No change in neurodevelopment at 11-years after extremely preterm birth

Neil Marlow DM FMedSci[1], Yanyan Ni PhD[1], Rebecca Lancaster MBBS[2], Emmi Suonpera BSc[1], Marialivia Bernardi PhD[1], Amanda Fahy PhD[1], Jennifer Larsen MBBS[2], Jayne Trickett PhD[2], John R Hurst PhD[3], Joan Morris PhD[4], Dieter Wolke PhD[5], Samantha Johnson PhD[2].

[1]Academic Department of Neonatology, UCL EGA Institute for Women's Health, University College London, London, UK;

[2]Department of Health Sciences, University of Leicester, Leicester, UK;

[3]UCL Respiratory, University College London, London UK;

[4]St Georges, University of London, London UK;

[5]Department of Psychology, University of Warwick, Coventry, UK

Corresponding Author:

> Professor Neil Marlow
> ORCID ID 0000-0001-5890-2953
> UCL Institute for Women's Health
> 74 Huntley Street
> London WC1E 6AU
> United Kingdom
> n.marlow@ucl.ac.uk
> +44 20 7679 0834

Running Head: Change in impairment following extremely preterm birth

Word Count: for manuscript 2562

Data sharing statement *Data are available subject to the EPICure Data Sharing Policy (www.epicure.ac.uk) and will be available as part of the RECAP preterm Cohort Platform (https://recap-preterm.eu)*

What is already known

- Survival following birth before 26 weeks of gestation has increased steadily since 1995 in the UK
- It is unclear whether long term outcomes have changed.
- Between 1995 and 2006, survival without impairment had increased and estimated developmental test scores had improved at 3 years of age

What this study adds

- At 11 years of age in two national cohorts, the prevalence of neurodevelopmental impairment was unchanged
- Mean test scores for IQ and academic attainment were similar in both cohorts
- Improvements in survival between 1995 and 2006 were not paralleled by improved long term outcomes

ABSTRACT

Objective To determine whether improvements in school age outcomes had occurred between two cohorts of births at 22 to 25 weeks of gestation to women resident in England in 1995 and 2006.

Design Longitudinal national cohort studies.

Setting School or home-based assessments at 11 years of age.

Participants EPICure2 cohort of births at 22-26 weeks of gestation in England during 2006: a sample of 200 of 1031 survivors were evaluated; outcomes for 112 children born at 22-25 weeks' gestation were compared to those of 176 born in England during 1995 from the EPICure cohort. Classroom controls for each group acted as a reference population.

Main outcome measures Standardised measures of cognition and academic attainment were combined with parent report of other impairments to estimate overall neurodevelopmental status.

Results At 11 years in EPICure2, 18% had severe and 20% moderate impairments. Comparing births at 22-25 weeks in EPICure2 (n=112), 26% had severe and 21% moderate impairment compared to 18% and 32% respectively in EPICure. After adjustment, the odds ratio of moderate or severe neurodevelopmental impairment in 2006 compared to 1995 was 0.76, 95%CI 0.45, 1.31; p=.32. IQ scores were similar in 1995 (mean 82.7; SD 18.4) and 2006 (81.4; SD 19.2), adjusted difference in mean z-scores 0.2SD (95%CI -0.2, 0.6), as were attainment test scores. The use of multiple imputation did not alter these findings.

Conclusion Improvements in care and survival between 1995 and 2006 are not paralleled by improved cognitive or educational outcomes or a reduced rate of neurodevelopmental impairment.

