

ORIGINAL ARTICLE

Effects of Hypothermia for Perinatal Asphyxia on Childhood Outcomes

Denis Azzopardi, M.D., Brenda Strohm, R.N., Neil Marlow, D.M., Peter Brocklehurst, F.F.P.H., Aniko Deierl, M.D., Ph.D., Oya Eddama, Ph.D., Julia Goodwin, Ph.D., Henry L. Halliday, M.D., Edmund Juszczak, M.Sc., Olga Kapellou, M.D., Malcolm Levene, M.D., Louise Linsell, M.Sc., Omar Omar, M.Sc., Marianne Thoresen, M.D., Ph.D., Nora Tusor, M.D., Andrew Whitelaw, M.D., and A. David Edwards, D.Sc., for the TOBY Study Group*

ABSTRACT

BACKGROUND

In the Total Body Hypothermia for Neonatal Encephalopathy Trial (TOBY), newborns with asphyxial encephalopathy who received hypothermic therapy had improved neurologic outcomes at 18 months of age, but it is uncertain whether such therapy results in longer-term neurocognitive benefits.

METHODS

We randomly assigned 325 newborns with asphyxial encephalopathy who were born at a gestational age of 36 weeks or more to receive standard care alone (control) or standard care with hypothermia to a rectal temperature of 33 to 34°C for 72 hours within 6 hours after birth. We evaluated the neurocognitive function of these children at 6 to 7 years of age. The primary outcome of this analysis was the frequency of survival with an IQ score of 85 or higher.

RESULTS

A total of 75 of 145 children (52%) in the hypothermia group versus 52 of 132 (39%) in the control group survived with an IQ score of 85 or more (relative risk, 1.31; $P=0.04$). The proportions of children who died were similar in the hypothermia group and the control group (29% and 30%, respectively). More children in the hypothermia group than in the control group survived without neurologic abnormalities (65 of 145 [45%] vs. 37 of 132 [28%]; relative risk, 1.60; 95% confidence interval, 1.15 to 2.22). Among survivors, children in the hypothermia group, as compared with those in the control group, had significant reductions in the risk of cerebral palsy (21% vs. 36%, $P=0.03$) and the risk of moderate or severe disability (22% vs. 37%, $P=0.03$); they also had significantly better motor-function scores. There was no significant between-group difference in parental assessments of children's health status and in results on 10 of 11 psychometric tests.

CONCLUSIONS

Moderate hypothermia after perinatal asphyxia resulted in improved neurocognitive outcomes in middle childhood. (Funded by the United Kingdom Medical Research Council and others; TOBY ClinicalTrials.gov number, NCT01092637.)

From the Centre for the Developing Brain, King's College London (D.A., N.T., A.D.E.), Institute for Women's Health, University College London (N.M., P.B.), Institute of Clinical Sciences, Imperial College London (D.A., A.D.), and Homerton University Hospital (O.K.), London, the National Perinatal Epidemiology Unit Clinical Trials Unit, University of Oxford, Oxford (B.S., O.E., J.G., E.J., L.L., O.O.), Royal Maternity Hospital, Belfast (H.L.H.), University of Leeds, Leeds (M.L.), and University of Bristol, Bristol (M.T., A.W.) — all in the United Kingdom. Address reprint requests to Dr. Azzopardi at the Centre for the Developing Brain, Department of Perinatal Imaging, St. Thomas Hospital, King's College London, Westminster Bridge Rd., London SE1 7EH, United Kingdom, or at denis.azzopardi@kcl.ac.uk.

*A complete list of members of the Total Body Hypothermia for Neonatal Encephalopathy Trial (TOBY) study group is provided in the Supplementary Appendix, available at NEJM.org.

N Engl J Med 2014;371:140-9.

DOI: 10.1056/NEJMoa1315788

Copyright © 2014 Massachusetts Medical Society

PERINATAL ASPHYXIAL ENCEPHALOPATHY is associated with a high risk of death or early neurodevelopmental impairment. Among survivors, cerebral palsy, functional disability, and cognitive impairment often develop later in childhood. The cost of this condition to patients, their families, and society is high.

