Educational Outcomes for Children at 7 to 9 Years of Age After Birth at 39 vs 40 to 42 Weeks’ Gestation

Richard J. Hiscock, MBBS; Jessica A. Atkinson, BBiomed(Hons); Stephen Tong, PhD; Susan P. Walker, PhD; Amber Kennedy, PhD; Jeanie Y. L. Cheong, PhD; Jon L. Quach, PhD; Lyle C. Gurrin, PhD; Roxanne Hastie, PhD; Anthea Lindquist, DPhil

Abstract

**IMPORTANCE** Birth at 39 weeks’ gestation is common and thought to be safe for mother and neonate. However, findings of long-term outcomes for children born at this gestational age have been conflicting.

**OBJECTIVE** To evaluate the association of birth at 39 weeks’ gestation with childhood numeracy and literacy scores at ages 7 to 9 years compared with birth at 40 to 42 weeks’ gestation.

**DESIGN, SETTING, AND PARTICIPANTS** In this Australian statewide, population-based cohort study using a causal inference framework based on target trial emulation, perinatal data on births between January 1, 2005, and December 31, 2011, were linked to educational outcomes at 7 to 9 years of age. Statistical analyses were performed from December 2022 to June 2023.

**EXPOSURE** Birth at 39 weeks’ gestation compared with birth at 40 to 42 weeks’ gestation.

**MAIN OUTCOMES AND MEASURES** Numeracy and literacy outcomes were assessed at 7 to 9 years of age using Australian National Assessment Program–Literacy and Numeracy data and defined by overall z score across 5 domains (grammar and punctuation, reading, writing, spelling, and numeracy). Multiple imputation and doubly robust inverse probability weighted regression adjustment were used to estimate population average causal effects.

**RESULTS** The study population included 155,575 children. Of these children, 49,456 (31.8%; 24,952 boys [50.5%]) were born at 39 weeks’ gestation and were compared with 106,119 (68.2%; 52,083 boys [49.1%]) born at 40 to 42 weeks’ gestation. Birth at 39 weeks’ gestation was not associated with altered educational outcomes for children aged 7 to 9 years compared with their peers born at 40 to 42 weeks’ gestation (mean [SE] z score, 0.0008 [0.0019] vs –0.0031 [0.0038]; adjusted risk difference, −0.004 [95% CI, −0.015 to 0.007]). Each educational domain was investigated, and no significant difference was found in grammar and punctuation (risk difference [RD], −0.006 [95% CI, −0.016 to 0.005]), numeracy (RD, −0.009 [95% CI, −0.020 to 0.001]), spelling (RD, 0.001 [95% CI, −0.011 to 0.0013]), reading (RD, −0.008 [95% CI, −0.019 to 0.003]), or writing (RD, 0.006 [95% CI, −0.005 to 0.016]) scores for children born at 39 weeks’ gestation compared with those born at 40 to 42 weeks’ gestation. Birth at 39 weeks’ gestation also did not increase the risk of scoring below national minimum standards in any of the 5 tested domains.

**CONCLUSIONS AND RELEVANCE** Using data from a statewide linkage study to emulate the results of a target randomized clinical trial, this study suggests that there is no evidence of an association of birth at 39 weeks’ gestation with numeracy and literacy outcomes for children aged 7 to 9 years.

Key Points

**Question** Is birth at 39 weeks’ gestation associated with adverse childhood educational outcomes compared with birth at 40 to 42 weeks’ gestation?

**Findings** In this cohort study of 155,575 births, using a causal inference framework based on target trial emulation, birth at 39 weeks’ gestation was not associated with adverse numeracy and literacy outcomes at school age compared with birth at 40 to 42 weeks.

**Meaning** This study suggests that birth at 39 weeks’ gestation does not affect primary school educational outcomes compared with birth at 40 to 42 weeks’ gestation.
Introduction

Birth at 39 weeks' gestation is becoming increasingly common.1 This trend is likely to be associated with the findings of the ARRIVE trial (A Randomized Trial of Induction Vs Expectant Management) published in 2018.2 This large randomized clinical trial (RCT) found that bringing birth forward to 39 weeks' gestation via induction of labor reduced the rates of cesarean delivery and improved women's experience of birth, without increasing the risk of adverse perinatal outcomes.

