

# Effects of repeated prenatal ultrasound examinations on childhood outcome up to 8 years of age: follow-up of a randomised controlled trial

*Lancet* 2004; 364: 2038–44 John P Newnham, Dorota A Doherty, Garth E Kendall, Stephen R Zubrick, Louis L Landau, Fiona J Stanley

School of Women's and Infants' Health, The University of Western Australia at King Edward Memorial Hospital, Subiaco, Western Australia 6008 (J P Newnham MD, D A Doherty PhD); TVW Telethon Institute for Child Health Research (G E Kendall PhD; F J Stanley MD); Curtin Centre for Developmental Health, Institute for Child Health Research (S R Zubrick PhD); and Faculty of Medicine and Dentistry, The University of Western Australia, Perth, Western Australia (L L Landau MD).

Correspondence to: Prof John P Newnham, School of Women's and Infants' Health, King Edward Memorial Hospital, 374 Bagot Road, Subiaco, Western Australia 6008. [jnewnham@obs gyn.uwa.edu.au](mailto:jnewnham@obs gyn.uwa.edu.au)

## Summary

**Background** Despite the widespread use of prenatal ultrasound studies, there are no published data from randomised controlled trials describing childhood outcomes that might be influenced by repeated ultrasound exposures. We previously undertook a randomised controlled trial to assess the effects of multiple studies on pregnancy and childhood outcomes and reported that those pregnancies allocated to receive multiple examinations had an unexplained and significant increase in the proportion of growth restricted newborns. Our aim was to investigate the possible effects of multiple prenatal ultrasound scans on growth and development in childhood. Here, we provide follow-up data of the children's development.

**Methods** Physical and developmental assessments were done on children whose pregnant mothers had been allocated at random to a protocol of five studies of ultrasound imaging and umbilical artery Doppler flow velocity waveform between 18 and 38 weeks' gestation (intensive group n=1490) or a single imaging study at 18 weeks' gestation (regular group n=1477). We used generalised logistic and linear regression models to assess the group differences in developmental and growth outcomes over time. Primary data analysis was done by intention-to-treat.

**Findings** Examinations were done at 1, 2, 3, 5, and 8 years of age on children born without congenital abnormalities and from singleton pregnancies (intensive group n=1362, regular group n=1352). The follow-up rate at 1 year was 85% (2310/2714) and at 8 years was 75% (2042/2714). By 1 year of age and thereafter, physical sizes were similar in the two groups. There were no significant differences indicating deleterious effects of multiple ultrasound studies at any age as measured by standard tests of childhood speech, language, behaviour, and neurological development.

**Interpretation** Exposure to multiple prenatal ultrasound examinations from 18 weeks' gestation onwards might be associated with a small effect on fetal growth but is followed in childhood by growth and measures of developmental outcome similar to those in children who had received a single prenatal scan.

## Introduction

Over the past three decades, the application of ultrasound technology to the clinical practice of obstetrics has achieved widespread acceptance. Ultrasound imaging and flow studies are used routinely in different gestational ages, and the benefits of this technology are welcomed both by health care professionals and the general community. However, there also is an acceptance that ultrasound examination of the developing embryo or fetus is without hazard, but the evidence that underpins this acceptance is far from comprehensive. Further, the widespread use of these tests in pregnancies in which there is little expectation of pathological findings increases our need to ensure that the benefits outweigh any potential for harm.

To our knowledge, there has never been a randomised controlled trial done for the specific purpose of investigating potential bio-effects of prenatal ultrasound in humans.<sup>1</sup> In the absence of such evidence, data from studies done for other purposes assume importance. Studies that have been assessed for this purpose include randomised controlled trials designed to investigate effectiveness of the various diagnostic tests, case-control studies, and population cohorts. These studies have

provided reasonable evidence that one or two scans do not cause any complications of pregnancy, growth restriction, or childhood malignancy.<sup>2–4</sup> Debate continues about findings that suggest a possible increase in non-right handedness in boys, resulting in lingering uncertainties about potential effects on the developing brain.<sup>5–8</sup>

In 1993, we published the results of a randomised controlled trial in which pregnant women were allocated at random to a protocol of five ultrasound imaging and Doppler blood flow studies from 18 to 38 weeks' gestational age, or to a single imaging scan at 18 weeks with further scans only as indicated on clinical grounds.<sup>9–11</sup> The purpose of the study was to investigate potential benefits in terms of pregnancy outcome. There were no measurable improvements in any maternal, fetal, or newborn outcome, indicating that a routine policy of multiple prenatal scans from mid-pregnancy onwards was not of benefit. However, an unexplained increase in the proportion of growth-restricted offspring was seen in the repeated ultrasound group, unaccompanied by any apparent morbidity. In view of long-standing reports from animal studies suggesting a possible link between prenatal ultrasound exposure and fetal growth,<sup>12,13</sup> our findings were judged to warrant

ongoing investigation. In this study, we compare the growth and development of the children in each of the two prenatal study groups from 1–8 years of age.

