



OPEN ACCESS

# Hypothermia for perinatal asphyxia: trial-based quality of life at 6–7 years

Helen Campbell,<sup>1</sup> Oya Eddama,<sup>1</sup> Denis Azzopardi,<sup>2,3</sup> A David Edwards,<sup>2</sup> Brenda Strohm,<sup>1</sup> Oliver Rivero-Arias<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/archdischild-2017-313733>).

<sup>1</sup>National Perinatal Epidemiology Unit (NPEU), Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>2</sup>Centre for the Developing Brain, King's College, London, UK

<sup>3</sup>Institute of Clinical Sciences, Imperial College, London, UK

## Correspondence to

Dr Oliver Rivero-Arias, National Perinatal Epidemiology Unit (NPEU), Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK; [oliver.rivero@npeu.ox.ac.uk](mailto:oliver.rivero@npeu.ox.ac.uk)

HC and OR-A contributed equally.

Received 13 July 2017

Revised 17 January 2018

Accepted 18 January 2018

Published Online First

6 March 2018

## ABSTRACT

**Objective** To assess the impact of hypothermic neural rescue at birth on health-related quality of life (HRQL) in middle childhood.

**Design** Six-year to 7-year follow-up of surviving children from the Total Body Hypothermia for Neonatal Encephalopathy (TOBY) Trial.

**Setting** Community study including a single parental questionnaire to collect information on children's HRQL.

**Patients** 145 children (70 in the control group, 75 in the hypothermia group) whose parents consented and returned the questionnaire.

**Interventions** Intensive care with cooling of the body to 33.5°C for 72 hours or intensive care alone.

**Main outcome measures** HRQL attributes and utility scores using the Health Utilities Index (HUI).

**Results** At 6–7 years, speech appeared disproportionately affected when compared with other aspects of HRQL but levels of normal emotional functioning were similar in both groups. The mean (SE) HUI3 HRQL scores were 0.73 (0.05) in the hypothermia group and 0.62 (0.06) in the control group; mean difference (95% CI) 0.11 (–0.04 to 0.26).

**Conclusions** Findings of non-significant differences were not unexpected; the study used data from long-term survivors in a neonatal trial and was underpowered. However, results favoured moderate hypothermia and so complement the clinical results of the TOBY Children study. The work provides further insight into the long-term HRQL impact of perinatal asphyxial encephalopathy and provides previously unavailable utility data with which to contemplate the longer term cost-effectiveness of hypothermic neural rescue.

**Trial registration number** This study reports on the follow-up of the TOBY clinical trial: ClinicalTrials.gov number NCT01092637.

## INTRODUCTION

Perinatal asphyxial encephalopathy is associated with a high risk of death or early neurodevelopmental impairment, and survivors frequently develop functional disabilities and cognitive impairments in later childhood.<sup>1</sup> Evidence from randomised trials demonstrates that moderate hypothermia to 33°C–34°C for 72 hours initiated within 6 hours of delivery in infants with clear evidence of asphyxial encephalopathy reduces the risk of death or disability at 18–24 months of age and increases the rate of survival without disability.<sup>2</sup>

Since 2010, guidelines have recommended hypothermia for the treatment of moderate to severe perinatal asphyxial encephalopathy.<sup>3–6</sup> Longer term

## What is already known on this topic?

- The benefits of hypothermic neural rescue in terms of cognitive and disability-free survival persist into middle childhood.
- Data on generic health-related quality of life (HRQL) utilities following hypothermia have not previously been reported.