In several randomized, controlled trials involving infants with clear evidence of asphyxial encephalopathy, moderate hypothermia (33 to 34°C) for 72 hours, initiated within 6 hours after delivery, has been shown to reduce the risk of death or disability at 18 to 24 months of age and to increase the rate of survival free of disability.¹ The observation that hypothermia reduces the proportion of infants with abnormalities as seen on neuroimaging supports the hypothesis that early improvements in outcome should persist.² However, early assessment of neurodevelopmental outcomes may be altered on later follow-up, when more precise assessment is possible, and improvements observed in the short term may be only temporary.^{3,4} Data on long-term outcomes after neonatal hypothermia are lacking. In one previous study of neurologic outcomes at the age of 6 to 7 years in children who had been treated with hypothermia soon after birth, the children in the hypothermia group had an increased rate of survival as compared with those in the control group, but there was no significant between-group difference in the rate of the composite primary end point of death or an IQ score lower than 70 (46 of 97 children [47%] in the hypothermia group vs. 58 of 93 [62%] in the control group; relative risk in the hypothermia group, 0.78; 95% confidence interval [CI], 0.61 to 1.01).⁵ In addition, there were no significant reductions in other neurodevelopmental disabilities.

The Total Body Hypothermia for Neonatal Encephalopathy Trial (TOBY) was a large, randomized, controlled trial of hypothermia for perinatal asphyxial encephalopathy.⁶ At 18 months, children who had been treated with hypothermia had reduced risks of cerebral palsy and improved scores on the Mental Developmental Index and Psychomotor Developmental Index of the Bayley Scales of Infant Development II (BSID-II) and on the Gross Motor Function Classification System.⁶ Here we report the results of evaluation of the study population at 6 to 7 years of age to determine whether the use of moderate hypothermia after perinatal asphyxia is associated with longer-

term benefits. The primary outcome was the frequency of survival with an IQ score of 85 or higher at 6 to 7 years, as assessed with the use of standardized tests.

METHODS

STUDY DESIGN

In the TOBY trial, 325 infants with a gestational age of at least 36 weeks who had moderate-to-severe asphyxial encephalopathy and abnormal results on amplitude-integrated electroencephalography (EEG) were randomly assigned within 6 hours after birth to receive standard care alone (control) or standard care with hypothermia to a rectal temperature of 33 to 34°C for 72 hours, followed by slow rewarming. Hypothermia was maintained by nursing the infant on a cooling blanket. The children were enrolled in the study from 2002 to 2006, and follow-up at 6 to 7 years was conducted from 2009 to 2013.

The National Research Ethics Service in the United Kingdom and the relevant ethics review board at each of the institutions outside the United Kingdom approved the TOBY protocol (available with the full text of this article at NEJM.org), and an independent trial steering committee oversaw the study. Written informed consent was obtained from the children's parents.

A letter of invitation, together with an information leaflet and consent form, was mailed to the parents of surviving children. After parental consent had been obtained, a psychologist and a pediatrician, both of whom were unaware of the study-group assignments, performed the assessments, usually at the child's school. The assessment comprised a neurologic examination and a neuropsychological assessment encompassing sensory function, cognition, memory, attention, and executive function, all areas that are likely to be affected by perinatal asphyxia. Questionnaire responses were requested from parents and teachers.

NEUROLOGIC EXAMINATION

A structured neurologic examination to detect signs of cerebral palsy and minor neurologic dysfunction was performed.^{7,8} Neuromotor function was assessed with the use of the Gross Motor Function Classification System and the Manual Ability Classification System; scores on the two assessments range from 1 to 5, with higher scores indicating greater impairment.^{9,10}

PSYCHOMETRIC ASSESSMENT

We used the Wechsler Preschool and Primary Scale of Intelligence III (WPPSI-III) test or the Wechsler Intelligence Scale for Children IV (WISC-IV), standardized for children in the United Kingdom (2004) or as appropriate in other countries, to evaluate general cognitive performance.¹¹ Results on these two tests include a general measure of IQ (the primary outcome) with quotients for verbal and nonverbal performance and processing speed. All results are expressed as an age-standardized score, with a population mean of 100 and a standard deviation of 15.

OTHER ASSESSMENTS

We also administered age-appropriate subtests selected from the following domains of the Developmental Neuropsychological Assessment II (NEPSY-II): attention and executive function, visuospatial processing, sensorimotor function, and memory and learning.¹² We report the mean scores for each domain, derived from the mean standardized subset scores.