## Participants and methods

### Enrolment

The children in this study and their pregnant mothers had been enrolled into a randomised controlled trial designed to investigate whether intensive use of ultrasound imaging and Doppler flow studies would improve pregnancy outcome expressed as days of neonatal stay and the rate of pre-term birth.<sup>9</sup> Recruitment was at King Edward Memorial Hospital which is the tertiary level perinatal centre for the state of Western Australia and was undertaken between May, 1989 and November, 1991. The criteria for enrolment were gestational age between 16 and 20 weeks, proficiency in English sufficient to understand the implications of participation, an expectation to deliver at the hospital, and an intention to remain in Western Australia in the coming years such that long-term follow-up was feasible. The study was approved by the institutional ethics committee, and written consent was obtained from every pregnant woman. There were 2834 singleton pregnancies resulting in 2743 surviving newborn children available for follow-up. Of these, 2714 were included in our analysis since 29 were excluded because of congenital anomalies.

A questionnaire was completed by every woman at the time of enrolment with assistance from a research midwife. The questionnaire contained 108 questions about the woman's social and economic circumstances, lifestyle, medical history, and environmental exposures. A second questionnaire was given to the woman to be completed by her partner, containing 30 questions about his physical size, education, occupation, and environmental exposures.

### Randomisation

After enrolment, women were allocated to a study group by draw of a sealed envelope prepared in blocks of 20 by computer-generated random numbers. Individuals allocated to the intensive ultrasound group had imaging and Doppler flow studies at about 18 weeks' gestation, and then at 24, 28, 34, and 38 weeks' gestation. Those allocated to the regular group had an imaging study at about 18 weeks' gestation with further scans only at the request of her attending clinician. Of the 2714 included in our analysis, 1352 of these children had previously been randomised to the regular ultrasound group and 1362 to the intensive group. Results of all studies were revealed to the women and their health-care providers. A further questionnaire was completed by every woman at 34 weeks' gestation to ascertain information on any new exposures during the pregnancy. Details of the women and their pregnancies in each group have been reported previously.<sup>9</sup>

### Pregnancy outcomes

Demographic and clinical outcomes were recorded from the questionnaires and medical records, with classifications determined for obstetric intervention, smoking practices, and socioeconomic conditions.

### Ultrasound studies

Ultrasound imaging studies were done with one of two General Electric 3600 machines (Milwaukee, USA) with 3.5 MHz linear array and 5 MHz sector transducers. At all examinations, measurements were made of the biparietal and occipito-frontal diameters, head and abdominal circumferences, and femur length. Gestational age was determined at the 16 to 18 week examination, and was based on the date of the last menstrual period if the age by that calculation did not differ from the estimation by biometry by more than 7 days, and by the biometry in the remaining cases. Women in both groups received an ultrasound examination at 18 weeks' gestation to ensure that determination of gestational age was identical in the two groups. Imaging studies at 18 weeks' gestation included full assessment of fetal anatomy. Doppler flow velocity waveform studies were done with a spectrum analyser (Medasonics SP25A; Mountain View, California) and a D10 bi-directional continuous wave Doppler system (total power output 3 mW, spatial peak temporal average 25 mW/cm<sup>2</sup>). The waveforms were obtained from an umbilical artery and an arcuate artery within the placental bed, located by ultrasound imaging and recognition of the characteristic audible signals.

### Childhood follow-up

All children were invited for follow-up studies at 1, 2, 3, 5, and 8 years of age. Physical measures were made of weight; height; circumferences of the head, chest, and mid-arm; and skin-fold thicknesses from the triceps, sub-scapular, supra-iliac, and abdominal regions. Methods used to assess development are shown in the panel.

### Statistical analysis

Losses to follow-up were examined for all singleton pregnancies (regular ultrasound group, n=1352; intensive ultrasound group, n=1362). Baseline summary statistics and group comparisons were restricted to the children seen on at least one occasion during follow-up (regular ultrasound group, n=1233; intensive ultrasound group, n=1272) with appropriate sample sizes specified for the individual analyses.

