allocated to standard care (mean 40.3 days). In contrast, participants allocated to DRIFT tended to stay longer in outlying hospitals (mean 38.5 days) after transfer from the Bristol NICU than participants allocated to standard care (mean 18.2 days). The total mean costs of the neonatal stay were higher in patients who had DRIFT, but the CI was wide and included zero (unadjusted mean difference £6556, 95% CI –£11,161 to £24,273). The finding was sensitive to adjustment for covariates, particularly birthweight. After adjustment for birthweight, gender and IVH grade,

TABLE 30 Cases available for economic analysis

	Trial arm			
	DRIFT		Standard care	
Case characteristics	*N*		*N*	
All Bristol participants	27		27	
Birthweight, mean (SD)		1045 (332)		1285 (502)
Gestation, weeks (SD)		27.2 (2.5)		28.0 (2.8)
IVH grade 4, *n* (%)		12 (44)		14 (52)
Male, *n* (%)		23 (85)		18 (67)
Initial neonatal and transfer costs	24		23	
Birthweight, mean (SD)		1059 (345)		1273 (518)
Gestation, weeks (SD)		27.3 (2.6)		28.0 (2.9)
IVH grade 4, *n* (%)		11 (46)		11 (48)
Male, *n* (%)		20 (83)		15 (65)
Post-neonatal hospital costs	18		16	
Birthweight, mean (SD)		1073 (346)		1375 (577)
Gestation, weeks (SD)		27.4 (2.5)		28.7 (2.9)
IVH grade 4, *n* (%)		8 (44)		5 (31)
Male, *n* (%)		18 (100)		9 (56)
10-year QoL and resource use	23		17	
Birthweight, mean (SD)		1074 (345)		1353 (568)
Gestation, weeks (SD)		27.5 (2.6)		28.5 (2.9)
IVH grade 4, *n* (%)		10 (43)		7 (41)
Male, *n* (%)		20 (87)		10 (59)

estimated mean costs of neonatal care were lower in patients who had DRIFT although CIs were still wide and included zero (adjusted mean difference –£3056, 95% CI –£19,449 to £13,335).

Postnatal hospital admissions and total NHS secondary care costs

Participants allocated to DRIFT spent an average of 19.4 days in hospital up to age 2 years and an average of 26.6 additional days in hospital between age 2 years and 31 March 2016 (see *Table 31*). Participants allocated to standard care spent fewer days in hospital than participants allocated to DRIFT (8.8 days, 0–2 years; 18.5 days, 2 years onwards; see *Table 31*). The most common HRG chapters for postnatal admission episodes were 'Diseases of childhood' (335 out of 573; 58.5%), 'Nervous system' (71 out of 573; 12.4%), 'Mouth, head, neck and ears' (34 out of 573; 5.9%) and 'musculoskeletal system' (31 out of 573; 5.4%).

The unadjusted total costs of hospital care after the initial neonatal stay were higher in participants allocated to DRIFT (unadjusted mean difference £3413, 95% CI –£12,408 to £19,234). This finding was very sensitive to adjustment for covariates, particularly gender and birthweight. After adjustment, the estimated mean cost among participants allocated to DRIFT was lower (adjusted mean difference –£9739, 95% CI –£27,558 to £8080). In sensitivity analysis 3, the adjusted mean cost difference was somewhat larger if costs were not discounted (adjusted mean difference –£12,348, 95% CI –£33,603 to £8907).

TABLE 31 Cost of neonatal and postnatal secondary care, by trial arm

	Trial arm							
	DRIFT			Standard care			Difference in mean costs, £ (95% CI)	
Secondary care	*n*	Mean (SD) units	Mean (SD) costs, £	*n*	Mean (SD) units	Mean (SD) costs, £	Unadjusted	Adjusted[a]
DRIFT	24	5.2 (1.7) days	1513 (276)					
Reservoir during NICU stay	24	0.54 (0.51)[b] procedures	902 (847)	23	1.0 (0.30)[b] procedures	1664 (502)		
VP shunts during NICU stay	24	0.25 (0.44)[b] procedures	610 (1080)	23	0.26 (0.45)[b] procedures	636 (1095)		
Bristol NICU stay	24	29.8 (27.3) days	26,850 (22,367)	23	40.3 (42.5) days	33,150 (35,474)		
Transfer NICU stay	27	38.5 (52.1) days	31,489 (34,705)	26	18.2 (26.5) days	15,382 (19,538)		
Subtotal neonatal costs	24		59,395 (29,411)	23		52,839 (30,820)	6556 (–11,161 to 24,273)	–3056 (–19,449 to 13,335)
Postnatal admissions (0–2 years)	18	19.4 (26.1) days	8768 (9021)	16	8.8 (10.4) days	5732 (8053)		
Postnatal admissions (2 years onward)	18	26.6 (63.8) days	15,293 (21,118)	16	18.5 (21.5) days	14,907 (19,116)		
Subtotal postnatal discounted costs	18		24,051 (22,540)	16		20,638 (22,679)	3413 (–12,408 to 19,234)	–9739 (–27,558 to 8080)
Total NHS inpatient costs	18		86,893 (39,829)	16		75,009 (44,274)	11,884 (–17,491 to 41,259)	–20,963 (–49,213 to 7269)

a Adjusted by gender, IVH grade and birthweight.
b Only those patients who had the procedure during the initial neonatal or transfer stay. Some patients had these procedures during readmissions. The costs of these readmission procedures are captured under postnatal admissions.

Total unadjusted NHS inpatient costs since birth were higher in participants allocated to DRIFT (unadjusted mean difference £11,884, 95% CI –£17,491 to £41,259). Again, this finding was very sensitive to adjustment for birthweight and gender (adjusted mean difference –£20,963, 95% CI –£49,213 to £7269).

Use of ambulatory health and social care at ten-year follow-up

There was little evidence of a difference in emergency and outpatient care in the last 12 months at the 10-year follow-up (*Table 32*). Parents of participants in both arms of the trial reported an average of just over 0.4 visits to the ED and just over 2.8 outpatient clinic visits. The adjusted difference in mean costs was marginally higher in participants allocated to DRIFT (adjusted mean difference £2, 95% CI –£264 to £267). The costs of other ambulatory care during the last 6 months were higher in participants randomised to standard care (adjusted mean difference –£108, 95% CI –£596 to £380), but the CI was wide. In free text, parents of participants in both arms noted a wide range of other therapies including orthotics, physiotherapy, occupational therapy, hydrotherapy and music therapy, although the frequency of these therapies was often not recorded.

Family income, expenses and child's educational needs

Overall, a similar proportion of parents/carers were employed at the 10-year follow-up (*Table 33*). Including 22 cases where a partner's employment status was also reported, 68% (23 out of 34) of parents/carers of DRIFT participants were employed and 64% (18 out of 28) of parents/carers of standard care participants were employed. The average working hours were 36.9 in the households of participants who received DRIFT (estimated weekly income of £580), compared with 38.1 (estimated weekly income of £599) in the households of participants who received standard care. However, a lower proportion of households of participants who received DRIFT had benefits as their main source of income (adjusted OR 0.23, 95% CI 0.04 to 1.22), although the CI included 1.

A similar percentage of parents of participants in both arms reported that their child had a SEN statement (see *Table 33*). However, a higher percentage of parents of participants in the standard care arm reported that their child attended a special unit or special school (adjusted OR 0.13, 95% CI 0.02 to 0.82). Owing to the high cost of special schooling, this is potentially economically important; the adjusted mean difference in estimated annual school costs was –£5321, 95% CI –£9772 to –£870. In sensitivity analyses, if higher NASS estimates of the costs of special schooling are used, the adjusted difference in estimated school costs becomes much higher –£35,122, 95% CI –£58,546 to –£11,699. Other family expenses reported by parents in free text included equine therapy, nappies, play equipment, transport, wheelchair equipment and insurance and home modifications.

Child's health-related quality of life

In adjusted analyses, both the EQ-5D-5L and HUI3 scores of HRQoL tended to be higher in survivors who were allocated to DRIFT than in those who were allocated to standard care (*Table 34*). However, the CIs around the adjusted mean differences in EQ-5D-5L score (0.06, 95% CI –0.11 to 0.22) and HUI3 score (0.13, 95% CI –0.09 to 0.35) included zero. Imputing a score of zero for the six children recruited in Bristol who were known to have died (two in the DRIFT arm, four in the standard care arm) led to similar conclusions: adjusted mean differences in EQ-5D-5L score (0.10, 95% CI –0.08 to 0.29) and HUI3 score (0.13, 95% CI –0.07 to 0.33). In contrast, EQ-5D-5L VAS scores were higher in participants allocated to receive standard care (adjusted mean difference –11.18, 95% CI –23.66 to 1.32).

Associations between cognitive status, quality of life and total NHS costs

We observed the expected positive correlation between better cognitive status at 10-year follow-up and higher QoL scores as measured by the HUI3 and EQ-5D-5L (*Figures 25–28*). The QoL scores ranged across almost the entire spectrum of scores on the HUI3 and EQ-5D-5L (see *Table 34*). The EQ-5D-5L VAS was an exception, showing a high ceiling effect (many scores of 1) and no evident correlation with cognitive status. We also observed the expected negative correlation between cognitive status at 10-year follow-up and NHS inpatient care costs over the child's lifetime.

TABLE 32 Child's ambulatory health and social care use at school age, by trial arm

	Trial arm							
	DRIFT			Standard care			Difference in mean costs, £ (95% CI)	
Visits (in the last)	*n*	Mean (SD) units	Mean (SD) costs, £	*n*	Mean (SD) units	Mean (SD) costs, £	Unadjusted	Adjusted[a]
ED (12 months)	23	0.43 (0.59)		17	0.41 (0.62)			
OP (12 months)	23	2.83 (3.05)		17	2.82 (2.96)			
Total ambulatory hospital care (12 months)	23		381 (382)	17		378 (360)	3 (–238 to 244)	2 (–264 to 267)
GP (6 months)	23	1.70 (1.94)		17	1.12 (1.32)			
Optician (6 months)	23	0.57 (0.66)		17	1.23 (1.03)			
Dentist (6 months)	23	1.13 (0.69)		17	1.29 (0.77)			
Paediatrician (6 months)	23	0.43 (0.66)		17	0.76 (0.75)			
Speech therapist (6 months)	23	1.52 (4.25)		17	1.76 (4.85)			
Hearing specialist (6 months)	23	0.30 (0.56)		17	0.35 (0.61)			
School nurse (6 months)	23	0.65 (2.50)		17	2.24 (4.94)			
Individual counselling (6 months)	23	0.26 (1.25)		17	0.71 (2.91)			
Social worker (6 months)	23	0.48 (1.08)		17	0.94 (1.98)			
Other[b] ambulatory care (6 months)	23	14.26 (28.87)		17	1.94 (3.47)			
Total ambulatory community care (6 months)	23		569 (626)	17		718 (736)	–148 (–586 to 287)	–108 (–596 to 380)

a Adjusted by gender, IVH grade and birthweight.
b Health visitor, child development centre, family counselling, home help, day centre, after-school club. The large mean in the DRIFT group is due to a small number of individuals reporting a high number of after-school club visits and home help.