Since deficits in working memory have been reported in children after neonatal encephalopathy and hypoxia,^{13,14} we also assessed scores from the following three subtests of the Working Memory Test Battery for Children: block recall, to assess nonverbal short-term memory; digit recall, to assess verbal short-term memory; and backward digit recall, to assess central executive function.¹⁵

Parents and teachers, who were aware of study-group assignments, completed the Strengths and Difficulties Questionnaire (www.sdqinfo.com) (on a scale of 0 to 40, with higher scores indicating more difficulties) and the Attention Deficit–Hyperactivity Disorder (ADHD) Rating Scale IV (on a scale of 0 to 54, with higher scores indicating more severe symptoms).¹⁶ Appropriate translations were available for the European centers. Parents were also asked to complete questionnaires assessing their child's behavior, everyday memory, use of health care services, and demographic information and provided information about their child's health status with the use of the Health Utilities Index (HUI) as proxy respondents. We converted the responses into multi-attribute utility scores using published algorithms (HUI2 and HUI3); these range from -0.03 to 1.00 for HUI2 and from -0.36 to 1.00 for HUI3; the lowest score represents the utility of the worst possible state, with 0.00 indicating death and 1.00 perfect health,

although negative scoring is possible. A change of 0.03 points in the mean overall score is considered to be clinically meaningful.^{17,18}

We obtained permission from parents to request information about educational attainment from the child's school. Teachers and parents then completed two detailed questionnaires — the Total Academic Achievement Score¹⁹ and the Strengths and Difficulties Questionnaire (with the latter including responses from both teachers and parents) — and provided information about any special educational needs.

STUDY OUTCOMES

The primary outcome was the frequency of survival with an IQ score of 85 or higher (1 SD below the standardized mean IQ score for the general population). Other prespecified outcomes were components of the primary outcome; the frequency of survival without neurologic abnormalities, which was defined as an IQ score of 85 or higher, a normal neurologic examination, normal vision, and normal hearing; the full-scale IQ score and subscale scores on the WPPSI-III or WISC-IV; NEPSY-II domain scores; overall memory score; score for total difficulties from the parental questionnaire; overall ADHD rating; prevalence of cerebral palsy; scores for gross motor function and manual ability; the grade of disability, which was categorized as mild disability (an IQ score of 70 to 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 to 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), or 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); teacher's score for academic achievement; and score on the Health Utility Index (HUI2 and HUI3).

PARTIAL ASSESSMENTS AND MISSING DATA

Two assessors who were unaware of the study-group assignments independently classified the children with missing or incomplete results on the WPPSI-III or WISC-IV into two IQ groups (<85 or ≥ 85) on the basis of additional information that was available. Children were classified into the group with an IQ score of 85 or higher if

they had an IQ score of 85 or higher on any full subscale of the WPPSI-III and were unable to complete the other sections because of physical impairments. We included these children's adjudicated outcomes in the primary outcome analysis.

For any partially completed scores on the Strengths and Difficulties Questionnaire and ADHD ratings from parents and teachers, pro rata estimation was used if 10% or less of the items were missing (i.e., up to two items missing on each scale). Children for whom we had no information after the 18-month assessment were not included in the primary outcome analysis.

SENSITIVITY ANALYSES

We used two tests to explore the sensitivity of the primary outcome to missing data. In the first test, we assigned nonparticipants an IQ score of 85 or higher at the age of 6 to 7 years for those with a Mental Developmental Index of 85 or higher at 18 months and an IQ score of less than 85 for those with a Mental Developmental Index of less than 85 at 18 months. In the second test, we assigned participating children who were unable to complete the full-scale WPPSI-III or WISC-IV to the group with an IQ score lower than 85.

Table 1. Demographic and Clinical Characteristics of the Children at Trial Entry.*

Characteristic	Outcomes at 6 to 7 Years Available			Outcomes at 6 to 7 Years Not Available
	Hypothermia Group (N=145)	Control Group (N=135)	All Children (N=280)	All Children (N=45)
Male sex — no. (%)	91 (63)	74 (55)	165 (59)	24 (53)
Gestational age				
Median (IQR) — wk	40.3 (39.3–41.3)	40.1 (39.0–41.1)	40.2 (39.0–41.1)	40.0 (37.7–41.1)
Missing data — no.	16	16	32	1
Birth weight				
Median (IQR) — g	3467 (3053–3883)	3351 (3060–3700)	3428 (3055–3790)	3142 (2740–3790)
Missing data — no.	1	1	2	0
Age at randomization — no. (%)				
<4 hr	45 (31)	52 (39)	97 (35)	8 (18)
4–6 hr	100 (69)	83 (61)	183 (65)	37 (82)†
Maternal pyrexia during labor — no./total no. (%)	10/143 (7)	9/131 (7)	19/274 (7)	1/43 (2)
Delivery complications — no./total no. (%)	102/143 (71)	98/133 (74)	200/276 (72)	34/45 (76)
Apgar score ≤5 at 10 min — no./total no. (%)	98/117 (84)	88/113 (78)	186/230 (81)	29/38 (76)
Clinical seizures — no./total no. (%)	84/141 (60)	66/130 (51)	150/271 (55)	25/43 (58)
Temperature at randomization				
Mean — °C	36.6±1.1	36.4±1.2	36.5±1.1	36.9±1.0†
Missing data — no.	0	2	2	0
Abnormalities on aEEG at randomization — no. (%)				
Moderate	62 (43)	57 (42)	119 (42)	13 (29)
Severe	83 (57)	78 (58)	161 (58)	32 (71)