Except for PPTV-R, developmental outcomes were analysed with logistic regression that modelled the probability of scoring outside the reference range for a development test. A generalised estimating equations (GEE) approach was used to assess the developmental outcomes collected over time, such as IMQ and CBCL, and when all available data were included in the analysis.

**Panel: Assessments to measure child development****Modified toddler temperament scale (TTS)**

A 97-item, self-administered parent questionnaire that measures infant temperament on nine dimensions: activity, rhythmicity, approach, adaptability, intensity, mood, persistence, distractibility, and threshold.<sup>14</sup> The TTS was administered at 1 year of age.

**Early language milestone scale—second edition (ELM scale-2)**

Designed to screen children up to 3 years of age for speech and language delay on the basis of assessments in the auditory expressive, auditory receptive, and visual domains, producing a global language score.<sup>15</sup> The ELM Scale-2 was administered at 1 year of age. At this age, the scale is principally a test of receptive language performance, and a child is deemed to have delayed language development if he or she is unable to make isolated sounds, label the appropriate parent, follow simple verbal commands, or point to desired objects.

**Infant/child monitoring questionnaires (IMQ)**

A series of parent-completed questionnaires about motor, communication, social/personal, and adaptive development.<sup>16</sup> The IMQ was administered at 1, 2, and 3 years of age.

**Denver developmental screening test—second edition (Denver II)**

An assessment designed to be administered in a clinical setting by a variety of professionals to screen pre-school children for possible developmental problems in the areas of personal-social, fine motor, gross motor, and language functioning.<sup>17</sup> Several test materials are needed, such as a small bell, a tennis ball, and a small plastic doll. The test was administered at 1 and 3 years of age.

**Language development survey (LDS)**

Identifies expressive language delay by the parent identifying words that the child uses spontaneously.<sup>18</sup> The LDS was administered at 2 years of age.

**Child behaviour checklist (CBCL)**

A multi-item instrument that assesses children according to eight syndrome constructs: withdrawn, anxious/depressed, somatic complaints, social problems, attention problems, thought problems, delinquent behaviour, and aggressive behaviour.<sup>19</sup> The instrument was used by a principal caregiver at 2, 5, and 8 years of age.

**Peabody picture vocabulary test—revised (PPVT-R)**

An achievement test of receptive vocabulary in English based on recognition of drawings that best symbolise words spoken by an examiner.<sup>20</sup> The PPVT-R was administered at 5 years of age.

Odds ratios for the intensive group relative to the regular group were used to summarise the possible effect of group allocation on the developmental outcomes and

	Groups		p*
	Regular ultrasound group (n=1352)	Intensive ultrasound group (n=1362)	
<b>Year 1</b>			
Q and E	1014 (75%)	955 (70%)	..
Questionnaires	41 (3%)	41 (3%)	..
Physical exam	76 (6%)	183 (13%)	..
Not seen	221 (16%)	183 (13%)	0.033
<b>Year 2</b>			
Q and E	301 (22%)	321 (24%)	..
Questionnaires	585 (43%)	620 (46%)	..
Physical exam	31 (2%)	26 (2%)	..
Not seen	435 (32%)	395 (29%)	0.073
<b>Year 3</b>			
Q and E	687 (51%)	785 (58%)	..
Questionnaires	303 (22%)	268 (20%)	..
Physical exam	49 (4%)	48 (4%)	..
Not seen	313 (23%)	261 (19%)	0.011
<b>Year 3</b>			
Q and E	921 (68%)	968 (71%)	..
Questionnaires	97 (7%)	103 (8%)	..
Physical exam	16 (1%)	19 (1%)	..
Not seen	318 (24%)	272 (20%)	0.025
<b>Year 8</b>			
Q and E	884 (65%)	953 (70%)	..
Questionnaires	90 (7%)	88 (7%)	..
Physical exam	16 (1%)	11 (1%)	..
Not seen	362 (27%)	310 (23%)	0.015
<b>Total</b>			
Not seen	119 (9%)	90 (7%)	0.032

Q=questionnaire. E=physical examination. \*The p values were obtained with  $\chi^2$  test for association between the group assignment and dichotomous outcome of every child being seen on at least one occasion at each time point.