## What this study adds?

- New data on HRQL utility at 6–7 years after hypothermia plus intensive care and intensive care alone for perinatal asphyxia.
- Further insight into how different aspects of HRQL are affected following treatment for perinatal asphyxia.

follow-up of children in three major randomised trials of hypothermia (the Total Body Hypothermia for Neonatal Encephalopathy (TOBY) Trial, the CoolCap Trial and the National Institute of Child Health and Human Development (NICHD) Trial) has also confirmed that the clinical benefit at 18 months persists or is maintained, at least in part, into middle childhood.<sup>7–9</sup>

The question as to whether such clinical benefits translate into improvements in health-related quality of life (HRQL) for children is important. Parents want to know whether hypothermia can bring about long-term improvements in the overall well-being of their children. Healthcare professionals want to know what aspects of HRQL continue to be affected by perinatal asphyxial encephalopathy and hypothermia. Health economists need estimates of long-term generic HRQL to be able to contemplate cost-effectiveness. The CoolCap and NICHD Trials have reported various long-term data including those relating to the functional independence, cognitive outcomes and physical, emotional and social well-being of children.<sup>7–10</sup> In this paper, we present the first comparative analysis of generic HRQL using data from the Health Utilities Index (HUI), which were collected during the 6-year to 7-year follow-up of surviving children in the TOBY Trial.<sup>11</sup>

## METHODS

### Study population

Parents of surviving children who participated in the TOBY Trial were invited to take part in the TOBY Children Follow-up Study when their child reached 6–7 years of age.<sup>9</sup> The findings of TOBY



**To cite:** Campbell H, Eddama O, Azzopardi D, et al. *Arch Dis Child* 2018;**103**:654–659.

Children Follow-up Study are reported elsewhere in detail; however, in brief, the study reported that at 6–7 years and when compared with the control group, a higher number of children in the hypothermia group had survived with an IQ score  $\geq 85$ , survived without neurological abnormalities and had significant reductions in the risk of cerebral palsy and of moderate or severe disability.<sup>9</sup>

As part of that study, parents received a postal questionnaire including the HRQL HUI,<sup>11</sup> and questions about the use of healthcare services by their child over the previous 6 months (to be reported elsewhere). The study included centres from the UK, Sweden, Hungary, Finland and Israel, but UK centres contributed most to the sample used in the analysis.

### Child HRQL

Parents completed the HUI questionnaire on behalf of their children.<sup>11</sup> HUI measures generic health status and has two different scoring systems: HUI3 and HUI2. HUI3 was used here and is the preferred measure as it provides a more detailed descriptive system. HUI3 includes questions on eight attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain), whereas HUI2 contains six attributes (sensation, mobility, emotion, cognition, self-care and pain). HUI2 was used for a secondary analysis.

For each HUI attribute, respondents select one of a number of prespecified descriptions covering a range of functioning levels from best or normal (level 1) to most severe (level 3, 4, 5 or 6 depending on the attribute and scoring system). For example, level descriptors for the hearing attribute on the HUI3 range from 'Able to hear what is said in a group conversation with at least three other people, without a hearing aid' (level 1) to 'Unable to hear at all' (level 5). When taken together, responses for each attribute provide a health state description for an individual, and for both the HUI3 and HUI2, these descriptive systems can be converted into a single index utility score using algorithms developed from a sample of general population preferences for a subset of HUI health states.<sup>12–15</sup> Scores lie on a scale where 0 is equivalent to dead and 1 is equivalent to perfect health. Negative scores are also permitted and indicate health states considered worse than death. HUI was developed in Canada and so utility conversion algorithms from that country are widely used.<sup>12–13</sup> Algorithms for the HUI2 in the UK are also available.<sup>14–15</sup> In this study, the Canadian HUI3 and HUI2 algorithms and the HUI2 UK algorithm were both used.

In addition to the single index utility score, disability categories for HUI3 can also be derived using prespecified criteria based on the functional levels within each attribute.<sup>16</sup>

### Statistical analysis

Baseline characteristics between groups were compared using the  $\chi^2$  test for differences in proportions, the Mann-Whitney U test for median differences and the test for trend for ordered groups developed by Cuzick.<sup>17</sup> This test was also used when comparing HUI3 and HUI2 attribute levels between trial arms.

Single index utility scores were summarised using per patient means and standard errors. When comparing trial arms, mean differences and 95% CIs were estimated. Utility data were left skewed but as parametric and non-parametric confidence intervals were similar, only parametric intervals are reported.<sup>18</sup> Parental responses from all study countries were included.