* Plus–minus values are means ±SD. There were no significant differences between the hypothermia group and the control group among children for whom follow-up data were available. The abbreviation aEEG denotes amplitude-integrated electroencephalography, and IQR denotes interquartile range.

† P<0.05 for the comparison with children who had available outcomes at 6 to 7 years.

STATISTICAL ANALYSIS

We calculated relative risks and 95% confidence intervals on the basis of the proportion of children with an IQ score of 85 or higher and the proportions with secondary outcomes at the age of 6 or 7 years in the hypothermia group, as compared with the control group. We prespecified a test of interaction between treatment and abnormality on amplitude-integrated EEG, which provides an index of the severity of the asphyxial insult, and examined the effect of maternal education (classified according to whether the mother had graduated from high school) on the primary outcome.

For normally distributed continuous outcomes, we present means and standard deviations for each group and the mean between-group differences plus 95% confidence intervals. For non-Gaussian distributions, we present the median and interquartile range for each group.

A two-sided P value of less than 0.05 was considered to indicate statistical significance. No adjustment was made for multiple testing of the secondary outcome measures.

RESULTS**PATIENTS**

Outcome data were available for 280 children (184 survivors and 96 children who died before

the assessment at 6 to 7 years of age). Of these children, 245 (88%) were from the United Kingdom, and the rest were from other countries. Parental educational level and socioeconomic status were similar in the two study groups, and the clinical characteristics of the children were similar at trial entry (Table 1; and Tables S1 through S4 in the Supplementary Appendix, available at NEJM.org). A total of 45 children did not participate in the follow-up study and were excluded from the primary analysis (Fig. 1). As compared with participants, nonparticipants had a higher rectal temperature at the time of randomization, were significantly less likely to enter the study within 4 hours after birth, and had a higher frequency of severe abnormalities on amplitude-integrated EEG, although the last difference was not significant; at 18 months, nonparticipants had lower scores on the Mental Developmental Index.

PRIMARY OUTCOME

Of the 184 surviving children who participated in the follow-up evaluation, 140 underwent full-scale IQ testing with the WPPSI-III or the WISC-IV. Another 41 children (18 in the hypothermia group and 23 in the control group) were unable to complete the test, in most cases because of physical impairment (37 children) or because they would not cooperate. Of the 41 children who did not