Since the ARRIVE trial, further maternal and neonatal benefits have been associated with giving birth (or “delivering”) at 39 weeks' gestation, including a reduced risk of perineal injury, operative vaginal birth, and neonatal intensive care unit admission.3 In addition, our team has reported no differences in early developmental outcomes (aged 4-6 years) for children born electively at 39 weeks' gestation compared with those expectantly managed.4 Although these findings are reassuring, there have been studies demonstrating poorer long-term outcomes beyond early childhood for those born prior to 40 weeks' gestation, even though they were born at term gestation (>37 weeks).5-7

The last trimester of pregnancy (from 28 weeks' gestation onward) is a period of rapid fetal brain development, with a 4-fold increase in brain size and significant growth in surface area.8,9 It is plausible that bringing birth forward by even 1 week may disrupt brain development and have long-lasting neurodevelopmental consequences for children. This notion is supported by recent findings. A study of 39 199 singleton births in the US demonstrated that children's neurocognitive performance improved with each week of gestation gained between 37 and 41 weeks.6 However, other studies have found no difference in cognitive outcomes for children born at 39 weeks' gestation compared with 40 weeks' gestation.10

Previous studies investigating long-term outcomes after birth at term gestations have been limited by the presence of strong confounding factors, which are difficult to account for using standard statistical analysis. Examples of such confounders include overall family socioeconomic position and parental educational level.7 Conducting an RCT to answer this question would address confounders but is not feasible due to the very large numbers needed to assess educational achievement as the primary outcome, as well as the lengthy follow-up required. Thus, to evaluate the association of birth at 39 weeks' gestation with childhood educational achievement, we used a framework for causal inference called target trial emulation to analyze statewide linked data.11 This approach aims to estimate the population average treatment effect (ATE) of an intervention on an outcome—the causal question of interest is first articulated in the form of a detailed protocol for a hypothetical RCT that, if conducted, would answer the question of interest. The components of the protocol, including statistical analysis, are then applied to the observational data under a set of identifiable assumptions12-14 to emulate the results of the target trial. Our study sought to emulate an RCT to answer the causal question: what is the effect of birth at 39 weeks' gestation compared with birth at 40 to 42 weeks' gestation on childhood educational outcomes at 7 to 9 years of age?

Methods

The first step in the target trial emulation process13 (eAppendix 1 in Supplement 1) is to develop a detailed protocol for a hypothetical RCT that, if conducted, would address our goal of population-level treatment comparison. Each component of this RCT formulation is assessed against the information and resources within our retrospective cohort study to determine how well we can emulate the target trial using the observational data available for analysis. This is effectively an exercise in harmonizing the analysis of our observational data with RCT data, if they were available, to eliminate as many sources of bias as possible. The details of our analytical framework were outlined in a prespecified statistical analysis plan, agreed on by all authors prior to commencement of the analysis (eAppendix 1 in Supplement 1). Ethical approval for the project was obtained from the human research ethics committees at Mercy Health. Each data custodian provided contractual
approval for data access and linkage. The Centre for Victorian Data Linkage approved the project and performed the linkage. Given the retrospective and deidentified nature of this study, the requirement for individual participant informed consent was waived by the Mercy Health Human Research Ethics Committee and data custodians. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Population
The hypothetical RCT population included all women with access to obstetric services who reached 39 weeks' gestation and, at that time, had no indication for delivery before 40 weeks' gestation. For our study (the target trial emulation), the source population included all singleton, nonanomalous live births in Victoria, Australia, between January 1, 2005, and December 31, 2011. Perinatal and demographic data were obtained from the validated statewide birth registry—the Victorian Perinatal Data Collection. Race and ethnicity were not included in our model as these factors are not associated with educational performance. Linkage of maternal-child pairs was performed using Victorian Perinatal Data Collection data and birth records (obtained from the Victorian Births, Deaths and Marriages registry). Postlinkage false matches and duplicates were removed. We excluded pregnancies with risk factors identifiable at 39 weeks' gestation that would typically constitute an indication for birth prior to 40 weeks' gestation. These risk factors included twins or higher-order multiple births, in vitro fertilization conception, types 1 and 2 diabetes, suspected fetal growth restriction and large for gestational age fetuses, placental abruption or significant antepartum hemorrhage, preeclampsia, previous cesarean delivery, and breech or cord or shoulder presentation.