Just over 5% of data were missing for utilities. Data were assumed to be missing at random, and multiple imputation (MI)

using chained equations was used to impute missing values (see online supplementary appendix).<sup>19–20</sup> Following imputation, Rubin's rule was used to generate combined estimates of means and SEs across MI datasets.<sup>21</sup> Analyses were conducted in Stata MP V.13.<sup>22</sup>

## RESULTS

### Study population

There were 229 TOBY Trial survivors at 6–7 years; of these, 184 (80%) parents/carers consented to participate in the main TOBY Children Study, and the remaining 45 were lost to follow-up. One hundred and forty-five of the 184 parents (79%) returned the postal questionnaire and form the sample for this study. One hundred and thirty of these 145 parents were from the UK (14 were from Hungary and 1 from Finland). Thirty-nine of 184 (21%) families either did not respond to or declined the questionnaire.

Seventy-five children were from the hypothermia arm and 70 from the control arm. There were no differences between arms in terms of baseline demographics or clinical characteristics at trial entry (table 1). At 6–7 years, a higher proportion of children in the hypothermia group had normal neurological function and an IQ score  $\geq 85$  and fewer had moderate or severe levels of disability.

A comparison of the 145 included children and the 39 non-participants revealed no significant differences (see online supplementary appendix table A1). Also, among the 39 non-participating families (and as observed for the participating families), a higher proportion of children in the hypothermia group than in the control group had an IQ  $\geq 85$  (16/23 (70%) vs 9/16 (56%),  $P=0.394$ ) and had normal neurological functioning (14/23 (61%) vs 6/16 (38%),  $P=0.151$ ).

When compared with the surviving children for whom primary outcome data were available at 6–7 years, the 45 children who were lost to follow-up had a higher frequency (although not significantly so) of severe abnormalities on amplitude-integrated electroencephalogram (EEG) at trial entry and lower scores on the Mental Development Index at 18 months. Eighteen of these 45 children (40%) were in the hypothermia arm and 27 (60%) were in the control arm.<sup>9</sup>

### Child HRQL

Table 2 shows the attribute levels recorded on the HUI3 (those for the HUI2 are shown in online supplementary appendix table A2). For each attribute, the proportion of children reported by their parents as functioning at level 1 was greater in the hypothermia group than in the control group, although not significantly so. Differences were particularly pronounced for the speech (74% vs 59%) and dexterity (76% vs 64%) attributes. The smallest difference was seen on the emotion attribute, with 85% of parents in the hypothermia arm and 84% in the control arm considering their children to be functioning at normal emotional levels.

Table 3 shows the HUI3 and HUI2 HRQL utility scores, by trial arm and for complete case and MI analyses. Following imputation, utility scores were universally lower. This is because children with missing data were less likely to have normal neurological function and an IQ  $\geq 85$  and were more likely to have multiple handicaps and moderate or severe levels of disability. Table 3 shows that for both analyses and for any utility conversion algorithm used, children in the hypothermia group, on average, presented with higher utility scores than children in the control group, although not significantly so.

**Table 1** Child demographics, clinical characteristics and neurological function at 6–7 years by trial arm for surviving children with questionnaire data

	Control group (n=70)	Hypothermia group (n=75)	P value
Baseline demographics and characteristics at trial entry			
Male sex, n (%)	41 (59)	47 (63)	0.61
Missing	0	0	
Age (years)			
Median (IQR)	6.3 (6.1–6.7)	6.3 (6.1–6.8)	0.36
Missing	3	1	
Gestational age (weeks)			
Median (IQR)	40.1 (39–41)	40.2 (39.3–41.4)	0.48
Missing	10	7	
Birth weight (g)			
Median (IQR)	3400 (3200–3850)	3460 (3172–3828)	0.80
Missing	0	0	
Delivery complications, n (%)	52 (75)	57 (77)	0.82
Missing	1	1	
Apgar score $\leq 5$ at 10 min, n (%)	42 (74)	41 (71)	0.72
Missing	13	17	
Outcomes at 6–7 years			
Normal neurological function, n (%)	35 (50)	50 (67)	0.04
Missing	0	0	
IQ $\geq 85$ , n (%)	43 (64)	59 (79)	0.06
Missing	3	0	
Overall disability*, n (%)			
None or mild	42 (60)	59 (81)	0.006
Moderate or severe	28 (40)	14 (19)	
Missing	0	2	