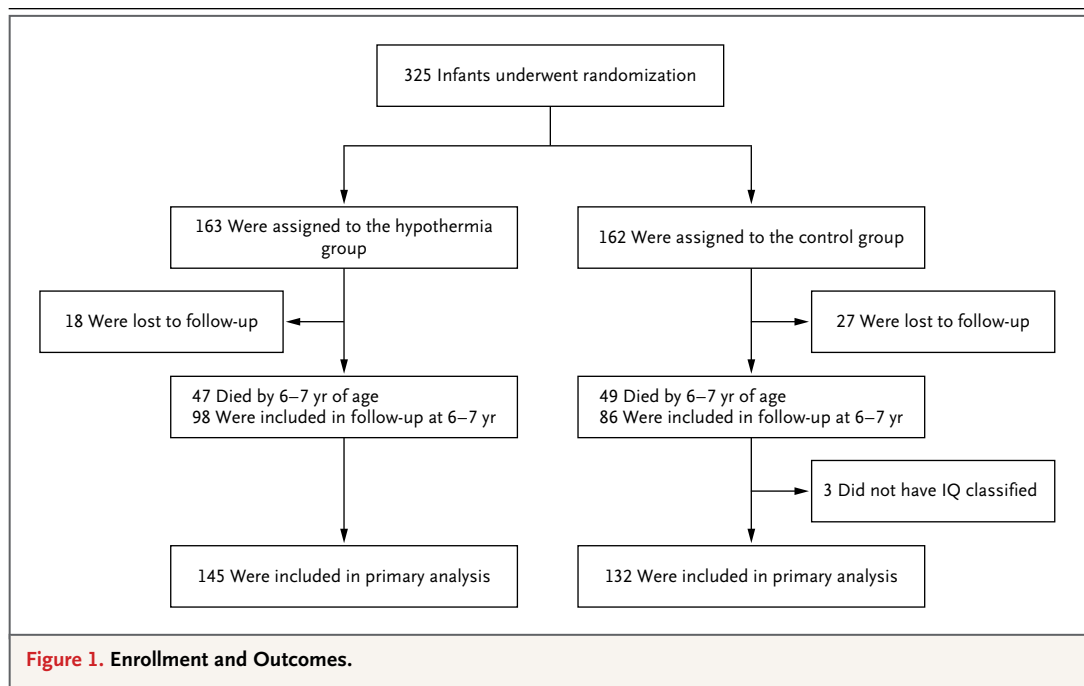


Table 2. Primary Outcome and Its Components in Children 6 to 7 Years of Age.*

	Hypothermia Group (N=163)	Control Group (N=162)	Relative Risk (95% CI)	P Value
Survival with IQ score ≥ 85 among all children who could be tested — no./total no. (%)	75/145 (52)	52/132 (39)	1.31 (1.01–1.71)	0.04
Death — no./total no. (%)	47/163 (29)	49/162 (30)	0.95 (0.68–1.33)	0.81
IQ score ≥ 85 among survivors — no./total no. (%)	75/98 (77)	52/83 (63)	1.22 (1.00–1.49)	0.05

* IQ scores could not be determined for 18 children in the hypothermia group and 30 in the control group. There was no significant interaction between treatment and results on amplitude-integrated EEG (risk ratio, 0.99; 95% CI, 0.59 to 1.67; $P=0.97$).

Table 3. Pediatric Assessment in Survivors.

Variable	Hypothermia Group no./total no. (%)	Control Group no./total no. (%)	Relative Risk (95% CI)	P Value
Grade of disability*				
No disability	65/96 (68)	37/83 (45)	1.52 (1.15–2.00)	0.002
Mild disability	10/96 (10)	15/83 (18)		
Moderate disability	8/96 (8)	11/83 (13)		
Severe disability	13/96 (14)	20/83 (24)		
Moderate or severe disability	21/96 (22)	31/83 (37)	0.59 (0.37–0.94)	0.03
Cerebral palsy†	21/98 (21)	31/86 (36)	0.59 (0.37–0.95)	0.03
Score on Gross Motor Function Classification System‡				
No abnormality	76/98 (78)	49/83 (59)	1.31 (1.07–1.62)	0.01
Level 1–2	6/98 (6)	13/83 (16)		
Level 3–5	16/98 (16)	21/83 (25)	0.65 (0.36–1.15)	0.14
Score on Manual Ability Classification System§				
No abnormality	75/98 (77)	51/83 (61)	1.25 (1.02–1.53)	0.04
Level 1–2	4/98 (4)	8/83 (10)		
Level 3–5	19/98 (19)	24/83 (29)	0.67 (0.40–1.13)	0.16
Visual impairment not corrected by eyeglasses	7/98 (7)	10/82 (12)	0.59 (0.23–1.47)	0.31
Blindness	1/98 (1)	1/82 (1)	0.84 (0.05–13.17)	1.00
Hearing impairment	4/98 (4)	8/83 (10)	0.42 (0.13–1.36)	0.15

* Two children in the hypothermia group could not be classified. $P=0.002$ for trend.

† Two children in the hypothermia group and two in the control group could not be classified and were subsequently found not to have cerebral palsy.

‡ Scoring on the Gross Motor Function Classification System is as follows: level 1, able to walk independently but may have some gait abnormalities; level 2, able to walk in most settings but with only minimal ability to perform gross motor skills such as running and jumping; level 3, walks with handheld assistive device and when seated may need seat belt for balance; level 4, requires physical assistance or powered mobility and needs adaptive seating; and level 5, severely limited in mobility, with limited ability to maintain antigravity head and trunk postures. $P=0.01$ for trend.