**Table 1: Summary of available information from questionnaires completed by parents, and from attendance for physical examination at scheduled visits**

were reported together with their 95% CIs. Group differences in PPTV-R were assessed with linear regression analysis. We also used linear regression modelling to assess growth differences between the groups over time with individual children modelled as random effects. Box-Cox transformations were used to achieve normality (the descriptive statistics were based on non-parametric summaries using medians and IQRs shown as Q1–Q3, and preliminary univariate tests implemented non-parametric inference). Primary data analysis was done on the intention-to-treat principle, and secondary analyses also compared the developmental and growth outcomes according to exposure to ultrasound during pregnancy. Secondary analyses according to exposure to ultrasound compared outcomes after a single ultrasound versus multiple ultrasound scans. Results of this analysis were analogous to those obtained when using the intention-to-treat approach and are not shown. Sensitivity analyses were done to assess the effects of missing data on group comparisons. Worst case and best case assessments were done for all outcomes that were considered for analysis and no adverse effects of repeated ultrasound examinations were seen when compared with the outcomes in the regular ultrasound group. SAS (version 8.2, SAS

	Girls			Boys		
	Regular group, median (IQR)	Intensive group, median (IQR)	p	Regular group, median (IQR)	Intensive group, median (IQR)	p
<b>Year 1</b>						
Weight (kg)	9.9 (9.2–10.7)	9.9 (9.1–10.8)	0.955	10.5 (9.8–11.4)	10.6 (9.8–11.5)	0.284
Height (cm)	77.0 (75.1–78.8)	77.0 (74.8–78.6)	0.467	78.3 (76.6–80.3)	78.2 (76.0–80.2)	0.270
Head circumference (cm)	48.6 (46.0–47.5)	46.7 (45.9–47.5)	0.458	48.1 (47.2–49.0)	48.0 (48.0–49.0)	0.826
<b>Year 2</b>						
Weight (kg)	12.5 (11.7–13.5)	12.3 (11.4–13.4)	0.469	13.1 (12.4–14.2)	13.4 (12.1–14.2)	0.453
Height (cm)	89.2 (86.5–91.8)	89.5 (87.0–91.3)	0.270	91.0 (88.0–93.0)	91.0 (88.6–92.9)	0.842
Head circumference (cm)	49.0 (48.0–50.0)	49.0 (48.1–50.0)	0.336	50.5 (49.5–51.0)	50.0 (49.0–51.0)	0.202
<b>Year 3</b>						
Weight (kg)	14.5 (13.5–15.6)	14.6 (13.5–15.9)	0.462	15.3 (14.1–16.3)	15.3 (14.1–16.5)	0.254
Height (cm)	95.3 (93.0–97.6)	95.9 (93.7–98.7)	0.197	97.4 (94.7–99.9)	96.7 (94.3–99.3)	0.610
Head circumference (cm)	50.2 (49.4–51.0)	50.2 (49.2–51.2)	0.730	51.5 (50.5–52.4)	51.5 (50.5–52.4)	0.673
<b>Year 5</b>						
Weight (kg)	20.7 (18.8–22.9)	20.7 (18.9–23.0)	0.889	21.3 (19.5–23.0)	21.2 (19.4–23.3)	0.450
Height (cm)	114.4 (112.2–118.3)	115.5 (112.4–118.5)	0.882	116.8 (113.6–120.0)	116.6 (113.2–120.0)	0.807
Head circumference (cm)	51.7 (50.7–52.5)	51.8 (50.7–52.5)	0.881	52.9 (51.9–53.7)	52.9 (51.9–53.9)	0.318
<b>Year 8</b>						
Weight (kg)	26.9 (24.0–30.9)	27.0 (24.3–30.5)	0.632	27.6 (25.1–30.0)	27.5 (24.4–30.7)	0.781
Height (cm)	128.6 (124.3–132.0)	128.6 (124.3–132.0)	0.955	130.2 (126.8–133.9)	129.7 (125.0–133.6)	0.068
Head circumference (cm)	52.5 (51.6–53.5)	52.5 (51.5–53.5)	0.775	53.6 (52.7–54.5)	53.6 (52.6–54.6)	0.683

p values were obtained in pairwise comparisons between the ultrasound groups using a longitudinal growth model adjusted for statistically significant covariates.

**Table 2: Measurements of physical size at every age assessed, by time**

Institute, Cary NC, USA) and SPSS (version 11.0, SPSS, Chicago IL, USA) statistical software were used for data analysis. p values less than 0.05 were judged significant.