IQR Inter-quartile range; IQ Intelligence Quotient; \*Overall disability: mild disability (an IQ score of 70–84, level 1 gross motor function (is able to walk independently but may have some gait abnormalities), or abnormality in one or both eyes with normal or nearly normal vision); moderate disability (an IQ score of 55–69, level 2 or 3 gross motor function (has minimal ability to perform gross motor skills or requires assistance with walking) or moderately reduced vision); severe disability (an IQ score of  $<55$ , level 4 or 5 gross motor function (needs adaptive seating or has severely limited mobility) or no useful vision).

The associated disability categories for the HUI3 are shown in table 4 and similarly show lower levels of disability (although not significantly so) in children randomised to hypothermia.

## DISCUSSION

This study is the first to offer a comparative assessment of HRQL using the HUI for children 6–7 years after randomisation to standard care with hypothermia or standard care alone for perinatal asphyxia. The HUI is a leading measure of generic HRQL and provides rich information on levels of functioning for key health-related attributes.<sup>11</sup> It also facilitates the calculation of single index utility scores that summarise a subject's health status relative to the anchors of death and perfect health and that can be used to estimate quality-adjusted life years for use in cost-effectiveness analyses.<sup>11</sup>

The data presented here showed that for each individual HUI3 attribute, more parents in the hypothermia arm than in the control arm categorised their child as functioning at normal age

**Table 2** Distribution of responses (n, (%)) for each of the HUI3 attributes by trial arm

Level	Vision		Hearing		Speech		Emotion		Pain		Ambulation		Dexterity		Cognition	
	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75	Control n=70	Hypoth. n=75
1	56 (81.16)	64 (85.33)	63 (94.03)	71 (98.61)	41 (58.57)	55 (74.32)	59 (84.29)	63 (85.14)	54 (77.14)	61 (82.43)	48 (68.57)	59 (78.67)	45 (64.29)	57 (76.00)	45 (67.16)	53 (71.62)
2	7 (10.14)	8 (10.67)	3 (4.48)	0	9 (12.86)	2 (2.70)	6 (8.57)	11 (14.86)	7 (10.00)	8 (10.81)	2 (2.86)	2 (2.67)	2 (2.86)	2 (2.67)	6 (8.96)	6 (8.11)
3	1 (1.45)	1 (1.33)	1 (1.49)	1 (1.39)	7 (10.00)	7 (9.46)	2 (2.86)	0	6 (8.57)	3 (4.05)	2 (2.86)	0	3 (4.29)	0	2 (2.99)	3 (4.05)
4	0	0	0	0	3 (4.29)	2 (2.70)	1 (1.43)	0	2 (2.86)	1 (1.35)	5 (7.14)	4 (5.33)	4 (5.71)	7 (9.33)	4 (5.97)	2 (2.70)
5	2 (2.90)	1 (1.33)	0	0	10 (14.29)	8 (10.81)	2 (2.86)	0	1 (1.43)	1 (1.35)	2 (2.86)	2 (2.67)	3 (4.29)	3 (4.00)	4 (5.97)	5 (6.76)
6	3 (4.35)	1 (1.33)	0	0	–	–	–	–	–	–	11 (15.71)	8 (10.67)	13 (18.57)	6 (8.00)	6 (8.96)	5 (6.76)
Missing	1	0	3	3	0	1	0	1	0	1	0	0	0	0	3	1
Test for trend	0.76, P=0.446		1.42, P=0.154		1.72, P=0.086		0.32, P=0.751		0.88, P=0.380		1.34, P=0.181		1.66, P=0.096		0.60, P=0.551	

Descriptive labels for level 1 on each attribute are as follows:

Vision: able to see well enough to read ordinary newspaper and recognise a friend on the other side of the street, without glasses or contact lenses.