§ Scoring on the Manual Ability Classification System is as follows: level 1, handles objects easily and successfully; level 2, handles most objects but with somewhat reduced quality or speed of achievement; level 3, handles objects with difficulty; needs help to prepare or modify activities; level 4, handles a limited selection of easily managed objects in adapted situations; level 5, does not handle objects and has severely limited ability to perform even simple actions. $P=0.04$ for trend.

undergo full-scale IQ testing, 35 (15 in the hypothermia group and 20 in the control group) were classified as having an IQ score lower than 85; 3 children in each group were classified as having an IQ score of 85 or higher by the expert assessors. For another 3 children, information was available from the parents only, and the IQ score could not be classified. Ultimately, 277 of 325 children in the TOBY population (85%) were included in the primary analysis. Scores on the Mental Developmental Index at 18 months were available for 39 children who were lost to follow-up at 6 to 7 years of age, and thus 316 children were included in the sensitivity analysis that relied on

earlier data for children for whom a current IQ measurement was missing.

The frequency of survival with an IQ score of 85 or higher was 52% (75 of 145 children) in the hypothermia group as compared with 39% (52 of 132 children) in the control group (relative risk, 1.31; 95% CI, 1.01 to 1.71; $P=0.04$) (Table 2). The number of children who would need to be treated with hypothermia in order to prevent 1 child from dying or having an IQ score lower than 85 was 8 (95% CI, 4 to 145). There was no significant interaction between treatment and grade of abnormality on amplitude-integrated EEG ($P=0.97$). The results of

Table 4. Psychometric Assessment in Survivors.*

	Hypothermia Group (N=98)	Control Group (N=86)	Difference in Means (95% CI)	P Value
Intelligence†				
Full-scale IQ score				
Mean	103.6±14.4	98.5±18.9	5.1 (-0.4 to 10.6)	0.07
Missing data (no.)	18	25		
Verbal IQ score				
Mean	105.2±15.6	101.1±17.3	4.0 (-1.6 to 9.5)	0.16
Missing data (no.)	21	25		
Nonverbal performance score				
Mean	101.1±15.0	96.7±19.0	4.4 (-1.2 to 10.1)	0.12
Missing data (no.)	19	22		
Processing speed score				
Mean	98.7±12.4	95.3±18.7	3.4 (-2.0 to 8.9)	0.22
Missing data (no.)	25	31		
Score on Developmental Neuropsychological Assessment‡				
Attention and executive function				
Mean	9.6±2.1	8.6±2.7	1.02 (0.12 to 1.92)	0.03
Missing data (no.)	32	38		
Memory and learning				
Mean	10.0±2.2	9.4±2.8	0.58 (-0.28 to 1.44)	0.18
Missing data (no.)	22	33		
Sensorimotor processing				
Mean	8.1±2.6	7.1±2.7	0.97 (-0.05 to 1.91)	0.06
Missing data (no.)	27	38		
Visuospatial processing				
Mean	10.4±3.1	9.6±3.5	0.82 (-0.35 to 2.0)	0.17
Missing data (no.)	23	34		

Table 4. (Continued.)				
	Hypothermia Group (N=98)	Control Group (N=86)	Difference in Means (95% CI)	P Value
Score on the Working Memory Test Battery for Children [§]				
Digit recall				
Mean	104.1±17.3	106.1±15.3	-2.0 (-8.3 to 4.3)	0.54
Missing data (no.)	36	39		
Block recall				
Mean	97.0±18.3	94.7±15.6	2.3 (-4.3 to 8.9)	0.49
Missing data (no.)	36	39		
Backward digit recall				
Mean	96.1±15.4	95.3±15.2	0.7 (-5.3 to 6.7)	0.81
Missing data (no.)	38	41		

* Plus-minus values are means ±SD. A positive between-group difference favors the hypothermia group, and a negative difference favors the control group.

† Intelligence was measured with the use of the Wechsler Preschool and Primary Scale of Intelligence III (WPPSI-III) or the Wechsler Intelligence Scale for Children IV (WISC-IV). Each score on the WPPSI-III and WISC-IV is standardized to have an average of 100, with scores of 115 and 85 representing 1 SD above and below the mean.

‡ Each subscale on the Developmental Neuropsychological Assessment II (NEPSY-II) ranges from 1 to 19, and a score of 5 or less is equivalent to a percentile rank below 25, which is below the expected level of ability for the child's age.

§ Scores on the Working Memory Test Battery for Children are standardized in the same manner as are scores on the WPPSI-III and WISC-IV.

an analysis adjusted for maternal educational level were materially unchanged, as were the results of a sensitivity analysis that included data from nonparticipants (relative risk, 1.31; 95% CI, 1.02 to 1.67) and a sensitivity analysis that included children who were not able to complete IQ testing and were assigned to the group with an IQ score lower than 85 (relative risk, 1.37; 95% CI, 1.04 to 1.80) (Tables S5 through S8 in the Supplementary Appendix).