(words 250/250)

Survival for babies born at extremely low gestational ages is continually increasing in the UK[1] and in many other countries.[2,3] Allied with this there is an expectation that long term outcomes will begin to improve. In our analysis of outcomes at 3 years of age for births ≤25 weeks of gestation in England, we identified that between 1995 and 2006 there was an increase in survival without impairment for liveborn babies and an 8-point (0.5 standard deviation) increase in developmental test scores, although the rates of severe impairments were unchanged.[4] Recent studies indicate that the prevalence of impaired outcomes in both early[5] and middle childhood[6] remains unchanged, and that educational and cognitive outcomes may even have deteriorated for extremely preterm/extremely low birthweight children born in 2005 compared with those born in the 1990s, despite improved survival.[6,7] Meta-analyses also indicate no improvement in cognitive outcomes for children born very preterm despite increasing survival.[8]

Predicting outcomes from infancy to middle childhood both in terms of categories of impairment[9] and absolute scores[10] remains relatively poor but is better from early school age[11,12], underscoring the need for longer term assessment. In light of the suggestion of improved outcomes in our population at 3 years[4], we report neurodevelopmental and educational outcomes at 11 years for a sample of extremely preterm (EP) births in England during 2006 (the EPICure2 Study) and compare the prevalence of impairments between this group and those of the original EPICure Study of EP births in 1995.[13] Based on the findings of our 3-year assessment, we hypothesized that we might see improved outcomes in middle childhood for EP babies born in 2006 compared with those born in 1995.

METHOD

Population The EPICure2 Study comprised births at 26 completed weeks of gestation or lower in England during the calendar year of 2006.[14] We undertook an evaluation of a sample of survivors at 11 years of age using similar methodology to that used in the 11-year follow up of the EPICure study

cohort.[13] Of 1041 EP survivors to discharge, invitations to take part in the 11-year assessment were sent to parents of 482 children admitted for care in 17 of the 45 NICUs and their networked hospitals operating in 2006[15], based on study centres in London and Leicester, respectively. The original EPICure population investigated outcomes for all children born in the UK and Ireland over 10 months in 1995. For the purposes of this comparison we restricted participants to those born to women resident in England at the time, comprising 176 of the 219 children examined at 11 years.

Study procedures Similar consent and evaluation processes were used in each study [13] (see Supplementary Appendix).

Assessment procedure IQ was measured using the Kaufman Assessment Battery for Children (1995 cohort: first edition; 2006: second edition) and academic attainment using the Wechsler Individual Achievement Test 2nd Edition (for details see Supplementary Appendix). Medical history was abstracted from parent-report questionnaires including the Gross Motor Function Classification System (GMFCS)[16] and the Manual Abilities Classification System (MACS).[17] Impairments were categorized in the same manner in both cohorts. Severe impairment comprised any one or more of an IQ score <-3 standard deviations of controls, cerebral palsy with GMFCS or MACS level 3-5, blindness or profound deafness; moderate impairment comprised any one or more of an IQ score <-2SD to -3SD of controls, GMFCS or MACS level 2, visual impairment but not registered blind or wearing hearing aids but not with profound hearing loss. Children with less severe impairments or no functional limitations were categorized with no or mild impairment. The overall level of impairment was categorized by the most severe impairment for each child. WIAT-II scores were categorized as impairments using the SD banded categories derived using contemporary controls. Socio-economic status was determined by the Index of Multiple Deprivation (IMD) derived from small area indices of income, employment, education, health, crime, housing and the living environment. IMD values for the 1995 cohort at 11 years were estimated using 2007 IMD codes (https://www.gov.uk/government/collections/english-indices-of-deprivation).

Data processing Data were entered and managed using the REDCap electronic data capture tool[18] hosted at University College London and data were transferred and combined with the main EPICure Study dataset for analysis with STATA (Version 15.1; www.stata.com). All analyses were carried out within the UCL Data Safe Haven or performed on anonymized data.

Analysis Because we had sampled the EPICure2 population, and to account for missing data, we performed regression and multiple imputation[19] as sensitivity analyses to account for selective dropouts and missing information at 11 years when estimating major outcomes in children who were not assessed by the research team. Missing data were imputed by chained equations using the STATA "MI" procedure. Imputation model variables, percentages of missing values and models for predicting missing values are listed for both cohorts in Appendix Table S5. Imputation models were based on the missing at random assumption. Although five imputed datasets may be sufficient on theoretical grounds[20], a larger number may reduce sampling variability from the imputation process.[21] Twenty imputed datasets were created for this analysis.