TABLE 33 Family income and child's educational needs

Income and educational needs	*n*	DRIFT	*n*	Standard care	Unadjusted (95% CI)	Adjusted[a] (95% CI)
First parent/carer employed	23	16 out of 23 (70%)	17	9 out of 17 (53%)	OR 2.03 (0.55 to 7.47)	OR 1.55 (0.36 to 6.75)
Partner employed[b]	11	7 out of 11 (64%)	11	9 out of 11 (82%)	OR 0.39 (0.05 to 2.77)	OR 0.60 (0.04 to 9.31)
Estimated weekly household income from employment	23	£580	17	£599	–£19 (–£287 to £325)	£0 (–£325 to £325)
Main source of income is benefits	23	4 out of 23 (17%)	17	7 out of 17 (41%)	OR 0.30 (0.07 to 1.28)	OR 0.23 (0.04 to 1.22)
Child has SEN statement	23	13 out of 23 (57%)	17	10 out of 17 (59%)	OR 0.91 (0.26 to 3.24)	OR 0.41 (0.08 to 2.10)
Child attends special unit/school	23	5 out of 23 (22%)	17	7 out of 17 (41%)	OR 0.40 (0.10 to 1.58)	OR 0.13 (0.02 to 0.82)
Estimated annual cost of schooling	23	£11,659	17	£14,164	–£2505 (–£7067 to £2056)	–£5321 (–£9772 to –£870)
Estimated annual cost of schooling SA1	23	£23,943	17	£43,249	–£19,305 (–£43,755 to £5144)	–£35,122 (–£58,546 to –£11,699)

a Adjusted by gender, IVH grade and birthweight.
b Twenty-two parents/carers provided information on their partner.

TABLE 34 Child's HRQoL

	Trial arm					
	DRIFT		Standard care		Difference in mean score (95% CI)	
HRQoL	*n*	Mean (SD) score	*n*	Mean (SD) score	Unadjusted	Adjusted[a]
EQ-5D-5L index score	23	0.70 (0.25)	18	0.70 (0.31)	0.00 (–0.18 to 0.17)	0.06 (–0.11 to 0.22)
EQ-5D-5L index score (SA2)[b]	25	0.64 (0.31)	22	0.57 (0.39)	0.07 (–0.14 to 0.27)	0.10 (–0.08 to 0.29)
HUI-3 index score	23	0.53 (0.34)	18	0.49 (0.43)	0.04 (–0.21 to 0.28)	0.13 (–0.09 to 0.35)
HUI-3 index score (SA2)[b]	25	0.49 (0.35)	22	0.40 (0.43)	0.08 (–0.14 to 0.31)	0.13 (–0.07 to 0.33)
EQ-5D-5L VAS score	23	78.6 (19.9)	18	90.8 (15.1)	–12.22 (–23.66 to –0.79)	–11.18 (–23.66 to 1.32)

a Adjusted by gender, IVH grade and birthweight.
b Including a score of zero for participants known to have died.

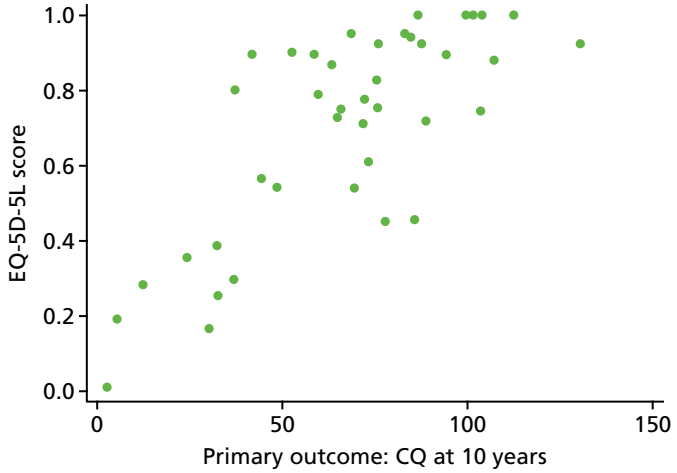


FIGURE 25 Association between cognition and EQ-5D-5L.

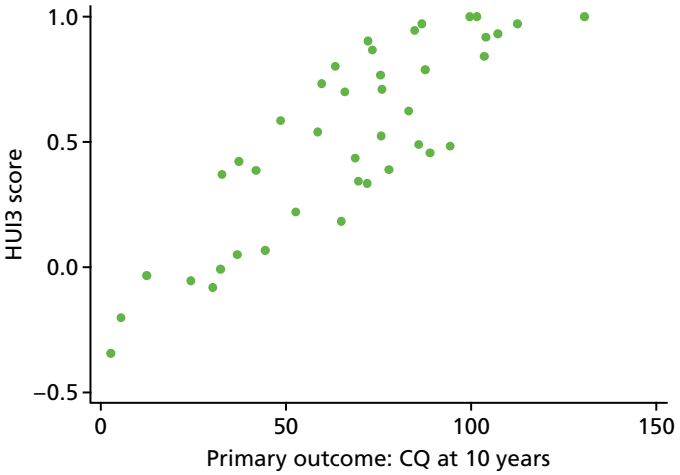


FIGURE 26 Association between cognition and HUI3.

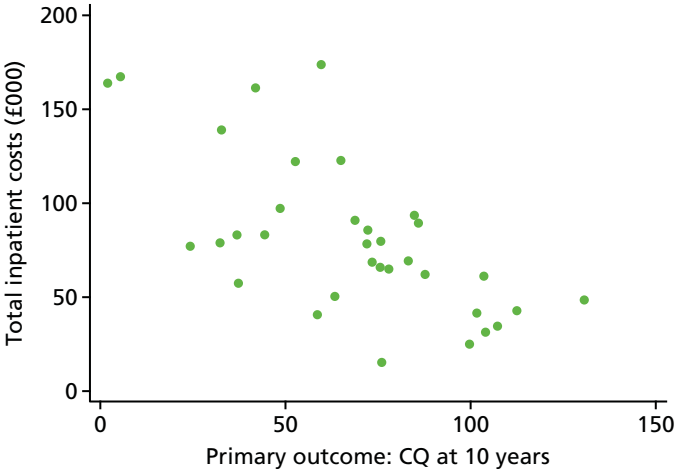


FIGURE 27 Association between cognition and inpatient costs.

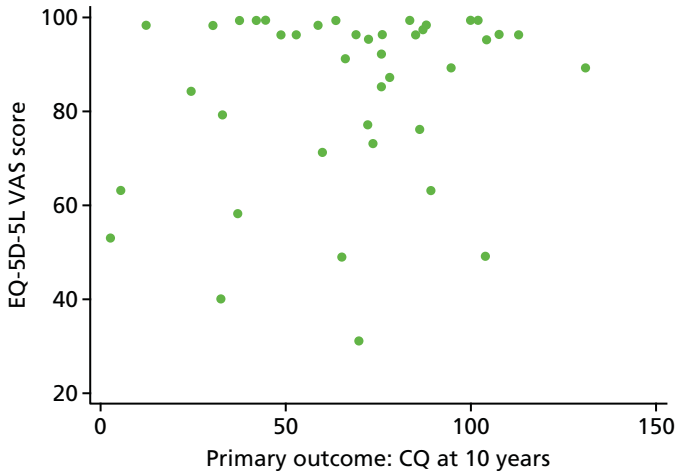


FIGURE 28 Association between cognition and EQ-5D-5L VAS.

Results of the decision analytical model

The parameters used to estimate the model are presented in *Table 35*. The results of this exploratory analysis (*Table 36*) indicate that DRIFT has the potential to be a cost-effective intervention at current NICE thresholds. At 18 years, the additional benefit (8.96 QALYs vs. 8.33 QALYs) resulting primarily from the lower mortality in the DRIFT arm justifies the higher NHS and social service costs (£112,341 vs. £102,611). The ICER (£15,621) is below the NICE thresholds of £20,000 to £30,000 per QALY and the INMB (£2711)

TABLE 35 Parameters of the decision analytical model derived from the DRIFT data

Parameter	Value	Distribution	Parameters[a]	
Cost of DRIFT (£)	1513	Log-normal	7.32	0.04
Cost of NICU stay (DRIFT) (£)	57,882	Log-normal	10.96	0.10
Cost of NICU stay (SC) (£)	52,839	Log-normal	10.87	0.12
Postnatal inpatient cost 0_2 (DRIFT) (£)	8876	Log-normal	9.06	0.24
Postnatal inpatient cost 0_2 (SC) (£)	5790	Log-normal	8.61	0.34
Postnatal inpatient cost 2_10 (DRIFT) (£)	18,209	Log-normal	9.76	0.32
Postnatal inpatient cost 2_10 (SC) (£)	18,245	Log-normal	9.76	0.31
Ambulatory care cost 12 months (DRIFT) (£)	381	Log-normal	5.92	0.21
Ambulatory care cost 12 months (SC) (£)	378	Log-normal	5.91	0.23
Community care cost 6 months (DRIFT) (£)	569	Log-normal	6.32	0.23
Community care cost 6 months (SC) (£)	718	Log-normal	6.55	0.24
EQ5D-5L decrement (DRIFT)	0.3031	Log-normal	–1.21	0.17
EQ5D-5L decrement (SC)	0.2983	Log-normal	–1.24	0.24
Mortality rate 0_2 (DRIFT)	0.0370	Beta	1.00	26.00
Mortality rate 0_2 (SC)	0.1111	Beta	3.00	24.00
Mortality rate 2 to 10 (DRIFT)	0.0417	Beta	1.00	23.00
Mortality rate 2 to 10 (SC)	0.0500	Beta	1.00	19.00

SC, standard care.
a For lognormal distributions, these are the mean of ln (x) and the SD of ln (x). For beta distributions, these are the alpha and beta parameters.

TABLE 36 Cost-effectiveness results from the decision analytical model

Decision analytical model	Trial arm					
	DRIFT		Standard care			
	Cost[a] (£)	QALYs[a]	Cost[a] (£)	QALYs[a]	ICER (£)	INMB (£)[b] (95% CI)[c]
Primary analysis	112,341	8.9566	102,611	8.3338	15,621	2711 (–52,397 to 58,445)
MSA1	94,677	5.7128	85,172	5.3452	25,856	–2152 (–39,195 to 36,015)
MSA2	98,833	9.3089	116,571	7.9178	DRIFT dominant	45,558 (–6289 to 97,203)
MSA3	219,182	8.9566	232,409	8.3338	DRIFT dominant	25,684 (–43,690 to 97,313)
MSA4	112,341	6.7946	102,611	5.8534	10,338	9095 (–57,309 to 78,277)

a Deterministic analysis, discounted values.
b At a willingness to pay threshold of £20,000 per QALY.
c Probabilistic analysis, discounted values.
MSA1, estimating costs and outcomes at age 10 years rather than age 18 years; MSA2, using costs and utility scores adjusted for baseline covariate, gender, IVH grade and birthweight; MSA3, including educational costs; MSA4, using HUI3 rather than EQ-5D-5L utility scores.

is positive. However, there is a high degree of uncertainty about the effect of DRIFT on both the costs and outcomes of care, as indicated by the very wide CI surrounding the INMB estimate and the flat cost-effectiveness acceptability curve, which shows that DRIFT has close to 0.5 probability of being cost-effective (*Figure 29*). However, in scenarios in which education costs (see *Table 36*, MSA3) are included or using costs and utility scores adjusting for gender, IVH grade and birthweight (see *Table 36*, MSA2), DRIFT has the potential to both save money and improve outcomes for children. In both scenarios, there is a high probability (> 0.75) that DRIFT is cost-effective at NICE thresholds.

Discussion

Main findings

We found no evidence that the DRIFT intervention either increased or decreased the cost of the initial neonatal stay. The initial cost of DRIFT was offset, to an extent, by the fact that fewer procedures to insert reservoirs were carried out. However, the costs of these procedures were small in comparison with the overall costs of NICU care. There was high between-patient variation in NICU length of stay and, therefore, in the costs of NICU care. We observed differences between trial arms in the distribution of NICU stay.