Hearing: able to hear what is said in a group conversation with at least three other people, without a hearing aid.

Speech: able to be understood completely when speaking with strangers or friends.

Emotion: happy and interested in life.

Pain: free of pain and discomfort.

Ambulation: able to walk around the neighbourhood without difficulty and without walking equipment.

Dexterity: full use of 2 hands and 10 fingers.

Cognition: able to remember most things, think clearly and solve day-to-day problems.

Hypoth., Hypothermia group.

**Table 3** HUI3 and HUI2 HRQL utility scores, by trial arm and for complete case and multiple imputation analyses

	Control group (n=70)					Hypothermia group (n=75)					Mean difference (95% CI)
HUI version	n	Missing	Mean	SD	SE	n	Missing	Mean	SD	SE	
Multiple imputation analysis											
HUI3 score (Canadian valuation)	70	0	0.618	–	0.058	75	0	0.730	–	0.049	0.112 (–0.038 to 0.262)
HUI2 score (Canadian valuation)	70	0	0.744	–	0.037	75	0	0.821	–	0.032	0.078 (–0.019 to 0.174)
HUI2 score (UK valuation)	70	0	0.733	–	0.038	75	0	0.809	–	0.032	0.076 (–0.021 to 0.174)
Complete case analysis*											
HUI3 score (Canadian valuation)	66	4	0.658	0.456	0.056	72	3	0.746	0.409	0.048	0.089 (–0.057 to 0.234)
HUI2 score (Canadian valuation)	66	4	0.772	0.292	0.036	71	4	0.829	0.266	0.032	0.057 (–0.038 to 0.151)
HUI2 score (UK valuation)	66	4	0.763	0.290	0.036	71	4	0.821	0.263	0.031	0.058 (–0.035 to 0.152)

SD standard deviation; SE standard error; CI parametric confidence intervals; \*Children for whom an HUI utility score was calculable from questionnaire responses.  
HRQL, health-related quality of life; HUI, Health Utilities Index.

appropriate levels (level 1) and that these differences fed through into higher average utility scores and lower disability levels for children randomised to hypothermia. None of the differences we observed however achieved statistical significance, and this was perhaps unsurprising because as the follow-up of a randomised trial and using a sample size fixed from the original trial, we were underpowered to detect such significant differences (an ex-postanalysis of power confirmed this). In addition, we cannot completely exclude the role of chance in our results. However, and when considered alongside the main clinical outcomes from the TOBY Children Study, the direction of the HRQL figures in favour of hypothermia appear intuitive.<sup>9</sup>

Differences in attribute levels between trial arms were particularly pronounced for the speech domain (table 2), with 74% and 59% of parents in the hypothermia and control groups, respectively, reporting their child to be functioning at normal levels (level 1). Also apparent is that the proportions of children categorised as normal on the speech attribute are noticeably lower than the proportions categorised as normal on the other attributes, suggesting that in general, for these children, speech may be disproportionately affected as compared with other aspects of HRQL. A similar finding has been observed elsewhere. The NICHD follow-up study assessed various measures

of intelligence for children 6–7 years after perinatal asphyxia and reported that verbal IQ scores generally showed greater deficits when compared with other measures such as performance IQ and processing speed quotients.<sup>10</sup>

Despite a greater proportion of children in the control arm having moderate to severe disabilities and abnormal neurological functioning, 84% were still reported as being ‘happy and interested in life’ (level 1) on the HUI3 emotion attribute as compared with 85% in the hypothermia arm. The similarity of these figures suggests that the level of disability faced by these children may not necessarily impact on their ability to be happy. In the NICHD follow-up study, parents also reported on their child’s emotional functioning at 6–7 years using the Child Health Questionnaire (CHQ).<sup>23</sup> Similar proportions in the hypothermia and control arms (75% vs 76%) reported they were ‘very satisfied’ with their child’s level of self-esteem.<sup>8</sup> Published research in children with cerebral palsy has similarly shown emotional functioning appears to be unrelated to the severity of the condition.<sup>24</sup> From parent-proxy responses on the HUI3, Kennes *et al*<sup>24</sup> showed correspondingly high proportions of children reported as being ‘happy and interested in life’ across five cerebral palsy severity categories. These findings are notable and may help parents of babies with disabilities as a consequence of perinatal asphyxia, who may anticipate that their child will suffer in the longer term. One must however acknowledge that these findings come from parent-proxy responses rather than directly from the children per se.