Study Permissions Approval was granted by both the UCL Research Ethics Committee (reference: 10175/001) and University of Leicester Research Ethics Committee (ref: 10225); UCL sponsored the study.

RESULTS

Of the 482 invitees, families of 220 agreed to participate. Due to difficulties in scheduling assessments, we evaluated 200 of the 1031 EPICure2 children known to have been alive at 3 years at a mean age of 11.8 years (sd: 0.5y). This represented a 41% response rate and the sample comprised 19% of the whole cohort; 112 were born ≤25 weeks of gestation, and therefore available to compare to outcomes from the 1995 cohort (Figure 1). From the original EPICure cohort of 219 participants evaluated at 11 years, 176 were born to women resident in England, 75% of the original population and examined at 10.9 (0.5) years.

Evaluation of dropouts Within each cohort, we compared the characteristics of dropouts with those assessed for births ≤25 weeks of gestation to women resident in England. We observed only small differences between perinatal characteristics, maternal age and IMD for the whole sample of English births ≤25 weeks of gestation and those evaluated in both cohorts (Table S1). In the 2006 cohort, the whole sample examined at 11 years, including those born at 26 weeks of gestation, was similar across a range of perinatal characteristics to the sample examined at 3 years of age.

Neurodevelopmental impairment. In EPICure2, including births at 26 weeks of gestation, 36 children (18%) had severe and 40 (20%) moderate impairment as defined above. Among the subgroup of children born before 26 weeks, 26% had severe impairment, 21% moderate and 53% no/mild impairment (Table 1; figure 2a). This compared to 18%, 32% and 49%, respectively, of births in England in EPICure. The relative risk ratio (RRR) of severe impairment in EPICure2 compared to EPICure was 1.34 (95%CI 0.73, 2.44) and for moderate impairment 0.62 (95%CI 0.35, 1.11). After adjustment for gestational age, birthweight z-score, male sex, multiple birth, maternal age and IMD at 11 years of age, RRRs were essentially unchanged. Combining the moderate and severe categories, the adjusted odds of neurodevelopmental impairment were 0.76 (96% CI 0.45, 1.31; *P*=.32).

There were no significant differences in the frequency of motor or sensory impairments between the two cohorts, but there was a reduction in the proportion of children with moderate cognitive impairment in EPICure2 compared with EPICure, before and after adjustment for baseline characteristics and multiple testing. This was offset by an increase in children with severe impairment (Table 1). Combining moderate and severe categories reduced these differences further. Considering the rates of impairment at each gestational week, there were no significant differences between the two cohorts, although there was a trend both towards higher rates of impairment (Figure 2a) and higher prevalence of milder levels of functional motor impairment (Figure 2b) at 23 weeks or below in EPICure2 (Table 2 & Table S2).

Cognitive scores Children in the control group had similar MPI scores in EPICure, using the KABC-I (mean 104; SD 11) to those in EPICure2 using the KABC-II (mean 103; SD 12; Table 2) and these were similar between the cohorts at each gestational week (Table S3). Using control group data as a contemporary reference to calculate z-scores in each cohort, there was a 1.9 SD (95% CI 1.6, 2.2) deficit for EP children versus controls in EPICure compared to a 1.8 SD (95% CI 1.4, 2.1) deficit in EPICure2; the difference in means of MPI z-scores between the two cohorts was 0.1 (95%CI -0.3, 0.5) points.

Academic attainment In contrast to MPI scores, both the EP and control groups had higher WIAT-II reading and mathematics scores in EPICure2 compared to those in EPICure. However, using z-scores calculated from cohort-specific controls as the reference, there was a 1.5 SD (95%CI 1.9, 1.2) deficit in reading scores for EP children in EPICure compared to a 1.3 SD deficit (95% CI 1.7, 0.9) in EPICure2, giving an adjusted difference between cohorts in mean z-scores of 0.2 SD (95%CI -0.2, 0.7; *P*=.3). Results were similar for mathematics with an adjusted difference of 0.3 SD (95%CI 0.0, 0.7; *P*=.06) (Table 2 & Table S4).