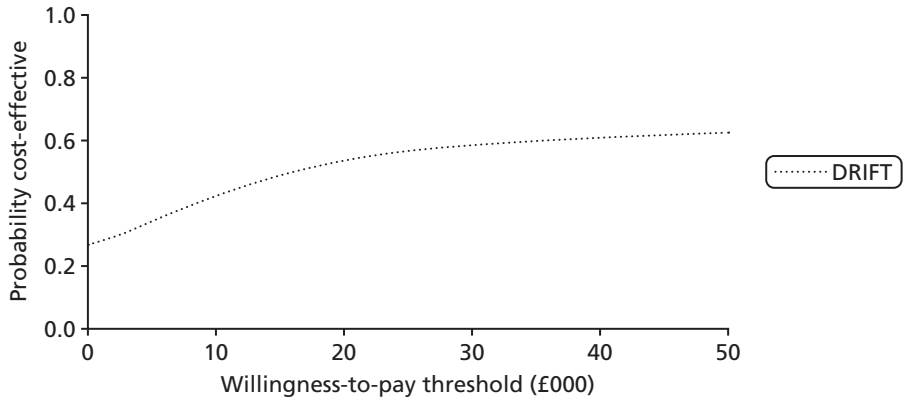


FIGURE 29 Cost-effectiveness acceptability curve.

Participants who received DRIFT spent fewer days in the Bristol NICU but more days in other NICUs after transfer from Bristol. It is possible that the lower number of reservoir procedures after DRIFT allowed these babies to be transferred back to the outlying NICU more quickly without decreasing total NICU stay. In adjusted analyses, we estimated that DRIFT might reduce the costs of neonatal care by approximately £3000, but the wide CI means that we cannot rule out the possibility that it increases costs.

In a subgroup (34 out of 54; 63%) of patients who survived to the 10-year follow-up, who lived in England and whose parents consented to data linkage, unadjusted analyses suggested that those who received the DRIFT intervention spent slightly more days in hospital after the initial neonatal stay. There was no strong evidence that the DRIFT intervention increased or decreased the costs of postnatal inpatient care or total NHS inpatient costs. The finding in unadjusted analyses that total costs of inpatient care were approximately £11,800 higher in participants who received DRIFT was very sensitive to adjustment for birthweight and gender. After adjustment, costs in the DRIFT arm were estimated to be approximately £21,000 lower. These findings should be interpreted cautiously owing to the sensitivity of the estimates, the wide CIs and the selective nature of the subgroup. It is worth noting that babies in the DRIFT arm included in the analysis of postnatal costs were less mature (on average 300 g smaller and born 1 week earlier) than those in the standard care arm. Therefore, other comorbidities due to immaturity at birth may have affected readmissions to hospital during childhood.

A subgroup of parents (41 out of 54; 76%) reported on QoL, recent health and social care, employment and educational needs. Children had wide-ranging health and social care needs at the 10-year follow-up; however, there was no evidence of economically important differences between participants who received DRIFT and those who received standard care. There was no evidence that DRIFT had an impact on household income, although there was a non-significant trend for a lower proportion of households of participants who received DRIFT to report having benefits as their main source of income. A lower proportion of parents of children in the DRIFT arm reported that their children attended a special school or unit. Because of the high costs and likelihood of ongoing need of special education, the potential savings (approximately £2550 and £5300 per annum in unadjusted and adjusted analysis, respectively) are likely to be very influential in the economic case for DRIFT; it is likely to become more cost-effective over the future lifetime of survivors. There was no evidence of differences in QoL scores (HUI3 and EQ-5D-5L) at 10-year follow-up between participants who received DRIFT and those who received standard care. The EQ-5D-5L VAS was higher (better) among participants who received standard care; however, this measure showed no correlation with the primary outcome of cognitive function. Previous work[65] has demonstrated that the EQ-5D-5L is more strongly correlated than the EQ-5D-5L VAS with disease-specific measures of QoL. It is also unclear how parents of children with lifelong cognitive and other health problems would interpret the VAS end point labels of 'best/worst health you can imagine'. For these reasons, we believe that the EQ-5D-5L and HUI3 are likely to be more valid indicators of patient outcomes in this population.

Exploratory analysis using a simulation model to interpolate and extrapolate costs and outcomes to age 18 years indicated that DRIFT has the potential to be cost-effective at conventional willingness-to-pay thresholds used by the NHS. In some scenarios (including education costs and using adjusted estimates of costs and outcomes), DRIFT was very likely to be cost-effective and might both save money and improve outcomes.

Strengths and weaknesses

We have provided the first evidence on the cost-effectiveness of the DRIFT procedure for infants with PHVD. The evidence is drawn from a RCT study design, which minimises the risk of selection bias. The long-term follow-up obtained in a high proportion of participants enabled us to compare the ongoing educational, health and social care needs of children. Our economic analysis is limited by a small sample size and in being restricted to one NICU. In microcosting the DRIFT intervention, we used unit costs derived from one hospital. Unit costs for the DRIFT intervention will vary somewhat from hospital to hospital; however, given the high cost of NICU and subsequent care, this variation is not likely to be pivotal in determining the cost-effectiveness of DRIFT.

Owing to the huge variance in health-care needs and costs of participants, a much larger trial would be needed to provide greater certainty about whether DRIFT increases or decreases NHS costs and other costs in the long term. Small RCTs are vulnerable to imbalances in important baseline covariates.[66] This was the case for birthweight and gender in this trial, which resulted in some of our findings being very sensitive to adjustment for these covariates. In the case of outcomes which were measured in only a subset of survivors, such as the cost of postnatal care, it is unclear whether or not the unadjusted analysis, which ignores baseline imbalances, or adjusted analysis, which may exacerbate attrition bias, will be more accurate. The data are not missing completely at random. We do not have postnatal cost data for the participants who died or for those who did not live in England, one of whom, in the standard care arm, was known to require permanent residential medical care.

Another challenge was the retrospective nature of several elements of the evaluation. As the original RCT had not included an economic evaluation, we were reliant on hospital notes being available and accurately recording relevant details. We were able to extract details on the Bristol NICU and any transfer NICU stay for most participants and supplemented this with HES data. However, no information on the Bristol NICU stay was available for some patients (7 out of 54; 13%) and they were excluded from the analysis; in other patients, some details (e.g. HDU days) had to be imputed. Other important elements of the economic evaluation, for example QoL between birth and school age, were also unavailable. Owing to the incomplete information, we chose to report a cost–consequence study based on available cases as our primary analysis rather than a more conventional cost-effectiveness analysis.

Our study illustrates the strengths and weaknesses of HES data for economic evaluation. Acquiring HES data involved a lengthy process of approval. However, without HES, we could not have built up such a detailed picture of inpatient care during the first 10 years of life. Parent recall would probably have been inaccurate over such a long period of time and it would have been impractical to identify all hospitals where care had been provided in order to extract data from notes. However, we had consent for data linkage from only 42 out of 54 (78%) parents and, as a number of families lived in Wales or emigrated, HES data were not comprehensive. Recent work[67] concluded that HES birth data offer a high-quality data set that captures the majority of English hospital births. We found that linked HES data had high, but not perfect, sensitivity for identifying the Bristol NICU episode and excellent specificity. It is unclear whether the absence of linked HES birth data is a failure of linkage (e.g. inaccurate record of date of birth and NHS number) or of absence of the episode from the HES data set. However, the imperfect sensitivity of data linkage suggests that our estimate of postnatal inpatient costs is likely to be conservative.

Accurate measurement of health and social care use is difficult in a group of children with such wide-ranging needs. We asked parents to quantify their child's use of a wide variety of professionals and services. However, the relatively large number of free-text comments, often without enough details to estimate costs, indicated that our estimate of ambulatory health and social care might be conservative. Conversely, as some of this care may have been provided as part of special schooling, there is also a risk of double counting care that is bundled in with educational costs.

The decision analysis model is exploratory. The cost-effectiveness findings have a high degree of uncertainty and include a number of strong underlying assumptions in interpolating and extrapolating some costs and outcomes between birth, 10-year follow-up and age 18 years. A larger trial with more frequent follow-up with parents and linkage to hospital and primary care records could reduce this uncertainty. Longitudinal cohort studies estimating the impact of neurological impairment acquired at birth on long-term outcomes and costs of care are also needed to determine the cost-effectiveness of interventions such as DRIFT over the lifetime of the patient.

Comparison with other studies

The HUI3 scores we found in the DRIFT trial lie between the HUI3 scores reported among approximately 11-year-old children with moderate (mean HUI3 score of 0.744) and severe (mean HUI3 score of 0.364) neurodevelopmental impairment following extremely preterm birth in the EPICure economic outcomes

study.[61] The EPICure economic outcomes study[61] also found that mean health and social care costs in the previous year were £1223 (at 2006/7 costs) among children with neurodevelopmental disability, increasing to £8241 if educational costs were included. This, and our work, highlights the importance of taking a broad perspective including cost of education in judging the cost-effectiveness of interventions in this group of patients.

Our study can capture only a subset of the total economic consequences of childhood disability. In their review of the literature, Stabile and Allin[68] identify the broader spill-over effects on parental employment, health and relationships as well as the long-term consequences for the child's employment and need for ongoing care and welfare benefits. They argue that many expensive interventions to prevent or reduce childhood disabilities might well be justified if all these economic consequences can be taken into account. The confidential inquiry into premature deaths of people with learning disabilities has highlighted the high proportion of adults requiring residential or nursing home care and needing 24-hour care.[19] However, as other authors[69] have noted, estimating the likely long-term return on investment of neonatal interventions is severely hampered by the methodological variability in studies investigating the economic costs to families who care for a child with disabilities.

Implications

The National Institute of Health and Care Excellence currently recommends that DRIFT should be used only in the context of research. If DRIFT were implemented more widely in specialist NICUs, it would be relatively straightforward to provide training to new teams in the NHS. One-to-one nurse staffing would need to be factored in to provide the frequent monitoring required for DRIFT. We found that the intervention itself has a relatively moderate financial cost, but the economic consequences of the procedure are potentially very large, particularly if it reduces the need for special education in the long term. Our findings suggest that DRIFT may increase cognitive status and reduce the need for special education at school age; however, more evidence is required to determine whether or not the intervention is cost-effective. A larger multicentre RCT with prospective economic evaluation would provide more definitive evidence. Economic modelling could initially extrapolate the results of such a RCT; long-term follow-up would also be required.

Chapter 5 Discussion and conclusions

Summary of findings

Infants who received DRIFT continued to demonstrate better cognitive ability at 10-year follow-up and effects were significant when taking into account birthweight, IVH grade and gender (two of which were unbalanced at baseline). The proportion of children who survived without severe cognitive disability was significantly higher with DRIFT in both adjusted and unadjusted analyses. Children who received the DRIFT intervention were nine times more likely to survive without severe cognitive disability and the NNT for DRIFT to prevent one death or one case of severe cognitive disability was only three.

However, there were no apparent differences in the secondary outcomes: parent-reported visual impairment, sensorimotor disability or emotional/behavioural difficulties.

High-resolution structural brain MRI at 10 years showed no evidence of residual damage associated with insertion of the DRIFT irrigation catheters. A larger proportion of the standard treatment arm required ventricular reservoirs and more residual frontal tracts associated with reservoirs were seen in the standard treatment arm. There was no difference in ongoing neurosurgical problems between the treatment arms at age 10 years.

Economic evaluation

Our findings suggest that DRIFT may increase cognitive status and reduce the need for special education at school age; however, more evidence is required to determine whether or not the intervention is cost-effective.

Although the DRIFT intervention has a relatively moderate financial cost, the economic consequences of the procedure are potentially very large, particularly if it continues to reduce the need for special education in the long term.

Because of the high costs and likelihood of ongoing need of special education, the potential savings are likely to be very influential in the economic case for DRIFT; it is likely to become more cost-effective over the future lifetime of survivors. The exploratory decision analysis model to age 18 years indicated that DRIFT has the potential to be cost-effective at conventional willingness-to-pay thresholds used by the NHS. In some scenarios, DRIFT may save money and improve outcomes owing to the possible reduction in the need for special education.

Strengths and limitations

The main strength of this study is the long-term follow-up to school age, which is more likely to give a valid conclusion for future function and cognitive ability. Long-term follow-up of this nature is challenging in neonatal clinical trials as families move around; it requires active buy-in from both parents and children and a significant time commitment from families. In the case of conditions such as PHVD, a significant proportion of survivors of which have severe neurodisabilities, the logistics and commitment around returning for long-term assessments understandably become even more challenging.