### Role of the funding source

The study sponsors had no role in study design; in the collection, analysis or interpretation of the data; in the writing of the report; or in the decision to submit the

paper for publication. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

### Results

Higher proportions of children in the intensive ultrasound group were seen at each of the scheduled follow-up visits (table 1). The percentages who did not

	Regular ultrasound group		Intensive ultrasound group		p
	n	All	n	All	
<b>TTS</b>					
Year 1	126 (12%)	1055	143 (13%)	1112	0.518
<b>ELM</b>					
Year 1	23 (2%)	959	12 (1%)	1021	0.039
<b>IMQ</b>					
Year 1	78 (7%)	1076	97 (9%)	1123	0.229
Year 2	130 (15%)	893	152 (16%)	933	0.305
Year 3	60 (6%)	932	75 (8%)	983	0.308
<b>Denver*</b>					
Year 1	130 (12%)	1036	141 (13%)	1086	0.764
Year 3	126 (22%)	559	177 (26%)	668	0.109
<b>LDS</b>					
Year 2					
>50 words	76 (10%)	762	71 (9%)	801	0.452
Combining words	60 (8%)	733	63 (8%)	793	0.863
<b>CBCL</b>					
Year 2	101 (11%)	928	117 (12%)	972	0.430
Year 5	201 (20%)	1005	232 (22%)	1058	0.282
Year 8	183 (19%)	964	199 (19%)	1019	0.758
<b>PPVT—R†</b>					
Year 5	106.4 (13.5)	844	105.5 (13.9)	892	0.054

TTS=toddler temperamental scale. ELM=early language milestone. IMQ=infant/child monitoring questionnaire. Denver=Denver development screening test. LDS=language development survey. CBCL=child behaviour checklist. PPVT-R=Peabody picture vocabulary test-revised. \*Year 3 follow-up comprised 1524 of 2021 (75.4% of a 70% sample) examinations; this sample included children considered at high risk as neonates (low birthweight, preterm birth, and other factors) and a random sample of other children; †Data are mean (SD). Only measurements from satisfactory childhood examinations are included, and scores outside of normal range are compared between the study groups.

**Table 3: Developmental measures administered up to 8 years of age**

attend or did not complete questionnaires were 16% (221/1352) and 13% (183/1362) at 1 year of age ( $p=0.033$ ) and 27% (362/1352) and 23% (310/1362) at 8 years of age in the regular and intensive groups, respectively ( $p=0.015$ ). Overall, 209 children were lost to follow-up during the 8 years (119 in the regular group and 90 in the intensive group;  $p=0.032$ ).

At enrolment, the two groups did not differ significantly in terms of parental racial origin, education level, smoking practices, or financial income. Distribution of sexes was equivalent between the groups with 623 (50.1%) and 655 (52.3%) boys followed-up in the regular and intensive ultrasound groups, respectively ( $p=0.38$ ). The mean ages of children at the scheduled visits were similar in both groups.

At birth, the length of newborns in the intensive ultrasound group was significantly less than that in the regular ultrasound group ( $p=0.011$ ). This difference was due to shorter lengths in the newborns in the intensive group between the 25th and 45th percentiles (difference of 0.5 cm). The weights and circumferences of the head of newborns were similar between the groups ( $p=0.37$  with medians of 3370 g [IQR 3070–3680] and 3380 g [3010–3685] in the regular and intensive ultrasound groups, respectively;  $p=0.056$  with medians 34.5 cm [33.5–35.5] for both groups). Heights, weights, and circumferences of the head at each age stratified by sex are summarised in table 2. Weights, circumferences of the head, chest, and mid-arm, and all skin-fold thicknesses did not differ significantly between the two groups at any age, either in univariate or multivariate analyses with adjustments for other covariates such as sex or prematurity.

There were no significant differences between the two groups in any of the measures of childhood development

at 1, 2, 3, 5, or 8 years of age, except ELM at 1 year ( $p=0.039$ ) (table 3). This difference in ELM at 1 year of age resulted from abnormal scores in 12 of the 1021 in the intensive ultrasound group and 23 of 959 in the regular ultrasound group. The significant difference remained after adjustment for sex, gestational age at birth, and the proportions with birthweight less than the 10th percentile (adjusted odds ratio 0.47, 95% CIs 0.23–0.96) (table 4).