Overall there were no significant differences in the proportion of children with impaired reading or mathematics scores (*P*=.13 and *P*=.13, respectively) although there was a significantly reduced RRR for moderate reading impairment in EPICure2 compared to EPICure (0.32; 95%CI 0.11, 0.93) after adjustment for baseline differences.

Results after multiple imputation There were small changes in both overall impairment rates and in the prevalence of cognitive impairment (Table 3). In particular the frequency of severe impairment was increased in the 2006 cohort, but with a corresponding fall in the moderate category, such that there were no statistically significant differences in the prevalence of no or mild impairment overall.

DISCUSSION

We report outcome at 11 years of age for the EPICure2 cohort of EP births in England in 2006. Outcome for babies born before 27 weeks of gestation is still attended by a significant risk of neurodevelopmental impairments, both in terms of deficits in IQ and academic attainment, and in motor impairment or cerebral palsy.

In this analysis we compared outcomes for babies born 22-25 weeks of gestation in two cohorts of births 11 years apart. The prevalence of impairment in reading in EPICure2 appeared lower relative to those of the original EPICure study of births in 1995 but relating cognitive and educational scores to those of contemporary controls led to the conclusion that there are few differences. Therefore, we concluded that there were no significant changes in the prevalence of neurodevelopmental impairments or academic attainment over 11 years, despite over a decade of improvements in neonatal care; this has to be set alongside the significant improvements we have seen in survival and that have continued since.

Comparing cohorts separated by 11 years has its challenges. The cognitive test used in the first cohort had been superseded necessitating use of the 2nd edition of the test in the 2006 cohort. Differences in content between the two tests may have impacted on the results despite a similar theoretical construction. The same test of academic achievement was used in both cohorts making these data directly comparable, however scores had risen markedly in the 2006 cohort among both groups. This may be due to a phenomenon sometimes termed the Flynn effect[22], whereby standardized population scores drift upwards over time, making it necessary to continually reappraise standardization data to produce contemporary normed scores. To counteract bias, we included standardized educational as well as cognitive tests to triangulate the differences observed and used contemporaneous classmate controls to examine deficits in both cohorts to account for any upward drift in test scores over time. Using these methods, we found few differences across standardised cognitive, reading or mathematics scores between the two cohorts. Both samples showed significant attrition and it seems unlikely that this occurs at random. In EPICure2 at 3 years of age, there was a higher dropout of families with social disadvantage[4], as found in most studies[23 24]

but the dropout analysis for both studies shows good matching on baseline characteristics. We used

multiple imputation to account for other differences between the two populations but found

essentially similar proportions with impairment in both cohorts.

Other studies have found little change between cohorts of EP or VP births over time. Most studies

do not report outcomes in middle childhood, reporting instead infant outcomes using

developmental tests. However the Victorian Infant Collaborative Study reported outcomes from

three cohorts of EP (<28 weeks' gestation)/ELBW infants born between 1991 and 2005 at eight years

of age and concluded that moderate/severe impairment had not changed over the three cohorts, in

contrast to their infant data at 2 years.[25] Their data suggested even poorer educational outcomes,

executive functions and motor performance at 8 years of age in EP/ELBW children born in 2005

compared with those in the 1990s.[7][26] Our data do not support this finding, instead identifying no

discernible change in either IQ or academic attainment between EP children born <26 weeks of

gestation in 2006 and 1995.

Among more mature preterm infants, a meta-regression identified 71 studies of school age

outcomes for very preterm populations (<32 weeks of gestation) and found no association between

birth year and cognitive outcomes.[8] The authors suggested that the major contributory factor was

differences in BPD rates between studies. It is interesting that the rate of use of supplemental

oxygen at 36 weeks, the most commonly used definition for "BPD" was very similar in both the

EPICure cohorts.[4]

We would have greater confidence if these national studies had achieved higher follow up rates.