Precise cognitive assessment in children with a very wide range of abilities is a significant challenge for trials. The approach to cognitive assessment in this study achieved a CQ in children of all abilities. Inclusion of educational outcome as a pragmatic outcome was also important as this gives some idea of the likely gains going forward into an independent adulthood.

Parent and family involvement and the organisation of the NHS ensured a very high follow-up rate at school age in the UK. Only two patients had an unknown survival status at 10 years and, for this reason, best- and worst-case scenarios were also explored.

The main limitation is the size of the trial and corresponding precision of the results. Given that this intervention was innovative and the condition rare, the sample size, naturally, was conservative. The intervention is invasive and, for safety insurance, the trial had stringent stopping criteria, which limited the achieved sample size. Although the safety reasons for this are understandable and justifiable, it unfortunately resulted in a smaller sample size than was required to give 80% power for the primary outcome (CQ). Reassuringly, the binary outcome survival without severe cognitive disability was significantly better in the DRIFT group both in adjusted and unadjusted analyses.

The majority of infants in the DRIFT trial and at 10-year follow-up were UK-based babies managed in Bristol. The infants managed in Bristol were referred and transferred from all over the UK (33 hospitals in total). Therefore, the UK trial cohort was probably representative of the wider UK population of preterm infants with IVH and PHVD. The DRIFT trial also demonstrated that it is feasible to train highly motivated teams in other centres to deliver the intervention within the context of a RCT. However, there remains uncertainty around the extension of DRIFT trial results to new potential centres in the UK, which will need to be resolved with further work on standardisation and training within the governance structure of a UK-wide trial.

Owing to the modest size of this trial, the small numbers did result in imbalances in important characteristics at baseline and, consequently, at 10-year follow-up. Infants in the DRIFT group were significantly smaller, less mature and more likely to be male and had more severe-grade haemorrhages. Therefore, prespecified adjustments were made in the primary, secondary and health economic analyses using these covariates consistent with the earlier work.

Interpretation of results

DRIFT is the first intervention for PHVD in preterm infants to demonstrate benefit in a RCT. DRIFT also demonstrates the proof of principle that washing away the debris of IVH in a controlled way reduces secondary brain injury.

Rates of severe cognitive (learning) disability in the standard treatment arm were 52%, similar to previously reported work[5] in younger children with PHVD. The proportion of children with severe cognitive disability was reduced with DRIFT to 21%. The CIPOLD study[19] in England highlighted the complex lifelong health and social care needs of individuals with significant learning disabilities; two-thirds of individuals lived in residential care homes, the majority with 24-hour paid nursing care. Children who received DRIFT were also more likely to attend mainstream schools. The reduction in severe cognitive disability seen with this intervention is likely to translate into the ability to lead more independent lives into adulthood.

A multicentre RCT comparing two treatment thresholds for ventricular reservoir insertion after PHVD, the Early vs. Late Ventricular Intervention Study (ELVIS; ISRCTN43171322), has recently ended recruitment. Short-term outcomes should be published shortly but long-term neurological outcomes will not be known before 2019.

Newer interventions are also being tested. A feasibility and safety study of endoscopic ventricular lavage showed fewer complications and need for VP shunts in larger (> 1000 g) preterm babies than standard treatment in historical controls.[29] However, as yet, neither short-term outcomes (complications and need for shunts) nor long-term effects on neurological function have been reported in any controlled trial.

Implications for health care

The school-age follow-up of the DRIFT trial strengthens the evidence of benefit found at the 2-year follow-up and adds further evidence of safety of the intervention. We conclude that DRIFT improves cognitive function when taking into account birthweight, grade of IVH and gender. The cost of the intervention is moderate and the reduction in the need for special education at school age is likely to translate into a cost-effective intervention over a lifetime.

In the years since the DRIFT trial, neonatal intensive care organisation in the NHS has evolved into a highly organised hub-and-spoke service mapped to large regions. Individual units in these operational delivery networks (ODNs) are connected by dedicated neonatal transport services. Cases with PHVD could feasibly be managed in this networked system by a small number of highly specialised units with neonatal neurocritical care and neurosurgical expertise. Although the equipment and consumables used in DRIFT are widely available, it needs to be acknowledged that DRIFT is potentially a high-risk intervention and specialised units will need intensive neurosurgical, medical and nursing training in delivering DRIFT safely.

Future research implications

The demographics of the population of infants with PHVD has evolved since the DRIFT trial, with significantly better survival in the 23- and 24-week gestation categories. A larger proportion of infants with PHVD is now extremely immature, small and clinically unstable. Further refinements in DRIFT may need to be studied in this very immature group of patients.

DRIFT has an effect on cognition but does not appear to improve motor function. The most likely explanation is that simple irrigation, although effective at reducing secondary neurotoxicity and damage, is not sufficient to promote tissue regeneration in critical motor pathways after significant parenchymal infarction. There is scope to supplement DRIFT with novel interventions to promote brain tissue repair in the future.

The role of any NHS implementation of DRIFT, ideally in a few specialised tertiary centres, delivered through the existing neonatal ODNs, will need to be studied prospectively in a multicentre trial. As well as measures of cognition and functional measures, the data from the 10-year outcomes indicate that any future studies should continue to collect data on vision, motor skills and education given the trends seen in the secondary outcomes, which the study was not powered to address.

Acknowledgements

Contributions of authors

Dr Karen Luyt (Consultant Senior Lecturer in Neonatal Medicine, Neonatal Neurology, School of Clinical Sciences, University of Bristol and chief investigator) made a substantial contribution to the conception and design of the study. She drafted the research protocol and supervised all data acquisition and analysis. She contributed to the interpretation of all of the trial outcomes and health economic analysis. She contributed to the assessment panel for the MRI scans. She made a major contribution to drafting and revision of the report for intellectual content.

Dr Sally Jary (Senior Research Associate in Child Development, Neonatal Neurology, School of Clinical Sciences, University of Bristol and co-investigator) made a contribution to the conception and design of the study. She contributed to the design of the cognitive assessment protocol. She performed all the sensorimotor assessments in all the children in Bristol, Glasgow and Bergen and acquired and analysed their data. She co-ordinated and supervised the assessment days in Bristol, including obtaining informed consent and guiding parental completion of all the questionnaires. She contributed to analysis and interpretation of the cognitive and other outcome data. She co-ordinated data collection of the health resource use data. She assisted with drafting and revision of the report for intellectual content.

Dr Charlotte Lea (NIHR Academic Clinical Fellow, Neonatal Neurology, School of Clinical Sciences, University of Bristol) contributed to the analysis and interpretation of the cognitive and other outcome data. She co-ordinated data collection of the health resource use data. She made a major contribution to finalising and cleaning the study database for final analysis. She acquired and cleaned all the educational outcome data and a substantial amount of the neonatal stay hospital data and contributed to the analysis and interpretation. She assisted with the health data acquisition from NHS Digital. She analysed the structural MRI data. She made a major contribution to drafting and revision of the report for intellectual content.

Miss Grace J Young (Research Associate in Medical Statistics, School of Social and Community Medicine, University of Bristol) performed all of the statistical analysis of the trial outcomes. She contributed to the interpretation of the analysis and wrote the statistical analysis report. She made a major contribution to drafting and revision of the report for intellectual content.

Dr David Odd (Consultant in Neonatal Medicine, North Bristol NHS Trust and co-investigator) made a contribution to the conception and design of the study. He contributed to the statistical analysis of the trial outcomes and the health economic analysis. He contributed to the interpretation of all the analysis. He made a major contribution to drafting and revision of the report for intellectual content.

Dr Helen Miller (Child Psychologist, Neonatal Neurology, School of Clinical Sciences, University of Bristol) made a contribution to the conception and design of the study. She contributed to the design of the cognitive assessment protocol. She assessed all the children in Bristol, Glasgow and Bergen and acquired and analysed their cognitive data. She contributed to the interpretation of the cognitive data. She assisted with revision of the report for intellectual content.

Dr Grazyna Kmita (Child Psychologist, Faculty of Psychology, University of Warsaw, Warsaw, Poland and co-investigator) made a contribution to the conception and design of the study. She contributed to the design of the cognitive assessment protocol. She assessed all the children in Poland and acquired and analysed their cognitive data. She contributed to the interpretation of the cognitive data. She assisted with revision of the report for intellectual content.

Miss Cathy Williams (Reader in Paediatric Ophthalmology, School of Social and Community Medicine, University of Bristol and co-investigator) made a contribution to the conception and design of the study. She designed the visual assessment protocol. She examined the children in Bristol and was responsible for acquisition and cleaning of the visual data. She contributed to the visual data analysis and interpretation. She contributed to drafting and revision of the report for intellectual content.

Dr Peter S Blair [Reader in Medical Statistics, School of Social and Community Medicine, University of Bristol and Bristol Randomised Trials Collaboration (BRTC) and co-investigator] made a contribution to the conception and design of the study. He supervised all of the statistical analysis of the trial outcomes. He contributed to interpretation of the analysis and to the statistical analysis report. He contributed to drafting and revision of the report for intellectual content.

Miss Aída Moure Fernández (Research Associate in Health Economics, School of Social and Community Medicine, University of Bristol) performed the hospital stay data acquisition and contributed to the health economic analysis. She contributed to drafting and revision of the report for intellectual content.

Professor William Hollingworth (Professor of Health Economics, School of Social and Community Medicine, University of Bristol and BRTC and co-investigator) made a contribution to the conception and design of the study. He supervised all, and performed a substantial part of, the health economic analysis. He was responsible for the data acquisition from NHS Digital. He performed the interpretation of the health economic analysis and drafted the health economic report. He contributed to drafting and revision of the report for intellectual content.

Dr Michelle Morgan (Child Psychologist, Community Children's Health Partnership, Bristol and co-investigator) made a contribution to the conception and design of the study. She contributed to the design of the cognitive assessment protocol and supervised the cognitive data acquisition. She assisted with revision of the report for intellectual content.

Dr Adam Smith-Collins (NIHR Academic Clinical Lecturer, Neonatal Neurology, School of Clinical Sciences, University of Bristol University of Bristol) made a contribution to the conception and design of the study. He designed the MRI protocol and acquired and analysed the MRI data. He contributed to the assessment panel for the MRI scans. He assisted with revision of the report for intellectual content.

Dr N Jade Thai [Senior Research Fellow, Clinical Research and Imaging Centre Bristol (CRICBristol), School of Clinical Sciences, University of Bristol] made a contribution to the conception and design of the study. She designed the MRI protocol and acquired and analysed the MRI data. She assisted with drafting the report and revision of the report for intellectual content.

Mr Steven Walker-Cox (Parent Representative, Neonatal Neurology, School of Clinical Sciences, University of Bristol University of Bristol and co-investigator) made a contribution to the conception and design of the study. He reviewed all the study literature and advised around the assessment days in Bristol. He represented the families and children on the trial steering group. He assisted with revision of the report for intellectual content.

Mr Kristian Aquilina (Consultant in Paediatric Neurosurgery, Great Ormond Street Hospital, London and co-investigator) made a contribution to the conception and design of the study. He contributed to the assessment panel for the MRI scans. He assisted with revision of the report for intellectual content.

Mr Ian Pople (Consultant in Paediatric Neurosurgery, University Hospitals Bristol NHS Trust and co-investigator) made a contribution to the conception and design of the study. He contributed to the assessment panel for the MRI scans. He supervised the costing of all the neurosurgical procedures. He assisted with revision of the report for intellectual content.

Professor Andrew Whitelaw (Professor of Neonatal Medicine, Neonatal Neurology, School of Clinical Sciences, University of Bristol, co-investigator of the follow-up study and chief investigator of the original DRIFT trial) made a substantial contribution to the conception and design of the study. He contributed to the interpretation of all of the trial outcomes and health economic analysis. He contributed to the assessment panel for the MRI scans. He assisted with drafting the report at all stages and revised it critically for intellectual content.