Exposure, Assignment Procedures, and Follow-Up Period
The hypothetical RCT cohort consisted of all eligible mothers, randomized in week 38 of pregnancy to 2 birth groups (birth at 39 weeks and 0 days to 39 weeks and 6 days vs 40 weeks and 0 days to 42 weeks and 6 days), who reached 39 weeks' gestation. In our study cohort, we were able to ascertain only if mothers birthed by elective (nonindicated) induction of labor or planned cesarean delivery at 39 weeks' gestation (39 weeks and 0 days to 39 weeks and 6 days completed) or were expectantly managed thereafter (allowed to await spontaneous labor between 40 weeks and 0 days and 42 weeks and 6 days). The follow-up period was 7 to 9 years, when the child had reached the eligible age for outcome assessment; therefore, postrandomization dropouts may have occurred. Reasons for dropout in this study would include stillbirth; neonatal, infant, or child death; or failure to perform National Assessment Program–Literacy and Numeracy (NAPLAN) assessment. Children with missing exposure data (gestational age at birth) were excluded from analysis.

Main Outcome Measure: NAPLAN Results
Childhood educational outcomes at 7 to 9 years of age were assessed using the NAPLAN, a universal, standardized test performed in all mainstream Australian schools in years 3, 5, 7 and 9. NAPLAN is a psychometric assessment across 5 educational domains: grammar and punctuation, numeracy, reading, spelling, and writing. Grade 3 NAPLAN (fourth year of primary school) results were investigated among our study cohort. An unweighted overall z score was calculated from the 5 domains and used as the primary outcome. A priori consensus determined that a mean z score difference of 0.2 SDs would be considered functionally important for student achievement, which aligns with established benchmarks in education. Secondary outcomes included (1) individual domain z scores and (2) individual domain scores (binary) below the published national minimum standard NAPLAN scores (by year of test).

Covariates
The multidisciplinary authorship team decided a priori which covariates should be considered for inclusion in the regression models required for statistical analysis in the target trial emulation. We then used a directed acyclic graph (eAppendix 1 in Supplement 1) to describe the direction and
structure of potential causal relationships between covariates and to identify those required in the selection (propensity score) model. The covariates included in that model were socioeconomic position (characterized by Socio-Economic Indexes for Areas [SEIFA] quintile, in which the lowest quintile represents the most deprived) and maternal education level. The covariates chosen for the regression adjustment (outcome) model included year of NAPLAN testing, child age at test, sex of child, and language background other than English (LBOTE). Potential mediators on the causal pathway (mode of birth and birth weight) between gestation and educational outcome were not included in the analysis models because doing so would potentially result in a biased estimate of the association between exposure and outcome. The directed acyclic graph was included in our prespecified statistical analysis plan (eAppendix 1 in Supplement 1).

Missing Data
In the setting of an RCT, missing outcome data or failure to perform the NAPLAN assessment can arise due to the following mechanisms: (1) informative or missing not at random, where the child is unable to complete some or all NAPLAN assessments due to individual child factors (coded as "exempt from sitting NAPLAN" in the database); and (2) noninformative or missing completely at random (MCAR) mechanisms (eg, child unwell on day of test). In our study setting, missing outcome data could also be considered missing at random (MAR), conditional on prespecified covariates included in the analysis model.

We calculated the frequency of missing data for exempt status in each of the treatment groups (MCAR vs MAR). It was predetermined that, if the frequencies of MCAR and MAR statuses were not substantially different among the treatment groups, then estimation bias could be managed conservatively by deterministically imputing all missing z scores as being equal to −4. If the distribution of exempt scores was nondifferential between exposure groups, the exempt missingness could not bias the final estimate of the population ATE and these children would be excluded from analysis (no difference was found; 2561 cases were excluded; eTable 1 in Supplement 1). For all other missing data (outcome, selection, or regression adjustment covariates), multiple imputation was performed using fully conditional specification accounting for maternal clustering (ie, children sharing the same mother) within the calculation of SEs (details shown in eAppendix 1 in Supplement 1).

Statistical Analysis
Statistical analyses were performed from December 2022 to June 2023. The distribution of maternal, birthing, and child characteristics were summarized using mean (SD), median (IQR), and number (percentage) according to type and distribution of data. Detailed description of missing data included the proportion of missingness across the 5 outcome domains, exposure, and model covariates, along with the total number of observations with complete data. Presentation of missing data patterns included graphical summaries (eTables 1-4 in Supplement 1).