Following large cohorts of children scattered across the country is expensive and fraught with

difficulty. Routine clinical follow up to two years remains patchy in the UK[27], despite being

recommended for many years. Recent guidance recommends that routine outcome evaluation in

the UK should include an IQ test at preschool age to provide a more robust measure for those born

EP.[28] An interesting option is to use nationally collected educational data which are considerably

more consistent and complete and which have been shown to be associated with gestational age.[29]

These data may provide the opportunity to evaluate changing patterns of outcome robustly, as long as evaluation criteria remain the same.

Difficulties in study retention and corresponding small sample sizes also make it difficult to explore interactions and sub-group analyses so it is possible that improvements are emerging among certain groups, but this is not affecting overall outcome. For example, it is possible that improvements in outcome have occurred among those from higher SES backgrounds with greater family resources, but such analyses are not possible in smaller datasets.

Despite major advances in improved organization and delivery of neonatal care between 1995 and 2006, and in survival, this has not yet been paralleled by improved investment in developmental follow-up or in educational support.[30] If improving neonatal care is not resulting in the gains in neurodevelopmental outcomes once expected, then perhaps it is time to look at improving the provision of longer term developmental, parent and educational support[31][32] as interventions to optimise long term outcomes.

2531 words

REFERENCES

1. Mactier H, Bates SE, Johnston T, et al. Perinatal management of extreme preterm birth before 27 weeks of gestation: a framework for practice. *Archives of Disease in Childhood Fetal and Neonatal Edition* 2020;105(3):232-39.

2. Norman M, Hallberg B, Abrahamsson T, et al. Association Between Year of Birth and 1-Year Survival Among Extremely Preterm Infants in Sweden During 2004-2007 and 2014-2016. *JAMA* 2019;321(12):1188-99.

3. Torchin H, Morgan A, Ancel P-Y. International comparisons of neurodevelopmental outcomes in infants born very preterm. *Seminars in Fetal and Neonatal Medicine* 2020;in press

4. Moore T, Hennessy EM, Myles J, et al. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. *BMJ* 2012;345:e7961. doi: 10.1136/bmj.e7961

5. Ancel PY, Goffinet F, Group E-W, et al. Survival and Morbidity of Preterm Children Born at 22 Through 34 Weeks' Gestation in France in 2011: Results of the EPIPAGE-2 Cohort Study. *JAMA Pediatrics* 2015;169(3):230-8.

6. Cheong J, Spittle A, Burnett A, et al. Have outcomes following extremely preterm birth improved over time? *Seminars in Fetal and Neonatal Medicine* 2020;in press

7. Burnett AC, Anderson PJ, Lee KJ, et al. Trends in Executive Functioning in Extremely Preterm Children Across 3 Birth Eras. *Pediatrics* 2018;141(1) doi: 10.1542/peds.2017-1958

8. Twilhaar ES, de Kieviet JF, Aarnoudse-Moens CS, et al. Academic performance of children born preterm: a meta-analysis and meta-regression. *Archives of Disease in Childhood Fetal and Neonatal Edition* 2018;103(4):F322-F30.

9. Marlow N, Wolke D, Bracewell M, et al. Neurologic and developmental disability at six years of age after extremely preterm birth. *The New England journal of medicine* 2005;352(1):9-19.

10. Linsell L, Johnson S, Wolke D, et al. Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: a prospective, population-based cohort study. *Archives of Disease in Childhood* 2018;103(4):363-70.

11. Breeman LD, Jaekel J, Baumann N, et al. Preterm Cognitive Function Into Adulthood. *Pediatrics* 2015;136(3):415-23.

12. Mangin KS, Horwood LJ, Woodward LJ. Cognitive Development Trajectories of Very Preterm and Typically Developing Children. *Child Development* 2017;88(1):282-98.

13. Johnson S, Fawke J, Hennessy E, et al. Neurodevelopmental disability through 11 years of age in children born before 26 weeks of gestation. *Pediatrics* 2009;124(2):e249-57. doi: 10.1542/peds.2008-3743

14. Costeloe KL, Hennessy EM, Haider S, et al. 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. doi: 10.1136/bmj.e7976

15. Marlow N, Bryan Gill A. Establishing neonatal networks: the reality. *Archives of Disease in Childhood Fetal and Neonatal Edition* 2007;92(2):F137-42.