## Discussion

A protocol of five ultrasound imaging and Doppler blood flow studies, when compared with a protocol of a single imaging study with further scans reserved for clinical indications, resulted in similar growth and development up to 8 years of age. The overall absence of any deleterious effects of prenatal scans on childhood development is consistent with results from other studies, although previous investigations were not based on the use of multiple scans.<sup>1–4</sup> In our study, the only significant difference in developmental outcomes in the group that had received multiple scans was a reduction in the proportion of infants at 1 year of age with abnormal scores on the ELM scale. This instrument provides an assessment of speech and language delay. At later ages, all tests of this aspect of development were similar in the two groups. The many end-points tested in this study render it likely that this single significant finding resulted from chance. We suggest that there is no evidence to support the notion that prenatal scans could improve language acquisition in infancy, but it is possible that the finding resulted from altered parenting in the intensive ultrasound group. Children in that arm of the study were more likely to attend for childhood follow-up examinations throughout the 8-year study period. The possibility remains that greater awareness of the study and attention to the pregnancy in those randomised to the intensive ultrasound group might have both increased rates of attendance for subsequent examinations and enhanced parental attention resulting in earlier acquisition of language. Irrespective of this possibility, there is no evidence from our results of any adverse developmental outcomes from multiple prenatal ultrasound exposures. The seemingly high incidence of problem outcomes in some of the other developmental measures shown in table 3 represent the cut-off levels described by the original creators of each instrument and the fact that these tests were developed for the purpose of screening rather than diagnosis.

Our results also provide reassurance that multiple prenatal ultrasound scans are not followed by smaller body size in infancy or childhood. We noted previously at the time of birth that the proportions of live infants with weight less than the 10th and 3rd centiles were greater in the intensive ultrasound group (relative risk 1.35, 95% CIs 1.09–1.67,  $p=0.006$ ; and 1.65, 1.09–2.49,  $p=0.02$ , respectively).<sup>9</sup> Mean birthweight in the intensive group

	Crude odds ratio (95% CI)	Adjusted odds ratio (95% CI)*	Multivariate odds ratio (95% CI)†
TTS	1.09 (0.84–1.41)	1.10 (0.85–1.42)	1.09 (0.84–1.41)
ELM	0.48 (0.24–0.98)	0.47 (0.23–0.96)	0.47 (0.23–0.96)
IMQ	1.16 (0.92–1.46)	1.11 (0.87–1.40)	1.11 (0.87–1.40)
Denver‡	1.13 (0.93–1.38)	1.12 (0.92–1.36)	1.12 (0.89–1.41)
LDS			
>50 words	0.88 (0.63–1.23)	0.82 (0.58–1.17)	0.79 (0.56–1.12)
Combining words	0.97 (0.67–1.40)	0.90 (0.62–1.31)	0.88 (0.60–1.28)
CBCL	1.04 (0.87–1.24)	1.03 (0.87–1.23)	1.03 (0.86–1.22)
PPVT-R§	1.27 (-0.02–2.56)	1.16 (-0.13–2.46)	0.93 (-0.32–2.18)

IUGR=intrauterine growth restriction. TTS=toddler temperamental scale. ELM=early language milestone scale. IMQ=infant/child monitoring questionnaire. LDS=language development survey. CBCL=child behaviour checklist. PPVT-R=Peabody picture vocabulary test—revised. \*Adjusted for IUGR, prematurity, and sex. †Adjusted for IUGR, prematurity, sex, and other significant predictors of outcome such as maternal education and the language most often spoken at home. ‡Denver developmental screening test. §Differences in means interval between the regular ultrasound versus intensive ultrasound groups and corresponding 95% CIs are shown.

**Table 4: Crude, adjusted, and multivariate odds ratios for abnormal outcomes for the intensive ultrasound group relative to those in the regular ultrasound group, by developmental measure**

was 25 g lower than in the regular group, a difference that was not significant. Babies in the intensive ultrasound group tended to be shorter when measured at birth ( $p=0.123$ ) and on day 2–3 of age ( $p=0.068$ ), but circumferences and skin-fold thicknesses were not significantly different between the two groups.<sup>9,10</sup> Principal components analysis showed a trend for a reduction for the skeletal component ( $p=0.085$ ) but not for the soft tissue component ( $p=0.332$ ) suggesting that if multiple ultrasound scans did indeed affect fetal growth, the mechanism was more likely to have been an effect on bone growth rather than a reduction in nutrient supply from the placenta. When measured at 1 year of age all measures of size were similar in the two groups.<sup>11</sup> Further, by calculation of the ratio of observed to expected weight at 1 year relative to the observed to expected weight at birth, infants in the intensive ultrasound group had shown more rapid growth in the first year of postnatal life relative to their birthweight, suggesting catch-up growth. This possible effect on fetal growth in a proportion of cases was supported by experiments with animal models in which ultrasound exposures have been shown to reduce birthweight.<sup>12,13</sup> Significant reductions in growth have been shown to follow multiple clinically relevant imaging studies in monkeys,<sup>13</sup> although similar studies by ourselves using sheep had no effect on fetal development.<sup>21</sup> It would, however, be incorrect to conclude from our present findings that multiple prenatal ultrasound scans influence growth before birth. The original purpose of our randomised controlled trial was to investigate the hypothesis that a protocol of multiple scans would improve pregnancy outcome and reduce the rate of preterm birth. The original study therefore did not prove that prenatal ultrasound scans affect fetal growth, but has provided evidence that further study is warranted. These results from the childhood follow-up confirm that even if ultrasound scans affected growth before birth, weight, height, circumferences, and skin-fold thicknesses from 1–8 years are unaffected.