The utility scores computed using the HUI3 system were lower than those estimated using HUI2 (table 3), a finding also observed elsewhere and which is likely attributable to the HUI3 having more severity levels per attribute and therefore facilitating a greater range for states describing severe impairment.<sup>25</sup> Regardless of the system used, the difference between trial arms in perceived HRQL favoured hypothermia and is likely to be large enough to be important to the individuals involved. The developers of the HUI have reported that differences of 0.05 or greater are meaningful, and smaller differences might well be too.<sup>11</sup> As alluded to above, the larger utility scores observed with hypothermia complement the clinical results of the TOBY Children Study and suggest that the objective assessment of neurological function is mirrored in a sense of overall well-being.<sup>9</sup> However, and also worthy of note, is that utility scores for children in this study were lower than normative values (0.93) that have been recorded for children in the general population.<sup>26 27</sup>

Like the TOBY Children Study, the CoolCap and NICHD follow-up studies also used parents as proxy respondents, although neither used the HUI.<sup>7 8</sup> The CoolCap study used

**Table 4** HUI3 disability categories\* by trial arm

Disability category	Control group n (%)	Hypothermia group n (%)
1: No disability	26 (39)	39 (54)
2: Mild disability	10 (15)	8 (11)
3: Moderate disability	6 (9)	7 (10)
4: Severe disability	24 (36)	18 (25)
Total	66 (100)	72 (100)
Missing	4	3
Test for trend	1.72, P=0.085	

Disability categories described as follows:

Level 1: no disability or perfect health. The associated utility score value is 1.00, and all attributes (dimensions or domains) of health status are at their highest functional level (level 1).

Level 2: mild disability. The associated utility scores range from 0.89 to 0.99, and at least one attribute is at reduced level of function that can be readily corrected and/or does not prevent any activities.

Level 3: moderate disability. The associated utility score ranges from 0.70 to 0.88, and at least one attribute is at reduced level of function that cannot be corrected and/or prevents some activities.

Level 4: severe disability. The associated utility score is less than 0.70, and at least one attribute is at reduced level of function that cannot be corrected and prevents many activities.

\*Category labels and descriptions taken from: Feng *et al*.<sup>29</sup>

HUI, Health Utilities Index.

the WeeFIM (Functional Independence Measure for Children) questionnaire, which includes cognition, mobility and self-care domains like the HUI; however, the range of items included in each attribute and how scores are calculated is different, making direct comparisons difficult.<sup>7 28</sup> In the NICHD follow-up study, parents completed the CHQ, rating physical health and self-esteem.<sup>8 23</sup> Although not directly comparable with the HUI3 ambulatory and HUI2 mobility domains that here showed a greater proportion of children in the hypothermia group to have normal ability (table 2 and online supplementary appendix table A2), the physical health domain on the CHQ also showed a greater proportion of children in the hypothermia group, rated as having 'excellent' physical health at 6–7 years.

This study has a number of limitations. Questionnaire responses were unavailable for around 20% of children (39/184) whose parents consented to the TOBY Children Study. Analyses revealed no significant differences between responders and non-responders and also showed that when looking between trial arms at differences in clinical outcomes, the direction and magnitude of the differences observed were not dissimilar for both responding and non-responding families. Had these children been included in the analysis, one might hypothesise that the findings would not have altered substantially.