All authors approved the final submitted version of the report.

Trial Steering Committee

Chairperson: Professor Neil Marlow.

Independent members: Dr Divyen Shah, Dr Topun Austin, Professor Stavros Petrou, Dr Catrin Tudur-Smith and Professor Andrew Wilkinson.

Parent representative: Steven Walker-Cox.

Other contributions

First, we need to thank the families and the children for their support and commitment to the DRIFT trial and for giving their time so generously to advance the knowledge and management of PHVD. This study would not have been possible without their significant contribution.

We would also like to thank all the clinicians from the various referring hospitals for assisting so generously with data provision.

We thank Dr Kathleen O'Reilly (Glasgow) and the paediatricians in Bergen for providing their support in tracing participants and data.

CRICBristol hosted all the assessments in Bristol, performed the MRI and generously hosted the showcase day for the families and children. Aileen Wilson (Lead Research Radiographer) made a significant contribution to scanning the children. Penny Warnes performed the visual assessments on the children in Bristol.

The BRTC provided expertise at every stage of the study.

Ethics review

Ethics approval was granted by the NHS Health Research Authority (14/SW/1078).

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review. Hospital Episode statistics data were provided by NHS Digital under agreement DARS-NIC-30560-W4V1T-v0.5.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. Larroche JC. Post-haemorrhagic hydrocephalus in infancy. Anatomical study. *Biol Neonate* 1972;**20**:287–99. https://doi.org/10.1159/000240472
2. Horbar JD, Carpenter JH, Badger GJ, Kenny MJ, Soll RF, Morrow KA, Buzas JS. Mortality and neonatal morbidity among infants 501 to 1500 grams from 2000 to 2009. *Pediatrics* 2012;**129**:1019–26. https://doi.org/10.1542/peds.2011-3028
3. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller A-B, Narwal R, *et al.* National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012;**379**:2162–72. https://doi.org/10.1016/S0140-6736(12)60820-4
4. Costeloe KL, Hennessy EM, Haider S, Stacey F, Marlow N, Draper ES. Short term outcomes after extreme preterm birth in England: comparison of two birth cohorts in 1995 and 2006 (the EPICure studies). *BMJ* 2012;**345**:e7976. https://doi.org/10.1136/bmj.e7976
5. Adams-Chapman I, Hansen NI, Stoll BJ, Higgins R, Network NR. Neurodevelopmental outcome of extremely low birth weight infants with posthemorrhagic hydrocephalus requiring shunt insertion. *Pediatrics* 2008;**121**:E1167–77. https://doi.org/10.1542/peds.2007-0423
6. Limbrick DD Jr, Mathur A, Johnston JM, Munro R, Sagar J, Inder T, *et al.* Neurosurgical treatment of progressive posthemorrhagic ventricular dilation in preterm infants: a 10-year single-institution study. Clinical article. *J Neurosurg Pediatr* 2010;**6**:224–30. https://doi.org/10.3171/2010.5.PEDS1010
7. van de Bor M, den Ouden L. School performance in adolescents with and without periventricular-intraventricular hemorrhage in the neonatal period. *Semin Perinatol* 2004;**28**:295–303. https://doi.org/10.1053/j.semperi.2004.08.007
8. Kaiser AM, Whitelaw AGL. Cerebrospinal-fluid pressure during post hemorrhagic ventricular dilatation in newborn-infants. *Arch Dis Child* 1985;**60**:920–4. https://doi.org/10.1136/adc.60.10.920
9. Savman K, Nilsson UA, Blennow M, Kjellmer I, Whitelaw A. Non-protein-bound iron is elevated in cerebrospinal fluid from preterm infants with posthemorrhagic ventricular dilatation. *Pediatr Res* 2001;**49**:208–12. https://doi.org/10.1203/00006450-200102000-00013
10. Savman K, Blennow M, Hagberg H, Tarkowski E, Thoresen M, Whitelaw A. Cytokine response in cerebrospinal fluid from preterm infants with posthaemorrhagic ventricular dilatation. *Acta Paediatr* 2002;**91**:1357–63. https://doi.org/10.1111/j.1651-2227.2002.tb02834.x
11. O'Connor A, Wilson CM, Fielder AR. Ophthalmological problems associated with preterm birth. *Eye* 2007;**21**:1254–60. https://doi.org/10.1038/sj.eye.6702838
12. Ramenghi LA, Ricci D, Mercuri E, Groppo M, De Carli A, Ometto A, *et al.* Visual performance and brain structures in the developing brain of pre-term infants. *Early Hum Dev* 2010;**86**(Suppl. 1):73–5. https://doi.org/10.1016/j.earlhumdev.2010.01.010
13. Cioni G, Bertuccelli B, Boldrini A, Canapicchi R, Fazzi B, Guzzetta A, Mercuri E. Correlation between visual function, neurodevelopmental outcome, and magnetic resonance imaging findings in infants with periventricular leucomalacia. *Arch Dis Child Fetal Neonatal Ed* 2000;**82**:F134–40. https://doi.org/10.1136/fn.82.2.F134
14. Rahi JS, Cable N, British Childhood Visual Impairment Study Group. Severe visual impairment and blindness in children in the UK. *Lancet* 2003;**362**:1359–65. https://doi.org/10.1016/S0140-6736(03)14631-4

15. Williams C, Northstone K, Sabates R, Feinstein L, Emond A, Dutton GN. Visual perceptual difficulties and under-achievement at school in a large community-based sample of children. *PLOS ONE* 2011;**6**:e14772. https://doi.org/10.1371/journal.pone.0014772

16. Fazzi E, Signorini SG, Bova SM, La Piana R, Ondei P, Bertone C, *et al.* Spectrum of visual disorders in children with cerebral visual impairment. *J Child Neurol* 2007;**22**:294–301. https://doi.org/10.1177/0883073807022003080 1

17. Centers for Disease Control and Prevention. Economic costs associated with mental retardation, cerebral palsy, hearing loss, and vision impairment. *MMWR Morb Mortal Wkly Rep* 2004;**53**:57–9.

18. Petrou S, Khan K. Economic costs associated with moderate and late preterm birth: primary and secondary evidence. *Semin Fetal Neonatal Med* 2012;**17**:170–8. https://doi.org/10.1016/j.siny.2012.02.001

19. Heslop P, Blair PS, Fleming P, Hoghton M, Marriott A, Russ L. The confidential inquiry into premature deaths of people with intellectual disabilities in the UK: a population-based study. *Lancet* 2014;**383**:889–95. https://doi.org/10.1016/S0140-6736(13)62026-7

20. Pople IK, Bayston R, Hayward RD. Infection of cerebrospinal fluid shunts in infants: a study of etiological factors. *J Neurosurg* 1992;**77**:29–36. https://doi.org/10.3171/jns.1992.77.1.0029

21. Drake JM, Kestle JRW, Milner R, Cinalli G, Boop F, Piatt J Jr, *et al.* Randomized trial of cerebrospinal fluid shunt valve design in pediatric hydrocephalus. *Neurosurgery* 1998;**43**:294–303. https://doi.org/10.1097/00006123-199808000-00068

22. Whitelaw A. Randomized trial of early tapping in neonatal posthaemorrhagic ventricular dilatation. *Arch Dis Child* 1990;**65**:3–10. https://doi.org/10.1136/adc.65.1_Spec_No.3

23. Johnson A, Wincott E, Grant A, Elbourne D. Randomized trial of early tapping in neonatal posthemorrhagic ventricular dilatation: results at 30 months. *Arch Dis Child* 1994;**71**:F147. https://doi.org/10.1136/fn.71.2.F147

24. Whitelaw A. Repeated lumbar or ventricular punctures in newborns with intraventricular hemorrhage. *Cochrane Database Syst Rev* 2001;**1**:CD000216. https://doi.org/10.1002/14651858.CD000216

25. Kennedy C, Campbell M, Elbourne D, Hope P, Johnson A, Darroch A, *et al.* International randomised controlled trial of acetazolamide and furosemide in posthaemorrhagic ventricular dilatation in infancy. *Lancet* 1998;**352**:433–40. https://doi.org/10.1016/S0140-6736(05)60352-2

26. Brouwer AJ, Groenendoal F, van den Hoogen A, Verboon-Maciolek M, Hanlo P, Rademaker KJ, de Vries LS. Incidence of infections of ventricular reservoirs in the treatment of post-haemorrhagic ventricular dilatation: a retrospective study (1992–2003). *Arch Dis Child Fetal Neonatal Ed* 2007;**92**:F41–3. https://doi.org/10.1136/adc.2006.096339

27. de Vries LS, Liem KD, van Dijk K, Smit BJ, Sie L, Rademaker KJ, *et al.* Early versus late treatment of posthaemorrhagic ventricular dilatation: results of a retrospective study from five neonatal intensive care units in the Netherlands. *Acta Paediatr* 2002;**91**:212–17. https://doi.org/10.1111/j.1651-2227.2002.tb01697.x

28. de Vries LS, Groenendaal F, Liem KD, Heep A, Brouwer AJ, van't Verlaat E, *et al.* Treatment thresholds for intervention in posthaemorrhagic ventricular dilation: a randomised controlled trial [published online ahead of print February 10 2018]. *Arch Dis Child Fetal Neonatal Ed* 2018.

29. Schulz M, Bührer C, Pohl-Schickinger A, Haberl H, Thomale UW. Neuroendoscopic lavage for the treatment of intraventricular hemorrhage and hydrocephalus in neonates. *J Neurosurg Pediatr* 2014;**13**:626–35. https://doi.org/10.3171/2014.2.PEDS13397

30. Whitelaw A, Pople I, Cherian S, Evans D, Thoresen M. Phase 1 trial of prevention of hydrocephalus after intraventricular hemorrhage in newborn infants by drainage, irrigation, and fibrinolytic therapy. *Pediatrics* 2003;**111**:759–65. https://doi.org/10.1542/peds.111.4.759

31. Whitelaw A, Evans D, Carter M, Thoresen M, Wroblewska J, Mandera M, *et al.* Randomized clinical trial of prevention of hydrocephalus after intraventricular hemorrhage in preterm infants: brain-washing versus tapping fluid. *Pediatrics* 2007;**119**:E1071–8. https://doi.org/10.1542/peds.2006-2841

32. Whitelaw A, Jary S, Kmita G, Wroblewska J, Musialik-Swietlinska E, Mandera M, *et al.* Randomized trial of drainage, irrigation and fibrinolytic therapy for premature infants with posthemorrhagic ventricular dilatation: developmental outcome at 2 years. *Pediatrics* 2010;**125**:E852–8. https://doi.org/10.1542/peds.2009-1960

33. Whitelaw A, Odd DE. Intraventricular streptokinase after intraventricular hemorrhage in newborn infants. *Cochrane Database Syst Rev* 2007;**4**:CD000498. https://doi.org/10.1002/14651858.CD000498.pub2

34. Elliot CD, Smith P. *British Ability Scales (Third Edition)*. London: GL Assessment Ltd; 2012.

35. Bayley N. *Scales of Infant & Toddler Development*. 3rd edn. San Antonio, TX: Harcourt Assessment Inc.; 2006.

36. Houliston MJ, Taguri AH, Dutton GN, Hajivassiliou C, Young DG. Evidence of cognitive visual problems in children with hydrocephalus: a structured clinical history-taking strategy. *Dev Med Child Neurol* 1999;**41**:298–306. https://doi.org/10.1017/S0012162299000675

37. Barnett A, Henderson SE, Sugden DA. *Movement Assessment Battery for Children (examiners manual)*. 2nd edn. London: Pearson Assessment; 2007.

38. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry* 1997;**38**:581–6. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

39. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol* 1997;**39**:214–23. https://doi.org/10.1111/j.1469-8749.1997.tb07414.x

40. Parekh SA, Field DJ, Johnson S, Juszczak E. Accounting for deaths in neonatal trials: is there a correct approach? *Arch Dis Child Fetal Neonatal Ed* 2015;**100**:F193–7. https://doi.org/10.1136/archdischild-2014-306730

41. Biering K, Hjollund NG, Frydenberg M. Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clin Epidemiol* 2015;**7**:91–106. https://doi.org/10.2147/CLEP.S72247

42. Client Service Receipt Inventory – Children's Version. URL: www.dirum.org/assets/downloads/634462397526425351-CSRI%20-%20Childrens%20Version.pdf (accessed 27 May 2018).

43. National Institute for Health and Care Excellence. *Drainage, Irrigation and Fibrinolytic Therapy (DRIFT) for Post-haemorrhagic Hydrocephalus In Preterm Infants*. URL: www.nice.org.uk/guidance/ipg412 (accessed 27 May 2018).

44. Coast J. Is economic evaluation in touch with society's health values? *BMJ* 2004;**329**:1233–6. https://doi.org/10.1136/bmj.329.7476.1233

45. Roostan E, Johnson C, Lamburne S. A nursing approach to caring for a baby and their family undergoing ventricular lavage for post haemorrhagic ventricular dilatation. *Infant* 2012;**8**:4–8.

46. Department of Health and Social Care (DHSC). *NHS Reference Costs 2014 to 2015*. London: DHSC; 2015.

47. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.

48. British Medical Association, Royal Pharmaceutical Society. *British National Formulary*. London: BMJ Group, and Pharmaceutical Press; 2016.

49. NHS Digital. *Hospital Episode Statistics (HES)*. URL: http://digital.nhs.uk/hesdata (accessed 11 October 2016).

50. The College of Optometrists. *Look After Your Eyes – What Does an Eye Examination Cost*. URL: https://lookafteryoureyes.org/seeing-clearly/eye-examination-cost/ (accessed 27 May 2018).

51. Romeo R, Knapp M, Scott S. Economic cost of severe antisocial behaviour in children – and who pays it. *Br J Psychiatry* 2006;**188**:547–53. https://doi.org/10.1192/bjp.bp.104.007625

52. Education Funding Agency. *Funding Allocation Pack: 2014 to 2015 Academic Year. A Guide For Mainstream New Provision Academies Opening Between 1 September 2014 and 31 March 2015*. 2014. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/324024/Funding_allocation_pack_2014_to_2015_academic_year_new_spon_provision_250614.pdf (accessed 23 May 2018).

53. Department for Education. *Official Statistics. LA and School Expenditure: 2014 to 2015 Financial Year*. 2016. URL: www.gov.uk/government/statistics/schools-education-and-childrens-services-spending-2014-to-2015 (accessed 11 September 2016).

54. Office for National Statistics. *Annual Survey of Hours and Earnings: 2015 Provisional Results*. 2015. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/annualsurveyofhoursandearnings/2015provisionalresults/relateddata (accessed 23 May 2018).

55. Clifford J, Theobold C. *Summary of Findings: Extension of the 2011 Cost Comparison Methodology to a Wider Sample. National Association of Independent Schools and Non-Maintained Special Schools*. 2012. URL: www.nasschools.org.uk/wp-content/uploads/sites/9/2014/08/NASS-Cost-Comparison-Report-October-2012.pdf (accessed 23 May 2018).

56. *National Tariff Payment System 2014/15*. London: Monitor & NHS England; 2013. URL: www.gov.uk/government/publications/national-tariff-payment-system-2014-to-2015 (accessed 23 May 2018).

57. National Institute for Health and Clinical Excellence (NICE). *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.

58. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, *et al.* Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002;**40**:113–28. https://doi.org/10.1097/00005650-200202000-00006

59. Health Utilities Inc. *Multi-Attribute Health Status Classification System: Health Utilities Index Mark 3 (HUI3) [updated December 2014]*. URL: www.healthutilities.com/hui3.htm (accessed 22 August 2016).

60. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;**20**:1727–36. https://doi.org/10.1007/s11136-011-9903-x

61. Petrou S, Johnson S, Wolke D, Marlow N. The association between neurodevelopmental disability and economic outcomes during mid-childhood. *Child Care Health Dev* 2013;**39**:345–57. https://doi.org/10.1111/j.1365-2214.2012.01368.x

62. Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. *Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England*. Research paper 16/01. London: Office of Health Economics; 2016.

63. Mihaylova B, Briggs A, O'Hagan A, Thompson SG. Review of statistical methods for analysing healthcare resources and costs. *Health Econ* 2011;**20**:897–916. https://doi.org/10.1002/hec.1653

64. O'Brien BJ, Briggs AH. Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods. *Stat Methods Med Res* 2002;**11**:455–68. https://doi.org/10.1191/0962280202sm304ra

65. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Qual Life Res* 2014;**23**:977–89. https://doi.org/10.1007/s11136-013-0537-z

66. Roberts C, Torgerson DJ. Understanding controlled trials: baseline imbalance in randomised controlled trials. *BMJ* 1999;**319**:185. https://doi.org/10.1136/bmj.319.7203.185

67. Ghosh RE, Ashworth DC, Hansell AL, Garwood K, Elliott P, Toledano MB. Routinely collected English birth data sets: comparisons and recommendations for reproductive epidemiology. *Arch Dis Child Fetal Neonatal Ed* 2016;**2016**:F451–7.

68. Stabile M, Allin S. The economic costs of childhood disability. *Future Child* 2012;**22**:65–96. https://doi.org/10.1353/foc.2012.0008

69. Anderson D, Dumont S, Jacobs P, Azzaria L. The personal costs of caring for a child with a disability: a review of the literature. *Public Health Rep* 2007;**122**:3–16. https://doi.org/10.1177/003335490712200102

Appendix 1 British Ability Scales version three, school age

DRIFT

Patient ID

Recognition of Designs

Raw Score _____

Item Set

- ○ 1-14
- ○ 1-19
- ○ 5-19
- ○ 1-24
- ○ 5-24
- ○ 10-24
- ○ 5-29
- ○ 10-29
- ○ 15-29
- ○ 10-34
- ○ 15-34
- ○ 20-34
- ○ 15-38
- ○ 20-38

Word Definitions

Raw Score _____

Item Set

○ 1-13 ○ 1-17 ○ 5-17 ○ 1-21 ○ 5-21 ○ 5-27 ○ 14-27 ○ 14-33 ○ 22-33
○ 14-38 ○ 22-38

Pattern Construction (Std/Alt)

Please select which pattern construction item set was used

Std ○
Alt ○

Raw Score _____

Item Set

○ 1-12 ○ 1-19 ○ 13-19 ○ 13-25 ○ 13-28 ○ 20-28 ○ 20-32 ○ 26-32 ○ 20-35

○ 26-35

Raw Score _____

Item Set

○ 1-12 ○ 1-19 ○ 13-19 ○ 13-25 ○ 13-28 ○ 20-28 ○ 20-32 ○ 26-32 ○ 20-35 ○ 26-35 ○ 26-38

Matrices

Raw Score _____

Item Set

○ 1-18 ○ 1-26 ○ 1-33 ○ 19-33 ○ 19-42 ○ 27-42 ○ 27-51 ○ 37-51

Verbal Similarities

Raw Score _____

Item Set

○ 1-12 ○ 1-20 ○ 8-20 ○ 8-27 ○ 15-27 ○ 15-32 ○ 21-32 ○ 15-37 ○ 21-37

Quantitative Reasoning

Raw Score _____

Item Set

○ 1-16 ○ 1-24 ○ 8-24 ○ 1-30 ○ 8-30 ○ 17-30 ○ 8-40 ○ 17-40 ○ 25-40

Word Reading A/B

please select which pattern construction item set

A ○
B ○

Raw Score _____

Item Set

○ 1-90 ○ 21-90 ○ 41-90 ○ 51-90

Raw Score _____

Item Set

○ 1-90 ○ 21-90 ○ 41-90 ○ 51-90

Recognition of Designs

Recognition of Designs - Ability Score _____

Recognition of Designs - T-score _____

Recognition of Designs - Percentile _____

Recognition of Designs - Difference from mean core T-score _____

Recognition of Designs - Significance _____
(p=0.05)

Recognition of Designs - Significance (y/n)Yes No

Recognition of Designs - Frequency (%) _____

Recognition of Designs - Age Equivalent _____
(yy:mm)

Word Definitions

Word Definitions - Ability Score _____

Word Definitions - T-score _____

Word Definitions - Percentile _____

Word Definitions - Difference from mean core T-score _____

Word Definitions - Significance _____ (p=0.05)

Word Definitions - Significance (y/n)

Word Definitions - Frequency (%) _____

Word Definitions - Age Equivalent _____ (yy:mm)

Pattern Constuction (Std/Alt)

Pattern Construction - Ability Score _____

Pattern Construction - T-score _____

Pattern Construction - Percentile _____

Pattern Construction - Difference from mean core T-score _____

Pattern Construction - Significance _____ (p=0.05)

Pattern Construction - Significance (y/n)

Pattern Construction - Frequency (%) _____

Pattern Construction - Age Equivalent _____ (yy:mm)

Matrices

Matrices - Ability Score _____

Matrices - T-score _____

Matrices - Percentile _____

Matrices - Difference from mean core T-score _____

Matrices - Significance _____ (p=0.05)

Matrices - Significance (y/n)

Matrices - Frequency (%) _____

Matrices - Age Equivalent _____ (yy:mm)

Verbal Similarities

Verbal Similarities - Ability Score _____

Verbal Similarities - T-score _____

Verbal Similarities - Percentile _____

Verbal Similarities - Difference from mean core T-score _____

Verbal Similarities - Significance _____ (p=0.05)

Verbal Similarities - Significance (y/n)

Verbal Similarities - Frequency (%) _____

Verbal Similarities - Age Equivalent _____ (yy:mm)

Quantitative Reasoning

Quantitative Reasoning - Ability Score _____

Quantitative Reasoning - T-score _____

Quantitative Reasoning - Percentile _____

Quantitative Reasoning - Difference from mean core T-score _____

Quantitative Reasoning - Significance _____ (p=0.05)

Quantitative Reasoning - Significance (y/n)

Quantitative Reasoning - Frequency (%) _____

Quantitative Reasoning - Age Equivalent _____ (yy:mm)

Mean T-Score

Mean T-score _____

Word Reading

Word Reading - Ability Score _____

Word Reading - Standard Score _____

Word Reading - Percentile _____

Word Reading - Difference from GCA score _____

Word Reading - Significance _____ (p=0.05)

Word Reading - Significance (y/n)

Word Reading - Frequency (%) _____

Word Reading - Age Equivalent _____ (yy:mm)

GCA and Cluster Standard Scores

Verbal Ability - Word Definitions _____

Verbal Ability - Verbal Similarities _____

Non-Verbal Reasoning Ability - Matrices _____

Non-Verbal Reasoning Ability - Quantitative Reasoning _____

Spatial Ability - Recognition of Designs _____

Spatial Ability - Pattern Construction _____

Sum of T-scores

Verbal Ability T-Scores _____

Non-Verbal Reasoning Ability T-score _____

Spatial Ability- T-score _____

Sum of T-Scores _____

SNC - Sum of T-scores _____

Standard Scores, Confidence Intervals and Percentiles

Verbal Ability - Standard Score _____

Verbal Ability - Confidence Interval _____

(-95%)

Verbal Ability - Percentile _____

Non-Verbal Reasoning Ability - Confidence Interval _____

(-95%)

Non-Verbal Reasoning Ability - Standard Score _____

Non-Verbal Reasoning Ability - Percentile _____

Spatial Ability - Standard Score _____

Spatial Ability - Confidence Interval _____

(-95%)

Spatial Ability - Percentile _____

GCA - Standard Score _____

GCA - Confidence Interval _____

(-95%)

GCA - Percentile _____

SNC - Standard Score _____

SNC - Confidence Interval _____

(95%)