The estimands for the primary outcome, presented as the ATE point estimate and 95% CIs, are defined as the between-treatment risk difference (RD) in mean standardized NAPLAN score. For the secondary outcomes, the estimands are defined as both the RD and relative risk (RR) for each of the 5 individual NAPLAN domains. Within the potential outcomes' framework, a causal interpretation using these estimands can be made, under the assumptions of (1) consistency: the outcome given a participant's observed treatment is the same as it would have been if that participant was randomized to receive that treatment in a trial; (2) ignorability: treatment groups are exchangeable after controlling for, or conditioning on, a set of covariates (ie, there are no important unmeasured confounders); and (3) positivity: the conditional probabilities of receiving either treatment or control must both be greater than 0 and less than 1 in any participant subgroup defined by a combination of covariate values. In practice, this is typically interpreted to mean that all treatments of interest are observed in every participant subgroup defined by a combination of covariate values.
The primary ATE estimator used an augmented doubly robust inverse probability–weighted adjustment (augmented inverse probability weighting [AIPW]) model combing an inverse probability–weighted (IPW) selection model (SEIFA and maternal education) with a regression adjustment model (age at test, year of test, sex, and LBOTE). Sensitivity analyses, using alternate estimators, included (1) an IPW selection model and (2) a regression-adjusted outcome model. All estimators used the same bootstrapped, multiply imputed data sets. Maternal clustering was accounted for in the analysis models. Statistical analysis was performed using Stata statistical software, release 17 (StataCorp LLC) including the teffects suite. Coding for imputation and analysis models is presented in eAppendix 2 and eAppendix 3 in Supplement 1. All P values were from 2-sided tests and results were deemed statistically significant at P < .05.

Results

From 2005 to 2011, 344 447 singleton births occurred in Victoria, Australia, that had NAPLAN outcome data available. After applying exclusion criteria, our population consisted of 158 136 children. Missing outcome data due to exemption from testing occurred among 2561 of 158 136 children (1.62%). The prevalence of missing outcome data due to exemption was essentially the same (to 2 decimal places as a percentage) across exposure groups, and in accordance with the predetermined statistical analysis plan, these cases were excluded (Figure; eTable 1 in Supplement 1). This left a total population of 155 575 for analysis, with 49 456 children (31.8%; 24 952 boys [50.5%]) born at 39 weeks’ gestation and 106 119 children (68.2%; 52 083 boys [49.1%]) born at 40 to 42 weeks’ gestation (Figure). Maternal and child baseline characteristics are shown in Table 1.

The frequency of missing outcome data ranged from 4.9% (n = 7558) for reading to 5.2% (n = 8094) for writing. For the selection model, 1.0% of SEIFA data and 3.9% of maternal education data were missing and were imputed (eTable 4 and eAppendix 2 in Supplement 1). There were no missing data for exposure or regression model covariates (age at test, year of test, sex, and LBOTE). The results of multiple imputation diagnostics generated for the first 5 bootstrap samples are presented in eAppendix 2 in Supplement 1.

Figure. Flowchart


Downloaded from jamanetwork.com by guest on 12/11/2023
Primary Outcome: NAPLAN z Score

We found that birth at 39 weeks' gestation was not associated with educational outcomes for children undertaking NAPLAN testing at 7 to 9 years of age compared with their peers born at 40 to 42 weeks' gestation (Table 2). With the use of the AIPW model on the bootstrap-imputed data sets, the estimated outcome mean z score was 0.0008 (SE 0.0019) for the 39-week cohort and −0.0031 (SE 0.0038) for 40- to 42-week cohort, with an adjusted RD of −0.004 (95% CI, −0.015 to 0.007). Using inverse probability weighting and regression adjustment methods yielded similar results (Table 2).