16. Palisano RJ, Hanna SE, Rosenbaum PL, et al. Validation of a model of gross motor function for children with cerebral palsy. *Physical Therapy* 2000;80(10):974-85.

17. Eliasson AC, Krumlinde-Sundholm L, Rosblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Developmental Medicine and Child Neurology* 2006;48(7):549-54.

18. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377-81.

19. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.

20. Allison PD. Multiple imputation for missing data: A cautionary tale. *Sociological Methods & Research* 2000;28:301-9.

21. Horton NJ, Lipsitz SR. Multiple imputation in practice: comparison of software packages for regression models with missing variables. *The American Statistician* 2001;55:244-54.

22. Flynn J. Searching for justice, The discovery of IQ gains over time. *American Psychlogist* 1999;54:5-20.

23. Tin W, Wariyar UK, Hey EN. Selection biases invalidate current low birthweight weight-for-gestation standards. The Northern Neonatal Network. *British Journal of Obstetrics and Gynaecology* 1997;104(2):180-5.

24. Wolke D, Waylen A, Samara M, et al. Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *The British Journal of Psychiatry* 2009;195(3):249-56.

25. Cheong JLY, Anderson PJ, Burnett AC, et al. Changing Neurodevelopment at 8 Years in Children Born Extremely Preterm Since the 1990s. *Pediatrics* 2017;139(6) doi: 10.1542/peds.2016-4086

26. Spittle AJ, Cameron K, Doyle LW, et al. Motor Impairment Trends in Extremely Preterm Children: 1991-2005. *Pediatrics* 2018;141(4) doi: 10.1542/peds.2017-3410

27. National Neonatal Audit Programme (NNAP) 2019 annual report on 2018 data. London: RCPCH, 2019.

28. National Institute for Health and Care Excellence. Developmental follow-up of children and young people born preterm. London: NICE 2019.

29. MacKay DF, Smith GC, Dobbie R, et al. Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren. *PLoS medicine* 2010;7(6):e1000289.

30. Johnson S, Gilmore C, Gallimore I, et al. The long-term consequences of preterm birth: what do teachers know? *Developmental Medicine and Child Neurology* 2015;57(6):571-7.

31. Wolke D, Johnson SJ, Mendonca M. The Life Course Consequences of Very Preterm Birth. *Annual Review of Developmental Psychology* 2019;1:66-92.

32. Wolke D, Jaekel J, Hall J, et al. Effects of sensitive parenting on the academic resilience of very preterm and very low birth weight adolescents. *The Journal of Adolescent Health* 2013;53(5):642-7.

Legends for Figures

Figure 1: Derivation of the two samples of extremely preterm children at 11 years born in England in 1995 and 2006

Figure 2 – Prevalence of categorization of overall impairment (a) and Gross Motor Function Classification System levels (b) by gestational age at birth at 11 years in the two EPICure cohorts born 22-25 weeks of gestation in 1995 and 2006, respectively.

Table 1: Categories of impairments among EPICure (1995) and EPICure2 (2006) cohorts of births at 22-25 weeks gestation in England at 11 years of age