It is also necessary to contemplate the potential implications for our findings of the 45/229 (20%) surviving TOBY children who were lost to follow-up. Analyses showed these children to have had more severe abnormalities on EEG at trial entry and lower scores on the Mental Development Index at 18 months than children for whom follow-up data were available.<sup>9</sup> With two-fifths of these 45 children in the hypothermia arm and three-fifths in the control arm, had they been included in the analysis, then the potential exists for a lowering of the absolute utility scores observed here and for an increase in the difference in utility between the trial arms in favour of hypothermia.

Finally, the issue of missing data must be acknowledged; although, in this study, the amount of missing data was small at just 5%. Nevertheless, the mean difference favouring hypothermia increased following imputation (table 3). Of the children with missing utility data, more in the control arm than in the hypothermia arm had moderate or severe levels of disability (4/4 vs 2/4) and multiple disabilities (4/4 vs 2/4), which likely translated into lower imputed levels of utility.

## CONCLUSIONS

This study is the first to report comprehensive HRQL using the HUI in children surviving 6–7 years following randomisation to hypothermia or standard care alone for perinatal asphyxia. Non-significant differences in HRQL favouring moderate hypothermia corroborate the clinical findings of the TOBY Children Study. This study also provides previously unavailable data of interest to parents, health policy makers and health economists who may now wish to re-evaluate the long-term cost-effectiveness of hypothermia for perinatal asphyxia.

**Acknowledgements** The authors would like to thank the children and parents who participated in this study and made this research possible. We are grateful to Omar Omar and Louise Linsell from the NPEU at the University of Oxford for kindly clarifying all our data queries. We also thank Ursula Bowler and Maggie Redshaw from the NPEU at the University of Oxford for comments on an earlier version of this manuscript. A full list of acknowledgements is provided in Azzopardi D, *et al.*

**Contributors** HC conducted aspects of the statistical analyses, wrote the manuscript with OR-A and revised and approved the final version for submission. OE conducted aspects of the statistical analysis and revised and approved the final manuscript for submission. DA was the chief investigator of the TOBY Children Study,

secured funding for the study, contributed to the design of the statistical analysis plan and revised and approved the final manuscript for submission. ADE was a principal investigator of the TOBY Children Study, secured funding for the study, contributed to the design of the statistical analysis plan and revised and approved the final manuscript for submission. BS was the study coordinator of the TOBY Children Study, monitored and supervised the data collection, clarified data-related issues and revised and approved the final manuscript for submission. OR-A designed the statistical analysis plan of the study, supervised the statistical analysis, conducted the multiple imputation method for the missing data, wrote the manuscript jointly with HC and revised and approved the final version for submission.

**Funding** This study was supported by the UK Medical Research Council and the National Institute for Health Research Biomedical Research Centres at Imperial College London, the University of Oxford and King's College London.

**Disclaimer** The study sponsor had no role in the study design; collection, analysis and interpretation of data; writing the report; or in the decision to submit the article for publication.