OTHER OUTCOMES

The rates of death did not differ significantly between the hypothermia group (in which 47 of 163 children [29%] died) and the control group (in which 49 of 162 children [30%] died) (Table 2). A total of 86 of the 96 deaths (90%) occurred before the 18-month assessment. A higher proportion of survivors in the hypothermia group than in the control group had an IQ score of 85 or higher (77% vs. 63%; relative risk, 1.22; 95% CI, 1.00 to 1.49; $P=0.05$) (Table 2). Significantly more children in the hypothermia group than in the control group survived without neurologic abnormalities (65 of 145 [45%] vs. 37 of 132 [28%];

relative risk, 1.60; 95% CI, 1.15 to 2.22) (Table S9 in the Supplementary Appendix). Among survivors, children in the hypothermia group had significantly reduced rates of cerebral palsy (21% vs. 36%, $P=0.03$) and moderate or severe disability (22% vs. 37%, $P=0.03$) and had significantly better scores for gross motor function and manual ability (Table 3). The rates of visual and hearing impairments did not differ significantly between the two groups, and head circumference and growth were similar in the two groups (Table S10 in the Supplementary Appendix).

Approximately 30% of children did not complete the psychometric tests. Among the children who completed the psychometric tests, there was no significant between-group difference with respect to IQ scores that were measured on a continuous scale and other scores, with the exception of 1 of the 11 scores compared (attention and executive function, $P=0.03$) (Table 4). There were also no significant differences between groups in mean scores on the index of health care status derived from parental assessments (HUI2 and HUI3) or in parental scores for children's strengths and difficulties and ADHD

scores (Tables S9, S10, and S11 in the Supplementary Appendix).

The mean difference in the academic achievement score favored the hypothermia group but was not significant. The proportion of children who required use of special educational resources was lower in the hypothermia group than in the control group (8.2% vs. 26.9%; relative risk, 0.30; 95% CI, 0.12 to 0.79; $P=0.01$) (Table S12 in the Supplementary Appendix).

DISCUSSION

Children with asphyxial encephalopathy who were treated with hypothermia shortly after birth were significantly more likely to survive with an IQ score of 85 or higher at 6 to 7 years of age than were children who did not undergo such therapy. Similar proportions of children in the two groups died, but a higher proportion of survivors in the hypothermia group had an IQ score of 85 or higher, and the frequency of moderate-to-severe disability was lower in this group than in the control group.

Data are lacking with respect to longer-term outcomes after hypothermia therapy for perinatal asphyxia. The CoolCap study, which evaluated the efficacy of selective head cooling on the basis of questionnaire data from parents of 62 of 135 children (46%) at 7 to 8 years of age, showed that childhood outcomes broadly correlated with the 18-month assessments, although the study was underpowered to examine the effect of hypothermia on cognitive function at an older age.²⁰ In the Childhood Outcomes after Hypothermia for Neonatal Encephalopathy study, sponsored by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), there was a high follow-up rate among children 6 to 7 years of age, with children in the hypothermia group having a lower rate of death than those in the control group. However, the study showed no significant differences in rates of disability or cognitive outcomes, perhaps because of a lack of statistical power.⁵ In our study, in which we used the same psychometric tests as were used in the NICHD study but used a different cutoff point for the IQ score in our primary outcome, we found significant reductions in the rates of moderate-to-severe cognitive deficiency in the hypothermia group, as compared with the control group, but similar rates of mortality in the two groups.

Our study was limited by the lack of availability of primary outcome data for 15% of the original TOBY trial population (11% of the hypothermia group and 18% of the control group). In anticipation of dropout, we prespecified a sensitivity analysis in which we substituted the Mental Developmental Index score at 18 months for the missing IQ score in order to examine the robustness of the results. The high correlation for neurologic outcomes between earlier assessment and later follow-up that we observed among children who had undergone the two assessments, as well as correlation with the findings of the CoolCap and NICHD studies, provides support for our approach, and the close agreement of the results of the primary and sensitivity analyses increases our confidence in the conclusions.

Several children were unable to complete the full-scale WPPSI-III or WISC-IV test, in most cases because of physical impairments. Assessors who were unaware of the study-group assignments classified these children as having an IQ score below 85 or an IQ of 85 or higher, which allowed them to be included in the primary analysis. Almost all these children (all but six) were classified as having an IQ score that was lower than 85. The results of our primary outcome analysis were not materially altered when all the children without results on full-scale IQ testing were classified as having an IQ score of less than 85.