Reassurances provided by our results do not lessen our need to undertake further studies of potential bio-effects of prenatal ultrasound scans. The children in this study seem to be the only published cohort from a randomised controlled trial in which a protocol of multiple ultrasound imaging and Doppler flow studies was compared with a single imaging scan and in which ascertainment of gestational age in the two groups has been identical. In view of the widespread and liberal use of this technology we are responsible for ensuring the safety of its use. Uncertainty remains about several potential issues, one of which is handedness.<sup>5–8</sup> Increased rates of non-right handedness have been observed in several studies, with the effect being seen in males. From the results of a meta-analysis including this sex-specific sub-group, Salvesen and Eik-Nes<sup>8</sup> suggested that prenatal ultrasound examinations might result in five extra non-right-handed boys in 100 male births.

Such a possibility could result from effects on neuronal migration, which is active at the time of many prenatal scans.<sup>22</sup> It is also possible that apparent changes in handedness result from confounding by changing community attitudes resulting in less encouragement for children to be non-left-handed, or from nothing more than chance and the consequence of statistical testing of multiple end-points. In the present study, a thorough assessment of handedness is included in the forthcoming 10-year examinations and will be the topic of a future report.

The significance of our incomplete evidence of the safety of prenatal ultrasound is amplified by the increasing power outputs of contemporary medical equipment. We have previously measured the spatial peak temporal average intensity outputs of imaging systems similar to those used in this study and recorded levels less than 5 mW/cm<sup>2</sup>. The output (SPTA [spatial peak, temporal average]) of the Doppler system was 25 mW/cm<sup>2</sup>. These outputs are well below the 100 mW/cm<sup>2</sup> level that had been stated by the Bio-effects Committee of the American Institute of Ultrasound in Medicine in 1991 to be the threshold below which there was no evidence of independently confirmed adverse significant biological effects in mammalian tissue.<sup>23</sup> Much of the more modern equipment however can produce intensities substantially greater than those in the present study.<sup>24</sup> In particular, the stand-alone continuous wave Doppler machine we used has been superseded by pulsed and colour Doppler systems with much higher output potentials. Further, the US Food and Drug Administration has transferred the responsibility for potential bio-effects from the manufacturer to the health-care professional who uses the machine by relaxing the output limits.<sup>23</sup> In some instances, the intensities have been increased by nearly eight times. These increased intensity levels are accompanied by mandatory display of appropriate data on the user's screen, but interpretation of the risk for an embryo or fetus is limited by our incomplete knowledge. Research in this field should continue and needs to include randomised controlled trials using contemporary equipment with pulsed and colour Doppler capabilities, and incorporate the full spectrum of gestational ages in which ultrasound is used in contemporary clinical practice.

#### Contributors

J P Newnham, the chief investigator, initiated and designed the trial, participated in design of childhood follow-up, and was the primary author of the manuscript. D A Doherty, the study's biostatistician, contributed to study design and undertook statistical analyses. G E Kendall was the study coordinator, contributed to study design and preparation of the manuscript, and had a major role in conduct of the study and interpretation of the findings. S R Zubrick had a main role in the psychological testing aspect of the design of childhood follow-up; he was also involved in data collection and preparation of the manuscript, and had a major role in interpretation of the findings. L L Landau and F J Stanley as co-chief investigators contributed to initiation and design of the trial, design of childhood follow-up, had a role in data interpretation and analysis, and helped prepare the manuscript.

**Conflict of interest statement**

We declare that we have no conflict of interest.

**Acknowledgments**

The study was funded by The Raine Foundation, Western Australia; the National Health and Medical Research Council of Australia; Healthways, Western Australia; and the Women and Infants Research Foundation, Western Australia.