Table 1. Baseline Characteristics of Exposure Cohorts

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>39 wk (n = 49 456)</th>
<th>40 to 42 wk inclusive (n = 106 119)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal age at birth, mean (SD), y</td>
<td>30.6 (5.2)</td>
<td>30.9 (5.2)</td>
</tr>
<tr>
<td>Nulliparity</td>
<td>22 700 (45.9)</td>
<td>55 943 (52.7)</td>
</tr>
<tr>
<td>Maternal clustering, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>84.6</td>
<td>85.3</td>
</tr>
<tr>
<td>2</td>
<td>14.4</td>
<td>13.6</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>≥4</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;12 y</td>
<td>5074 (10.3)</td>
<td>10 193 (9.6)</td>
</tr>
<tr>
<td>12 y</td>
<td>4923 (10.0)</td>
<td>10 642 (10.0)</td>
</tr>
<tr>
<td>Certificate or diploma</td>
<td>18 470 (37.3)</td>
<td>40 994 (38.6)</td>
</tr>
<tr>
<td>Bachelor degree or above</td>
<td>19 071 (38.6)</td>
<td>40 202 (37.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>1918 (3.9)</td>
<td>4088 (3.9)</td>
</tr>
<tr>
<td>Socio-Economic Indexes for Areas, quintiles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (indicates most deprived)</td>
<td>7879 (16.0)</td>
<td>16 971 (16.0)</td>
</tr>
<tr>
<td>2</td>
<td>6843 (13.8)</td>
<td>14 764 (13.9)</td>
</tr>
<tr>
<td>3</td>
<td>9878 (20.0)</td>
<td>21 438 (20.2)</td>
</tr>
<tr>
<td>4</td>
<td>11 907 (24.0)</td>
<td>25 365 (23.9)</td>
</tr>
<tr>
<td>5</td>
<td>12 895 (26.1)</td>
<td>27 478 (25.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>54 (0.1)</td>
<td>103 (0.1)</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
<td>25.3 (5.3)</td>
<td>25.8 (5.5)</td>
</tr>
<tr>
<td>Child characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language background other than English</td>
<td>13 284 (28.9)</td>
<td>23 705 (22.3)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24 952 (50.5)</td>
<td>52 083 (49.1)</td>
</tr>
<tr>
<td>Female</td>
<td>24 504 (49.5)</td>
<td>54 036 (50.9)</td>
</tr>
<tr>
<td>Birth weight, mean (SD), g</td>
<td>3449 (463)</td>
<td>3632 (475)</td>
</tr>
<tr>
<td>Mode of birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unassisted vaginal birth</td>
<td>35 096 (71.0)</td>
<td>69 043 (65.1)</td>
</tr>
<tr>
<td>Vacuum vaginal birth</td>
<td>3131 (6.3)</td>
<td>8464 (8.0)</td>
</tr>
<tr>
<td>Forceps vaginal birth</td>
<td>4592 (9.3)</td>
<td>11 087 (10.5)</td>
</tr>
<tr>
<td>Unplanned cesarean delivery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In labor</td>
<td>4199 (8.5)</td>
<td>14 930 (14.1)</td>
</tr>
<tr>
<td>No labor</td>
<td>433 (1.0)</td>
<td>1126 (1.1)</td>
</tr>
<tr>
<td>Planned cesarean delivery</td>
<td>1989 (4.0)</td>
<td>1444 (1.4)</td>
</tr>
<tr>
<td>Missing</td>
<td>16 (0.03)</td>
<td>25 (0.02)</td>
</tr>
<tr>
<td>Age at NAPLAN testing, mean (SD), y</td>
<td>7.8 (0.4)</td>
<td>7.8 (0.4)</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); NAPLAN, National Assessment Program for Literacy and Numeracy.

Total cohort = 155 575; excludes children with missing outcome data due to being exempt from testing.
Secondary Outcomes: Grammar, Numeracy, Reading, Spelling, and Writing

Next, we investigated each individual domain of the NAPLAN testing (grammar and punctuation, numeracy, reading, spelling, and writing) (Table 3). Compared with birth at 40 to 42 weeks’ gestation, birth at 39 weeks’ gestation was not associated with a change in grammar (RD, −0.006 [95% CI, −0.016 to 0.005]), numeracy (RD, −0.009 [95% CI, −0.020 to 0.001]), reading (RD, −0.008 [95% CI, −0.019 to 0.003]), spelling (RD, 0.001 [95% CI, −0.011 to 0.0013]), or writing (RD, 0.006 [95% CI, −0.005 to 0.016]) achievement for children tested at 7 to 9 years of age.

Finally, we assessed the association of birth at 39 weeks’ gestation with the risk of children scoring below the national minimum standard in each of the tested domains. We found that birth at 39 weeks’ gestation did not alter the risk of scoring below the national minimum standard for grammar (RR, 0.99 [95% CI, 0.93-1.06]), numeracy (RR, 1.08 [95% CI, 0.98-1.19]), reading (RR, 1.04 [95% CI, 0.97-1.11]), spelling (RR, 1.03 [95% CI, 0.96-1.10]), or writing (RR, 1.02 [95% CI, 0.93-1.12]) compared with peers born at 40 to 42 weeks’ gestation (Table 4).