SNC - Percentile _____

Appendix 2 British Ability Scales version three, early years

DRIFT

Page 1 of 6

Patient ID _____

Verbal Comprehension

Raw Score _____

Item Set

○ '1-12 ○ '1-23 ○ '6-23 ○ '6-30 ○ '13-30 ○ '13-40 ○ '13-23/31-40

Picture Similarities

Raw Score _____

Item Set

○ '1-23 ○ '1-28 ○ '12-28 ○ '12-35 ○ '19-35

Naming Vocabulary

Raw Score _____

Item Set

○ 1-17 ○ 1-24 ○ 11-24 ○ 1-31 ○ 11-31 ○ 18-31 ○ 11-36 ○ 18-36

Pattern Construction (Std/Alt)

Please select which pattern construction item set was used

Std ○
Alt ○

Raw Score_____

Item Set

○ 1-12 ○ 1-19 ○ 13-19 ○ 13-25 ○ 13-28 ○ 20-28 ○ 20-32 ○ 26-32 ○ 20-35 ○ 26-35

Raw Score_____

Item Set

○ 1-12 ○ 1-19 ○ 13-19 ○ 13-25 ○ 13-28 ○ 20-28 ○ 20-32 ○ 26-32 ○ 20-35 ○ 26-35 ○ 26-38

Matrices

Raw Score_____

Item Set

○ 1-18 ○ 1-26 ○ 1-33 ○ 19-33 ○ 19-42 ○ 27-42 ○ 27-51 ○ 37-51

Copying

Raw Score_____

Item Set

○ 1-12 ○ 1-15 ○ 5-15 ○ 1-20 ○ 5-20 ○ 11-20

Verbal Comprehension

Verbal Comprehension - Ability Score _____

Verbal Comprehension - T-score _____

Verbal Comprehension - Percentile _____

Verbal Comprehension - Difference from mean core T-score _____

Verbal Comprehension - Significance _____ (p=0.05)

Verbal Comprehension - Significance (y/n)Yes No

Verbal Comprehension - Frequency (%) _____

Verbal Comprehension - Age Equivalent (yy:mm) _____

Picture Similarities

Picture Similarities - Significance _____

(p=0.05)

Picture Similarities - Significance (y/n)Yes No

Picture Similarities - Frequency (%) _____

Picture Similarities - Age Equivalent _____

(yy:mm)

Naming Vocabulary

Naming Vocabulary - Ability Score _____

Naming Vocabulary - T-score _____

Naming Vocabulary - Percentile _____

Naming Vocabulary - Difference from mean core T-score _____

Naming Vocabulary - Significance _____

(p=0.05)

Picture Similarities - Ability Score _____

Picture Similarities - T-score _____

Picture Similarities - Percentile _____

Picture Similarities - Difference from mean core T-score _____

Naming Vocabulary - Significance (y/n)

Naming Vocabulary - Frequency (%) _____

Naming Vocabulary - Age Equivalent _____

(yy:mm)

Pattern Construction (Std/Alt)

Pattern Construction - Ability Score _____

Pattern Construction - T-score _____

Pattern Construction - Percentile _____

Pattern Construction - Difference from mean core T-score _____

Pattern Construction - Significance _____ (p=0.05)

Pattern Construction - Significance (y/n)

Matrices

Matrices - Ability Score _____

Matrices - T-score _____

Matrices - Percentile _____

Matrices - Difference from mean core T-score _____

Matrices - Significance _____ (p=0.05)

Matrices - Significance (y/n)

Matrices - Frequency (%) _____

Matrices - Age Equivalent _____ (yy:mm)

Copying

Copying - Ability Score _____

Copying - T-score _____

Copying - Percentile _____

Copying - Difference from mean core T-score _____

Copying - Significance _____

(p=0.05)

Copying - Significance (y/n)

Copying - Frequency (%) _____

Copying - Age Equivalent _____

(yy:mm)

Mean T-Score

Mean T-score _____

GCA and Cluster Standard Scores

Verbal Ability - Verbal Comprehension _____

Non-verbal reasoning ability - Picture Similarities _____

Verbal Ability- Naming Vocabulary _____

Spatial Ability - Pattern Construction _____

Non-Verbal Reasoning Ability- Matrices _____

Spatial Ability - Copying _____

Sum of T-scores

Verbal Ability T-Scores _____

Non-Verbal Reasoning Ability T-score _____

Spatial Ability T-score _____

Sum of T-Scores _____

SNC - Sum of T-scores _____

Standard Scores, Confidence Intervals and Percentiles

Verbal Ability - Standard Score _____

Verbal Ability - Confidence Interval _____

(-95%)

Verbal Ability - Percentile _____

Non-Verbal Reasoning Ability - Confidence Interval _____

(-95%)

Non-Verbal Reasoning Ability - Standard Score _____

Non-Verbal Reasoning Ability - Percentile _____

Spatial Ability - Standard Score _____

Spatial Ability - Confidence Interval _____

(-95%)

Spatial Ability - Percentile _____

GCA - Standard Score _____

GCA - Confidence Interval _____

(-95%)

GCA - Percentile _____

SNC - Standard Score _____

SNC - Confidence Interval _____

(95%)

SNC - Percentile _____

Appendix 3 Bayley Scales of Infant and Toddler Development version three

Confidential

Patient ID _____

Age of child being tested: _____

(YY:MM)

Gestation week at birth _____

COGNITIVE SCORE

Total Raw Score _____

Scaled Score _____

Composite Score _____

Percentile Rank _____

Confidence Interval (95%) _____

Developmental Age Equivalent (months) _____

Receptive Communication

Total Raw Score _____

Scaled Score _____

Developmental Age Equivalent (months) _____

Expressive Communication

Total Raw Score _____

Scaled Score _____

Developmental Age Equivalent (months) _____

LANGUAGE SCORE

Total Raw Score _____

Scaled Score _____

Composite Score _____

Percentile Rank _____

Confidence Interval (95%) _____

Developmental Age Equivalent (months) _____

Fine Motor

_____

Total Raw Score

Scaled Score _____

Developmental Age Equivalent (months) _____

Gross Motor

_____

Total Raw Score

Scaled Score _____

Developmental Age Equivalent (months) _____

MOTOR SCORE

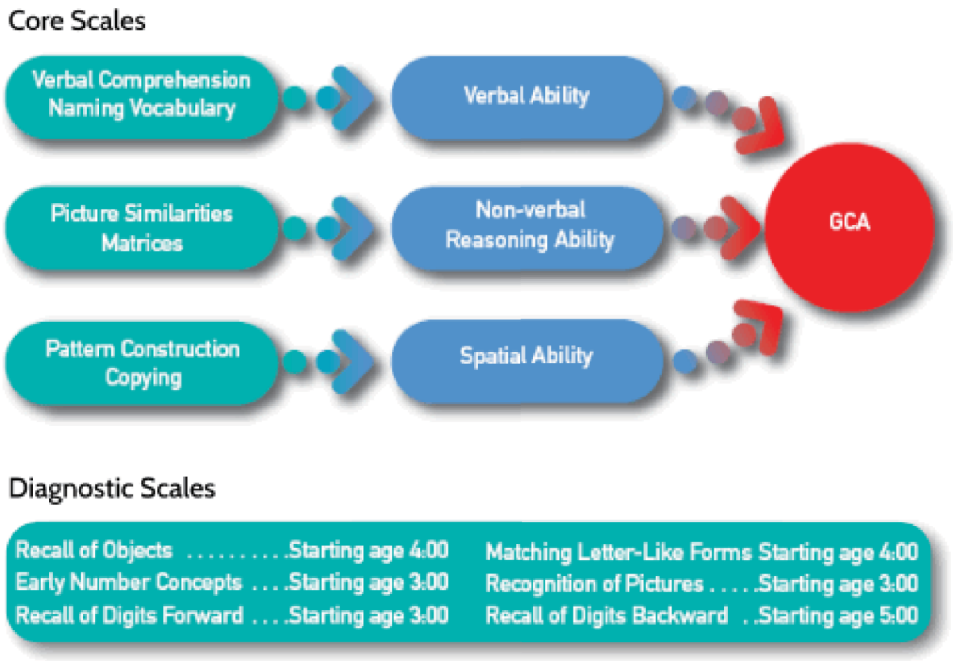
Total Raw Score	_____
Scaled Score	_____
Composite Score	_____
Percentile Rank	_____
Confidence Interval (95%)	_____
Developmental Age Equivalent (months)	_____

The Early Years Battery

The Early Years Battery covers all the scales that are appropriate for use with pre-school children, and those in the early school years – normed for the age range 3:00 to 7:11 years, but generally used up to 5:11 years. The scales use appealing artwork and manipulable objects to assess reasoning, perception and memory, together with understanding of basic quantitative concepts.

The Early Years Battery has six Core Scales that yield three Cluster Scores (based on two scales each) and the GCA, complemented by six Diagnostic Scales. Of the Diagnostic Scales, three (Matching Letter-Like Forms, Recall of Objects and Recall of Digits Backward) are normed from 4:00 or 5:00 years onwards as they are too demanding for most children under these ages. The structure of the Early Years Battery has been simplified, to reflect that of the School Age Battery with just one level.

The composition of the Early Years Battery is illustrated below:

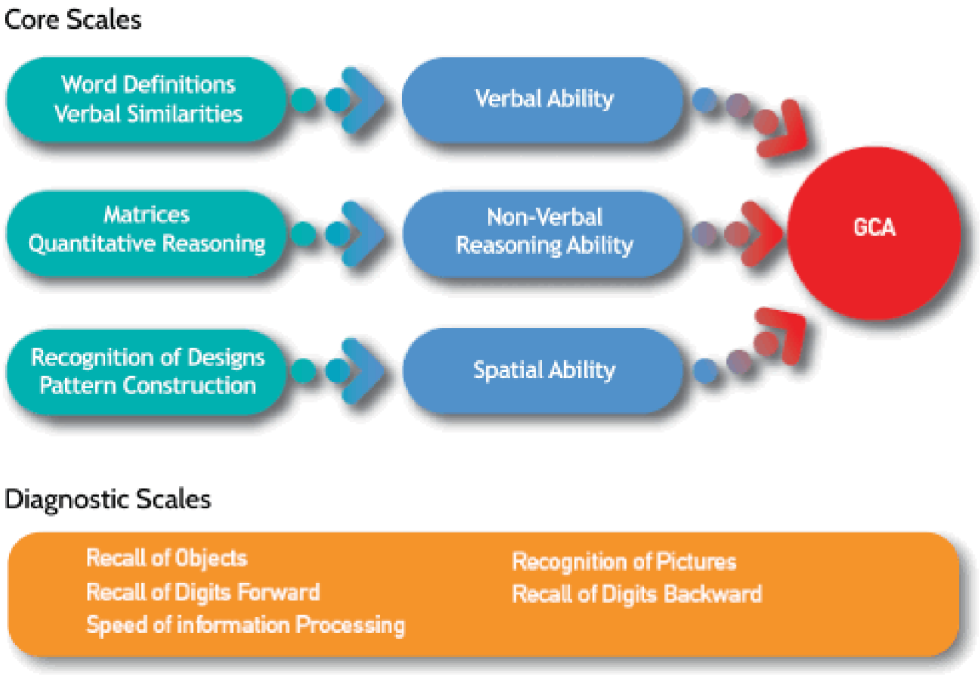


The School Age Battery

The School Age Battery covers the Cognitive Scales that are designed for use with school-age children. These scales assess reasoning, perception, processing speed and memory using verbal, numerical and figural materials. The latter includes abstract shapes, pictures and three dimensional materials.

The School Age Battery was normed from age 5:00 to 17:11 years and is generally used from 6:00 years as a few scales are quite demanding for younger children. Core Scales combine to produce three Cluster Scores and, in combination, the GCA. There are five Diagnostic Scales in this battery.

The composition of the School Age Battery is illustrated below:



Appendix 4 Visual questionnaire

Confidential

Visual Assessment

Please complete the survey below.