**Competing interests** None declared.

**Patient consent** Parental/guardian consent obtained.

**Ethics approval** The National Research Ethics Service in the UK and the relevant ethics review board at each of the institutions outside the UK approved the TOBY protocol.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- de Vries LS, Jongmans MJ. Long-term outcome after neonatal hypoxic-ischaemic encephalopathy. *Arch Dis Child Fetal Neonatal Ed* 2010;95:F220–F224.
- Jacobs SE, Berg M, Hunt R, *et al.* Cochrane Neonatal Group. Cooling for newborns with hypoxic ischaemic encephalopathy. *Cochrane Database Syst Rev* 2013;69:CD003311.
- Papile LA, Baley JE, Benitz W, *et al.* Committee on Fetus and Newborn. Hypothermia and neonatal encephalopathy. *Pediatrics* 2014;133:1146–50.
- Kattwinkel J, Perlman JM, Aziz K, *et al.* American Heart Association. Neonatal resuscitation: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Pediatrics* 2010;126:e1400–13.
- National Institute for Health and Care Excellence. *Therapeutic hypothermia with intracorporeal temperature monitoring for hypoxic perinatal brain injury* ([www.nice.org.uk/ipg347](http://www.nice.org.uk/ipg347)). London: National Institute for Health and Care Excellence, 2010.
- Perlman JM, Wyllie J, Kattwinkel J, *et al.* Neonatal Resuscitation Chapter Collaborators. Neonatal resuscitation: 2010 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science with Treatment Recommendations. *Pediatrics* 2010;126:e1319–44.
- Guillet R, Edwards AD, Thoresen M, *et al.* CoolCap Trial Group. Seven- to eight-year follow-up of the CoolCap trial of head cooling for neonatal encephalopathy. *Pediatr Res* 2012;71:205–9.
- Shankaran S, Pappas A, McDonald SA, *et al.* Eunice Kennedy Shriver NICHD Neonatal Research Network. Childhood outcomes after hypothermia for neonatal encephalopathy. *N Engl J Med* 2012;366:2085–92.
- Azzopardi D, Strohm B, Marlow N, *et al.* TOBY Study Group. Effects of hypothermia for perinatal asphyxia on childhood outcomes. *N Engl J Med* 2014;371:140–9.
- Pappas A, Shankaran S, McDonald SA, *et al.* Hypothermia Extended Follow-up Subcommittee of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Cognitive outcomes after neonatal encephalopathy. *Pediatrics* 2015;135:e624–e634.
- Horsman J, Furlong W, Feeny D, *et al.* The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;1:54.
- Torrance GW, Feeny DH, Furlong WJ, *et al.* Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Med Care* 1996;34:702–22.
- Feeny D, Furlong W, Torrance GW, *et al.* Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002;40:113–28.
- Stevens K, McCabe C, Brazier J, *et al.* Multi-attribute utility function or statistical inference models: a comparison of health state valuation models using the HUI2 health state classification system. *J Health Econ* 2007;26:992–1002.
- McCabe CJ, Stevens KJ, Brazier JE. Utility scores for the Health Utilities Index Mark 2: an empirical assessment of alternative mapping functions. *Med Care* 2005;43:627–35.
- Furlong W, Feeny D, Torrance G. *Health Utilities Index: procedures manual for self-administered questionnaires*. Canada: Health Utility Index, 2012.

- 17 Cuzick J. A Wilcoxon-type test for trend. *Stat Med* 1985;4:87–90.
- 18 Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ* 1997;6:327–40.
- 19 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99.
- 20 Faria R, Gomes M, Epstein D, *et al.* A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics* 2014;32:1157–70.
- 21 Little RJ, Rubin DB. *Statistical Analysis with Missing Data*. 2nd edn. Hoboken, NJ: Wiley, 2002.
- 22 StataCorp LP. *Stata Statistical Software [program]*. College Station, TX: StataCorp LP, 2013.
- 23 Wake M, Salmon L, Reddihough D. Health status of Australian children with mild to severe cerebral palsy: cross-sectional survey using the Child Health Questionnaire. *Dev Med Child Neurol* 2003;45:194–9.
- 24 Kennes J, Rosenbaum P, Hanna SE, *et al.* Health status of school-aged children with cerebral palsy: information from a population-based sample. *Dev Med Child Neurol* 2002;44:240–7.
- 25 Neumann PJ, Sandberg EA, Araki SS, *et al.* A comparison of HUI2 and HUI3 utility scores in Alzheimer's disease. *Med Decis Making* 2000;20:413–22.
- 26 Petrou S, Kupek E. Estimating preference-based health utilities index mark 3 utility scores for childhood conditions in England and Scotland. *Med Decis Making* 2009;29:291–303.
- 27 Stade BC, Stevens B, Ungar WJ, *et al.* Health-related quality of life of Canadian children and youth prenatally exposed to alcohol. *Health Qual Life Outcomes* 2006;4:81.
- 28 Msall ME, DiGaudio K, Rogers BT, *et al.* The Functional Independence Measure for Children (WeeFIM). Conceptual basis and pilot use in children with developmental disabilities. *Clin Pediatr* 1994;33:421–30.
- 29 Feng Y, Bernier J, McIntosh C, *et al.* Validation of disability categories derived from Health Utilities Index Mark 3 scores. *Health Rep* 2009;20:43–50.