Discussion

Using an inferential framework based on target trial emulation, we found no association between birth at 39 weeks’ gestation and childhood numeracy and literacy scores at 7 to 9 years of age compared with children born at 40 to 42 weeks’ gestation. Investigating individually tested domains,
we also found no difference in any domain scores, nor in the risk of children scoring below the national minimum standard.

Given the increasing number of children born at 39 weeks' gestation, these findings are reassuring and are in keeping with other, smaller studies from comparable settings. Our findings suggest that the practice shift toward birth at 39 weeks' gestation is not only safe in the short term for mother and baby but also has no adverse effects on both early developmental outcomes and later primary school educational attainment.

Our findings are particularly reassuring given that some medical indications for planned 39 weeks' gestation were likely to have been underreported. This means that our 39-week group may have inadvertently included more women with high-risk pregnancies, which have been associated with an increased risk of developmental vulnerability, yet this was not seen in our findings.

**Strengths and Limitations**

This study has some strengths. The major strengths of our study lie in our population-wide cohort and analysis of more than 150,000 children, the high proportion of matched outcome data, and our use of a formal, mathematical framework for causal inference. Here we have used recently developed methods for multiple imputation of missing data based on the work of Bartlett and Hughes, as well as von Hippel and Bartlett. The strength of our approach is highlighted by the recent contrasting findings by Selvaratnam et al. Using a similar Australian cohort, Selvaratnam et al report a 39% increased likelihood of poor educational outcomes for children in grade 3 who were born after elective induction of labor at 39 weeks' gestation. Their report concluded possible harm from birth at 39 weeks' gestation. However, the authors used multivariate logistic regression and simply excluded individuals with missing data, rather than performing imputation. These outcome regression approaches to causal inference are highly sensitive to residual confounding, which may explain the disparity with our findings. These contrasting results demonstrate the importance of appropriate considerations during study design and modeling when investigating questions of clinical decision-making.

Our study also has some limitations, which have been carefully considered and mitigated throughout our analysis, where possible. First, our target population is women with access to quality obstetric and neonatal care in well-resourced settings. As such, our findings likely apply to the most populous states in Australia and comparable settings worldwide, but not to more rural and less-resourced areas in Australia and globally. Second, our use of a school-based outcome assessment limits our cohort to children attending school. Our cohort will have excluded a small percentage of children with a disability significant enough not to attend mainstream school, which may have introduced selection bias. However, while we recognize this limitation, our study was not designed to examine outcomes of severe disability or developmental delay, but rather an overall measure of educational achievement.

In addition, using school-based outcomes, our study was inherently designed to examine outcomes for liveborn children. Live birth bias is a recognized limitation of observational studies investigating periconception and antenatal exposures. In our study, the outcome of stillbirth is a potential alternative end point. However, stillbirth is not directly relevant to our research question, which aimed to compare the school-aged educational outcomes of children born after elective birth with those managed expectantly. Last, children born at 42 weeks' gestation were included in our expectant management group. Given that previous studies have suggested that birth at 42 weeks' gestation may be associated with worse outcomes, it is possible that inclusion of this gestational age in our control group may be masking an adverse effect of birth at 39 weeks' gestation. However, we are reassured that the possible effect of this would be small, with only 3.1% of children in our population born at 42 weeks.
Conclusions

Our findings in this Australian statewide, population-based study using a causal inference framework based on target trial emulation revealed no association of birth at 39 weeks' gestation with children's primary school-aged educational outcomes. These results provide reassurance to families and clinicians that planned birth at 39 weeks' gestation was not associated with advanced primary school-aged educational achievement.
REFERENCES


SUPPLEMENT 1.
- eAppendix 1. Statistical Analysis Plan
- eTable 1. Children Exempt From NAPLAN Testing by Exposure Status
- eTable 2. Pattern of Missing Data Across Cohort
- eTable 3. Pattern of Missing Among Domain Outcomes Only
- eTable 4. Missing Data in Selection Model Covariates by Exposure
- eAppendix 2. Details of Multiple Imputation Model
- eAppendix 3. Details of Analysis Models

SUPPLEMENT 2.
- Data Sharing Statement