Thank you!

Section 1 - Questions relating to use of vision in day-to-day situations (Houliston et all; DMCN, 1999, 41: 298-306)

Recognition questions

	1 - Never	2	3	4	5 - Always
1.1 Parent recognition	○	○	○	○	○
1.2 Family recognition	○	○	○	○	○
1.3 Friend recognition	○	○	○	○	○
1.4 Photograph recognition	○	○	○	○	○
1.5 Self-photograph recognition	○	○	○	○	○
1.6 Shape recognition	○	○	○	○	○
1.7 Object recognition	○	○	○	○	○

Colour Questions

	1-Never	2	3	4	5-Always
1.8 Colour naming	○	○	○	○	○
1.9 Colour matching	○	○	○	○	○

Orientation

	1-Never	2	3	4	5-Always
1.10 Finding way in home	○	○	○	○	○
1.11 Asking way in home	○	○	○	○	○
1.12 Losing objects at home	○	○	○	○	○
1.13 Finding way in new places	○	○	○	○	○
1.14 Asking way in new places	○	○	○	○	○

Visually Guided Movement Questions

	1-Never	2	3	4	5-Always
1.15 Reaching and grasping objects	○	○	○	○	○

Confidential

1.16 Distinguishing step from line	○	○	○	○	○
1.17 Misjudging doorways/corridors	○	○	○	○	○
1.18 Can manage steps and kerbs?	○	○	○	○	○

Movement Questions

	1-Never	2	3	4	5-Always
1.19 Seeing moving objects	○	○	○	○	○
1.20 Finding objects while moving	○	○	○	○	○

Complex Scenes Questions

	1-Never	2	3	4	5-Always
1.21 Finding objects in complex pictures/scenes	○	○	○	○	○
1.22 Finds objects pointed out in distance?	○	○	○	○	○

Neglect Questions

	1-Never	2	3	4	5-Always
1.23 Eating food from part of plate	○	○	○	○	○

Section 2 - Acuity and Contrast testing

2.1a Binocular: Distance acuity single optotype (Binocular) _____

2.1b Test used _____

2.1c Glasses? ○ Yes ○ No

2.2a Binocular: Distance acuity _____

2.2b Test used _____

2.2c Glasses? ○ Yes ○ No

2.3a Binocular: Near acuity single optotype (Binocular) _____

2.3b Test used _____

2.3c Glasses? ○ Yes ○ No

2.4a Binocular: Near acuity crowded optotype (Binocular) _____

Confidential

2.4b Test used _____

2.4c Glasses?
- ○ Yes
- ○ No

2.5a Binocular: Comment on binocular acuity tests

_____

2.6a Right: Distance acuity single optotype _____

2.6b Test used _____

2.6c Glasses?
- ○ Yes
- ○ No

2.7a Right: Distance acuity crowded optotype _____

2.7b Test used _____

2.7c Glasses?
- ○ Yes
- ○ No

2.8a Left: Distance acuity single optotype _____

2.8b Test used _____

2.8c Glasses?
- ○ Yes
- ○ No

2.9a Left: Distance acuity crowded optotype _____

2.9b Test used _____

2.9c Glasses?
- ○ Yes
- ○ No

2.10a Right: Near acuity single optotype _____

2.10b Test used _____

2.10c Glasses?
- ○ Yes
- ○ No

2.11a Right: Near acuity crowded optotype _____

2.11b Test used _____

2.11c Glasses?
- ○ Yes
- ○ No

2.12a Left: Near acuity single optotype _____

2.12b Test used _____

2.12c Glasses?
- ○ Yes
- ○ No

2.13a Left: Near acuity crowded optotype _____

2.13b Test used _____

2.13c Glasses?
- ○ Yes
- ○ No

Confidential

2.14a Right/Left: Comment on monocular acuity testing

2.15a Binocular: Best near VA achieved at preferred distance

2.16a Binocular: Best comfortable font size

2.17a Binocular: Contrast sensitivity

_____
(%)

2.17b Test used _____

2.17c Glasses?
- ◯ Yes
- ◯ No

2.18a Right: Contrast sensitivity

_____
(%)

2.18b Test used _____

2.18c Glasses?
- ◯ Yes
- ◯ No

2.19a Left: Contrast sensitivity

_____
(%)

2.19b Test used _____

2.19c Glasses?
- ◯ Yes
- ◯ No

2.20a Binocular: Comment on contrast sensitivity

Section 3. Visual Fields, Cover Tests and Eye Movements

3.1 Binocular Visual field
- ◯ Normal
- ◯ Variably reduced
- ◯ Definitely reduced

Confidential

3.2 Comment on Binocular visual field

_____
(Describe field loss)

3.3 Right visual field
- ◯ Normal
- ◯ Variably reduced
- ◯ Definitely reduced

3.4 Left visual field
- ◯ Normal
- ◯ Variably reduced
- ◯ Definitely reduced

3.5 Comment on monocular field test

_____
(Describe field loss)

3.6 Cover test unaided at 33 cm
- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.7 Cover test unaided at distance
- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.8 Alternate Cover test unaided at 33 cm
- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.9 Alternate Cover test unaided at 6 m
- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.10 Cover test with glasses at 33 cm
- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

Confidential

3.11 Cover test with glasses at 6 m

- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.12 Alternate Cover test with glasses at 33 cm

- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.13 . Alternate Cover test with glasses at 6 m

- ◯ Normal
- ◯ Right ET
- ◯ Left ET
- ◯ Alt ET
- ◯ Right XT
- ◯ Left XT
- ◯ Alt XT
- ◯ Other

3.14 Comment on Cover/Alt cover testing

_____

3.15 Nystagmus

- ◯ None
- ◯ In PP
- ◯ At extremes of gaze

3.16 Describe Nystagmus

_____

3.17 Ocular motility

- ◯ Full
- ◯ Abnormal

3.18 Describe eye movements

_____

3.19 Smooth pursuits-Horizontal slow

◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

3.20 Smooth pursuits-Horizontal fast

◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

3.21 Smooth pursuits-Vertical slow

◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

Confidential

3.22 Smooth pursuits-Vertical fast

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3.23 Comment on smooth pursuits _____

3.24 Saccades - Horizontal small

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3.25 Saccades - Horizontal large

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3.26 Saccades -Vertical small

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3.27 Saccades - Vertical large

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3.28 Comment on saccades

3.29 FF-OKN - stripes moving to patient left

○ Normal
○ Reduced
○ Absent

3.30 FF-OKN - stripes moving to patient right

○ Normal
○ Reduced
○ Absent

3.31 Drum-OKN- stripes moving to patient left

○ Normal
○ Reduced
○ Absent

3.32 Drum-OKN- stripes moving to patient right

○ Normal
○ Reduced
○ Absent

3.33 Drum-OKN- stripes moving up

○ Normal
○ Reduced
○ Absent

3.34 Drum-OKN- stripes moving down

○ Normal
○ Reduced
○ Absent

3.35 Comment on OKN

_____

3.36 VOR - chair spinning to examiners right

_____
(Number of beats before stabilises)

3.37 VOR - chair spinning to examiners left

_____
(Number of beats before stabilises)

Confidential

3.38 Comment on VOR

_____

Section 4. Visuoperceptual/cognitive tests (all binocular)

4.1 Stereopsis _____

4.2 Comment on stereopsis

_____

4.3 Contour integration _____

4.4 Comment on contours

_____

4.5 Postbox

- ◯ Normal
- ◯ Some problems
- ◯ Couldn't do

4.6 Comment on postbox

_____

4.7 Rectangles -open

- ◯ Normal
- ◯ Some problems
- ◯ Couldn't do

4.8 Rectangles -closed

- ◯ Normal
- ◯ Some problems
- ◯ Couldn't do

4.9 Comment on rectangles

_____

4.10 Right eye sphere _____

4.11 Right eye cylinder (+ve) _____

4.11a Right eye axis _____

4.12 Left eye sphere _____

4.13 Left eye cylinder (+ve) _____

4.13a Left eye axis _____

4.14 Accommodation

- ◯ Normal
- ◯ Abnormal

4.15 Comment on accommodation

_____

Motion - RDK _____

Motion - Walker _____

Confidential

Perceptual tests worse than developmental age

- ◯ Yes
- ◯ No
- ◯ Tests not done
- ◯ Other

Refraction source

- ◯ Autorefraction at study clinic
- ◯ Own OO
- ◯ Hospital notes
- ◯ Other
- ◯ Refraction not done

Appendix 5 Movement Assessment Battery for Children-2

Patient ID_____

MOVEMENT ABC SCORES

Manual Dexterity - Component Score _____

Manual Dexterity - Standard Score _____

Manual Dexterity - Percentile _____

Aiming & Catching - Component Score _____

Aiming & Catching - Standard Score _____

Aiming & Catching - Percentile _____

Balance - Component Score _____

Balance - Standard Score _____

Balance - Percentile _____

TOTAL TEST SCORE

Total test score - Component score _____

Total test score - Standard score _____

Total test score - Percentile _____

Appendix 6 Strengths and Difficulties Questionnaire

Strengths and Difficulties Questionnaire P 4-17

For each item, please mark the box for Not True, Somewhat True or Certainly True. It would help us if you answered all items as best you can even if you are not absolutely certain or the item seems daft! Please give your answers on the basis of the child's behaviour over the last six months.

Child's Name ..... Male/Female

Date of Birth.....

	Not True	Somewhat True	Certainly True
Considerate of other people's feelings	☐	☐	☐
Restless, overactive, cannot stay still for long	☐	☐	☐
Often complains of headaches, stomach-aches or sickness	☐	☐	☐
Shares readily with other children (treats, toys, pencils etc.)	☐	☐	☐
Often has temper tantrums or hot tempers	☐	☐	☐
Rather solitary, tends to play alone	☐	☐	☐
Generally obedient, usually does what adults request	☐	☐	☐
Many worries, often seems worried	☐	☐	☐
Helpful if someone is hurt, upset or feeling ill	☐	☐	☐
Constantly fidgeting or squirming	☐	☐	☐
Has at least one good friend	☐	☐	☐
Often fights with other children or bullies them	☐	☐	☐
Often unhappy, down-hearted or tearful	☐	☐	☐
Generally liked by other children	☐	☐	☐
Easily distracted, concentration wanders	☐	☐	☐
Nervous or clingy in new situations, easily loses confidence	☐	☐	☐
Kind to younger children	☐	☐	☐
Often lies or cheats	☐	☐	☐
Picked on or bullied by other children	☐	☐	☐
Often volunteers to help others (parents, teachers, other children)	☐	☐	☐
Thinks things out before acting	☐	☐	☐
Steals from home, school or elsewhere	☐	☐	☐
Gets on better with adults than with other children	☐	☐	☐
Many fears, easily scared	☐	☐	☐
Sees tasks through to the end, good attention span	☐	☐	☐

Do you have any other comments or concerns?

Please turn over - there are a few more questions on the other side

Overall, do you think that your child has difficulties in one or more of the following areas: emotions, concentration, behaviour or being able to get on with other people?

No	Yes-minor difficulties	Yes-definite difficulties	Yes-severe difficulties
☐	☐	☐	☐

If you have answered "Yes", please answer the following questions about these difficulties:

- How long have these difficulties been present?

Less than a month	1-5 months	6-12 months	Over a year
☐	☐	☐	☐

- Do the difficulties upset or distress your child?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

- Do the difficulties interfere with your child's everyday life in the following areas?

	Not at all	Only a little	Quite a lot	A great deal
HOME LIFE	☐	☐	☐	☐
FRIENDSHIPS	☐	☐	☐	☐
CLASSROOM LEARNING	☐	☐	☐	☐
LEISURE ACTIVITIES	☐	☐	☐	☐

- Do the difficulties put a burden on you or the family as a whole?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

Signature ..... Date .....

Mother/Father/Other (please specify:)

Thank you very much for your help

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library