	EPICure 1995 N=176	EPICure2 2006 N=112	2006 vs 1995 Overall * P	Individual categories			
				Unadjusted RRR (95% CI)	P	Adjusted* RRR (95% CI)	P
Overall Impairment				N=288		N=283	
No/mild	49% (87/176)	53% (59/112)	0.85	-	-	-	-
Moderate	32% (57/176)	21% (24/112)		0.62 (0.35,1.11)	0.11	0.58 (0.31,1.08)	0.08
Severe	18% (32/176)	26% (29/112)		1.34 (0.73,2.44)	0.35	1.14 (0.57,2.27)	0.71
Combined mod/severe	51% (89/176)	47% (53/112)		0.88 (0.55,1.41)	0.59	0.76 (0.45,1.31)	0.32
Motor function							
No CP or GMFCS /MACS 1	81% (143/176)	80% (90/112)	0.96	-	-	-	-
GMFCS /MACS = 2	8.0% (14/176)	14% (16/112)		1.82 (0.85,3.90)	0.13	1.65 (0.71,3.82)	0.24
GMFCS /MACS 3-5	11% (19/176)	5% (6/112)		0.50 (0.19,1.30)	0.16	0.41 (0.14,1.16)	0.09
Combined mod/severe	19% (33/176)	20% (22/112)		1.06 (0.58,1.93)	0.85	0.93 (0.48,1.82)	0.84
Cognitive Impairment							
None ($\geq$-1SD)	30% (53/176)	37% (41/112)	0.80	-	-	-	-
Mild (<-1 to -2 SD)	29% (51/176)	26% (29/112)		0.74 (0.40,1.35)	0.32	0.78 (0.40,1.48)	0.44
Moderate (<-2 to -3SD)	26% (45/176)	13% (14/112)		0.40 (0.19,0.83)	0.01	0.32 (0.15,0.70)	0.004
Severe (<-3SD)	15% (27/176)	25% (28/112)		1.34 (0.69,2.61)	0.39	1.16 (0.54,2.49)	0.70
Combined mod/severe	41% (72/176)	38% (42/112)		0.87 (0.53,1.41)	0.56	0.69 (0.40,1.18)	0.18
Reading Impairment				N=279		N=276	
None (>-1SD)	48% (82/171)	56% (60/108)	0.13	-	-	-	-
Mild (<-1 to -2 SD)	19% (33/171)	23% (25/108)		1.04 (0.56,1.92)	0.91	1.01 (0.53,1.92)	0.99
Moderate (<-2 to -3SD)	12% (21/171)	5% (5/108)		0.33 (0.12,0.91)	0.03	0.32 (0.11,0.93)	0.04
Severe (<-3SD)	21% (35/171)	17% (18/108)		0.70 (0.36,1.36)	0.29	0.62 (0.30,1.27)	0.19
Combined mod/severe	33% (56/171)	21% (23/108)		0.56 (0.32,0.97)	0.04	0.51 (0.28,0.93)	0.03
Maths Impairment				N=283		N=278	
None (>-1SD)	28% (49/173)	36% (39/110)	0.13	-	-		
Mild (<-1 to -2 SD)	27% (46/173)	29% (32/110)		0.87 (0.47,1.62)	0.67	0.86 (0.45,1.65)	0.65
Moderate (<-2 to -3SD)	19% (33/173)	14% (15/110)		0.57 (0.27,1.20)	0.14	0.59 (0.27,1.29)	0.18
Severe (<-3SD)	26% (45/173)	22% (24/110)		0.67 (0.35,1.28)	0.23	0.57 (0.28,1.16)	0.12
Combined mod/severe	45% (78/173)	36% (39/110)		0.67 (0.41,1.09)	0.11	0.62 (0.36,1.07)	0.09

Key: CP: cerebral palsy; GMFCS: Gross Motor Function Classification system; MACS: Manual Abilities Classification System; *Wilcoxon Rank Sum Test

Table 2: Mean scores for IQ and academic attainment for the two EPICure cohorts at 11 years of births in England