We found no significant between-group differences for many secondary outcomes assessed, including 10 of 11 psychometric test scores. Prioritizing assessment of IQ as the primary outcome meant that some other tests were curtailed if the assessor determined that the child could not maintain adequate performance. Thus, our study was not adequately powered for many of these comparisons. Most point estimates favored the hypothermia group, though the differences were modest.

In conclusion, our study provides evidence that the benefits of moderate hypothermia after perinatal asphyxia persist into middle childhood.

Supported by the United Kingdom 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.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Dr. Maggie Redshaw for her assistance in the classification of children with missing or partial assessments. A full list of acknowledgments is provided in the Supplementary Appendix.

REFERENCES

1. Jacobs SE, Berg M, Hunt R, Tarnow-Mordi WO, Inder TE, Davis PG. Cooling for newborns with hypoxic ischaemic encephalopathy. *Cochrane Database Syst Rev* 2013;1:CD003311.
2. Rutherford M, Ramenghi LA, Edwards AD, et al. Assessment of brain tissue injury after moderate hypothermia in neonates with hypoxic-ischaemic encephalopathy: a nested substudy of a randomised controlled trial. *Lancet Neurol* 2010;9:39-45.
3. Barnett AL, Guzzetta A, Mercuri E, et al. Can the Griffiths scales predict neuromotor and perceptual-motor impairment in term infants with neonatal encephalopathy? *Arch Dis Child* 2004;89:637-43.
4. Schmidt B, Anderson PJ, Doyle LW, et al. Survival without disability to age 5 years after neonatal caffeine therapy for apnea of prematurity. *JAMA* 2012;307:275-82.
5. Shankaran S, Pappas A, McDonald SA, et al. Childhood outcomes after hypothermia for neonatal encephalopathy. *N Engl J Med* 2012;366:2085-92. [Erratum, *N Engl J Med* 2012;367:1073.]
6. Azzopardi DV, Strohm B, Edwards AD, et al. Moderate hypothermia to treat perinatal asphyxial encephalopathy. *N Engl J Med* 2009;361:1349-58.
7. Touwen BCL. Examination of the child with minor neurological dysfunction. 2nd ed. London: SIMP/Heinemann, 1979.
8. Cans C. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Dev Med Child Neurol* 2000;42:816-24.
9. Eliasson AC, Krumlinde-Sundholm L, Rösblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol* 2006;48:549-54.
10. Rosenbaum PL, Palisano RJ, Bartlett DJ, Galuppi BE, Russell DJ. Development of the Gross Motor Function Classification System for cerebral palsy. *Dev Med Child Neurol* 2008;50:249-53.
11. Wechsler D. Wechsler Pre-school and Primary Scale of Intelligence — third UK edition (WPPSI-III UK). London: Harcourt, 2004.
12. Korkman M, Kirk U, Kemp S. NEPSY-II: a developmental neuropsychological assessment. 2nd ed. San Antonio, TX: Psychological Corporation, 2007.
13. de Haan M, Mishkin M, Baldeweg T, Vargha-Khadem F. Human memory development and its dysfunction after early hippocampal injury. *Trends Neurosci* 2006; 29:374-81.
14. Marlow N, Rose AS, Rands CE, Draper ES. Neuropsychological and educational problems at school age associated with neonatal encephalopathy. *Arch Dis Child Fetal Neonatal Ed* 2005;90:F380-F387.
15. Pickering SJ, Gathercole SE. Working Memory Test Battery for Children. San Antonio, TX: Psychological Corporation, 2001.
16. Reid R, DuPaul GJ, Power TJ, et al. Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. *J Abnorm Child Psychol* 1998;26:187-98.
17. Drummond M. Introducing economic and quality of life measurements into clinical studies. *Ann Med* 2001;33:344-9.
18. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;1:54.
19. Wolke D, Rizzo P, Woods S. Persistent infant crying and hyperactivity problems in middle childhood. *Pediatrics* 2002;109: 1054-60.
20. Guillet R, Edwards AD, Thoresen M, et al. Seven- to eight-year follow-up of the CoolCap trial of head cooling for neonatal encephalopathy. *Pediatr Res* 2012;71:205-9.

Copyright © 2014 Massachusetts Medical Society.

POSTING PRESENTATIONS FROM MEDICAL MEETINGS ONLINE

Online posting of an audio or video recording of an oral presentation at a medical meeting, with selected slides from the presentation, is not considered prior publication. Authors should feel free to call or send e-mail to the *Journal's* Editorial Offices if there are any questions about this policy.