**References**

- Salvesen KA, Eik-Nes SH. Is ultrasound safe? A review of epidemiological studies of human exposure to ultrasound. *Ultrasound Obstet Gynecol* 1995; **6**: 293–98.
- Salvesen KA, Jacobsen G, Vatten LJ, Eik-Nes SH, Bakketeig LS. Routine ultrasonography in utero and subsequent growth during childhood. *Ultrasound Obstet Gynecol* 1993; **3**: 6–10.
- Kieler H, Haglund B, Waldenstrom U, Axelsson O. Routine ultrasound screening in pregnancy and the children's subsequent growth, vision and hearing. *Br J Obstet Gynaecol* 1997; **104**: 1267–72.
- Salvesen KA, Eik-Nes SH. Ultrasound during pregnancy and birthweight, childhood malignancies and neurological development. *Ultrasound Med Biol* 1999; **25**: 1025–31.
- Salvesen KA, Vatten LJ, Eik-Nes SH, Hugdahl K, Bakketeig LS. Routine ultrasonography in utero and subsequent handedness and neurological development. *BMJ* 1993; **307**: 159–64.
- Kieler H, Axelsson O, Haglund B, Nilsson S, Salvesen KA. Routine ultrasound screening in pregnancy and the children's subsequent handedness. *Early Hum Dev* 1998; **50**: 233–45.
- Kieler H, Cnattingius S, Haglund B, Palmgren J, Axelsson O. Sinistrality – a side effect of prenatal sonography: a comparative study of young men. *Epidemiology* 2001; **12**: 618–23.
- Salvesen KA, Eik-Nes SH. Ultrasound during pregnancy and subsequent childhood non-right handedness: a meta-analysis. *Ultrasound Obstet Gynecol* 1999; **13**: 241–46.
- Newnham JP, Evans SF, Michael CA, Stanley FJ, Landau LI. Effects of ultrasound during pregnancy: a randomised controlled trial. *Lancet* 1993; **342**: 887–91.
- Evans S, Newnham J, MacDonald W, Hall C. Characterisation of the possible effect on birthweight following frequent prenatal ultrasound examinations. *Early Hum Develop* 1996; **45**: 203–14.
- McDonald W, Newnham J, Gurrin L, Evans S. Effect of frequent prenatal ultrasound on birth-weight: follow up at 1 year of age. *Lancet* 1996; **348**: 482.
- O'Brien W. Dose dependent effect of ultrasound on fetal weight in mice. *J Ultrasound Med* 1983; **2**: 1–8.
- Tarantal AF, Hendrickx AG. Evaluation of the bioeffects of prenatal ultrasound exposure in the Cynomolgus Macaque (*Macaca fascicularis*) I. Neonatal/infant observations. *Teratology* 1989; **39**: 137–47.
- Oberklaid F, Prior M, Sanson A, Sewell J, Kyrios M. Assessment of temperament in the toddler age group. *Pediatrics* 1990; **85**: 559–66.
- Coplan J. Early language milestones: second edition. Texas: Pro-ed, 1993.
- Bricker D, Squires J. The effectiveness of parental screening of at-risk infants: the infant monitoring questionnaires. *Topics in Early Childhood Special Education* 1989; **9**: 67–85.
- Frankenburg W, Dodds J, Archer P, et al. Denver II training manual: second edn. Denver: Denver Developmental Materials, 1992.
- Rescorla L. The language development survey: a screening tool for delayed language in toddlers. *J Speech Hear Dis* 1989; **54**: 587–99.
- Achenbach TM. Manual for the child behaviour checklist/4–18 and 1991 profile. Burlington, VT: University of Vermont Department of Psychiatry, 1991.
- Dunn L, Dunn L. Peabody picture vocabulary test—revised: manual. Circle Pines: American Guidance Service, 1981.
- Harding K, Newnham J, Evans S, Carpenter D, Brogden F, Marshall L. A study of the bioeffects of multiple prenatal ultrasound scans in fetal sheep. *J Matern Fetal Invest* 1996; **6**: 77–82.
- Geschwind N, Galaburda AM. Cerebral lateralization: biological mechanisms, associations and pathology. I. A hypothesis and a program for research. *Arch Neurol* 1985; **42**: 428–59.
- Barnett SB, Ter Haar GR, Ziskin MC, Rott H-D, Duck FA, Maeda K. International recommendations and guidelines for the safe use of diagnostic ultrasound in medicine. *Ultrasound Med Biol* 2000; **26**: 355–66.
- Henderson J, Wilson K, Jago JR, Whittingham TA, Dunn T. A survey of the acoustic outputs of diagnostic ultrasound equipment in current clinical use. *Ultrasound Med Biol* 1995; **21**: 699–705.