Kaufmann ABC scores	EPICure 1995 22-25 weeks		EPICure2 2006 22-25 weeks		[c] v [a]		[a] v [b]		[c] v [d]		EPICure2 2006 22-26 weeks
	EPICure [a] (n=176) Mean (SD)	Comparison [b] (n=153) Mean (SD)	EPICure2 [c] (n=112) Mean (SD)	Comparison [d] (n=143) Mean (SD)	Δ means (95%CI)	Adjusted Δ * (95%CI)	Δ means (95%CI)	Adjusted Δ $^{\varphi}$ (95%CI)	Δ means (95%CI)	Adjusted Δ$^{\varphi}$ (95%CI)	(n=200) Mean (SD)
Mental processing index	83 (18) [n=174]	104 (11) [n=153]	81 (19) [n=112]	103 (12) [n=143]	-1.3 (-5.8, 3.2)	-0.7 (-5.0, 3.6)	-21 (-25, -18) ‡	-21 (-25, -17) **	-22 (-26, -18) ‡	-21 (-25, -17) ‡	85 (19) [n=200]
MPI z-score ¥	-1.9 (1.7)		-1.8 (1.6)		0.1 (-0.3, 0.5)	0.2 (-0.2, 0.6)	-1.9 (-2.3, -1.6) ‡	-1.9 (-2.2, -1.6) ‡	-1.8 (-2.2, -1.5) ‡	-1.8 (-2.1, -1.4) ‡	-1.5 (1.6) [n=200]
WIAT-II Composite Scores											
Reading	79 (21) [n=171]	99 (12) [n=153]	88 (22) [n=108]	103 (11) [n=140]	9.1 (3.9, 14) †	9.3 (4.2,14) ‡	-19 (-23, -16) ‡	-18 (-22, -14) ‡	-15 (-19, -11) ‡	-14 (-19, -10) ‡	92 (21) [n=195]
Reading z-score ¥	-1.6 (1.7)		-1.3 (2.0)		0.2 (-0.2, 0.7)	0.2 (-0.2, 0.7)	-1.6 (-1.9, -1.3) ‡	-1.5 (-1.9, -1.2) ‡	-1.4 (-1.8, -1.0) ‡	-1.3 (-1.7, -0.9) ‡	-1.0 (1.9) [n=195]
Mathematics	70 (21) [n=173]	99 (15) [n=153]	83 (28) [n=110]	109 (16) [n=139]	13 (7.2, 18) ‡	14 (7.9, 19) ‡	-28 (-32, -24) ‡	-27 (-32, -23) ‡	-26 (-32, -21) ‡	-25 (-31, -20) ‡	87 (27) [n=197]
Maths z-score ¥	-1.9 (1.4)		-1.6 (1.7)		0.3 (-0.1, 0.7)	0.3 (-0.0, 0.7)	-1.9 (-2.1, -1.6) ‡	-1.8 (-2.1, -1.5) ‡	-1.6 (-2.0, -1.3) ‡	-1.6 (-1.9, -1.2) ‡	-1.4 (1.7) [n=197]

* Multiple linear regression models were used to adjust for gestation, BW z score, male sex, multiple birth, maternal age and IMD at 11y.

$^{\varphi}$ Multiple linear regression models were used to adjust for male sex and IMD at 11y.

† P=.001; ‡ P<0.001

Δ Difference in means

¥ Z scores calculated using control group scores as reference to adjust for temporal changes

Table 3: Comparison of outcomes at 11 years for babies born at 22-25 weeks' gestation in England in the two EPICure cohorts: results of imputation and a sensitivity analysis including maternal age at delivery in the imputation model (see Table 8 (Appendix) for variables and missingness).

	Original results		Imputed results[¶]		
	1995	**2006**	**1995**	**2006**	**2006***
	N=176	N=112	N=260	N=584	N=584
Overall disability	% (95%CI)	% (95%CI)	% (95%CI)	% (95%CI)	% (95%CI)
No/mild	49% (42, 57)	53% (43, 62)	45% (38, 52)	45% (39, 51)	44% (37, 50)
Moderate	32% (26, 40)	21% (15, 30)	34% (27, 41)	23% (18, 28)	24% (15, 33)
Severe	18% (13, 25)	26% (19, 35)	21% (15, 27)	32% (27, 37)	33% (24, 41)
Cognitive disability					
No/mild	59% (52, 66)	63% (53, 71)	53% (47, 60)	55% (48, 62)	55% (46, 64)
Moderate	26% (20, 33)	13% (8, 20)	28% (21, 34)	17% (10, 24)	17% (9, 24)
Severe	15% (11, 22)	25% (18, 34)	19% (13, 25)	28% (21, 35)	29% (22, 36)

*¶ 20 imputed datasets were used. *After adding maternal age at delivery in the imputation models*