Maternal cannabis use and birth weight: a meta-analysis

D. R. ENGLISH, G. K. HULSE, E. MILNE, C. D. J. HOLMAN & C. I. BOWER

Department of Public Health, University of Western Australia, Australia

Abstract

Aims. To estimate the effect of maternal cannabis use on birth weight. Design. Meta-analysis of published observational studies adjusted for cigarette smoking. Separate analyses were performed for studies of low birth weight and mean birth weight. We used fixed and random effects models, but in all cases the results were identical. Setting. From the Medline database, we identified 10 studies in which the results were adjusted for cigarette smoking. In seven studies, information on cannabis use was collected prenatally. Five studies reported results for differences in mean birth weight associated with maternal cannabis use.

Participants. 32,483 women giving birth to live-born infants. Measurements. Mean birth weight and odds ratio for low birth weight. Findings. Three analyses of the studies on mean birth weight were conducted to avoid double-counting women from one study. The largest reduction in mean birth weight for any cannabis use during pregnancy was 4 g (95% confidence interval (CI) 83–14 g), with considerable heterogeneity among the five studies. Mean birth weight was increased by 62 g (95% CI 8 g reduction–132 g increase; p heterogeneity 0.59) among infrequent users (< weekly) whereas cannabis use at least four times per week had a 131 g reduction in mean birth weight (95% CI 52–209 g reduction; p heterogeneity 0.25). From the five studies of low birth weight, the pooled odds ratio for any use was 1.09 (95% CI 0.94–1.27, p heterogeneity 0.19). Conclusions. There is inadequate evidence that cannabis, at the amount typically consumed by pregnant women, causes low birth weight.

Introduction

Does consumption of cannabis during pregnancy cause low birth weight babies? Epidemiological research on this question has produced apparently inconsistent results (see Tables 1–3). One reason for the apparent inconsistency is low precision in most studies.

To overcome problems of small sample sizes, we undertook a meta-analysis of published data on the association of cannabis use during pregnancy and birth weight. Meta-analysis enables a single summary measure of effect to be calculated from any number of individual studies (Greenland, 1987). Cigarette smoking is known to cause low birth weight (English, Holman & Milne, 1995) and is positively associated with use of cannabis (Linn et al., 1983; Fried, Watkinson & Willan, 1984; Day, Wagener &
Taylor, 1991). Therefore, control for smoking is essential to avoid confounding. Thus, we included only those reports that presented results adjusted for smoking.

**Methods**

We searched for articles indexed in the Medline database from 1966 to November 1995. The following terms were used in the literature search: cannabis, substance abuse, fetal-development, pregnancy-complications, neonatal diseases and abnormalities, infant-newborn, birth weight and all sub-headings.

The search was restricted to articles published in English. Commentaries, letters and abstracts were excluded, as they did not usually present sufficient information for inclusion in a meta-analysis. Review articles and reference lists from the articles obtained were used to confirm that all relevant papers had been identified.

Ten studies of cannabis use during pregnancy and birth weight were included in the analysis (Table 1). The outcome was reported as differences in mean birth weight in five studies (Hingson et al., 1982; Fried et al., 1984; Kline, Stein & Hutzler, 1987; Zuckerman et al. 1989; Day et al., 1991), low birth weight (i.e. < 2500 g) in three studies (Linn et al., 1983; Kliegman et al., 1994; Shiono et al., 1995) and both mean birth weight and low birth weight in the remaining two studies (Hatch & Bracken, 1986; Teitelman et al., 1990). Both these studies reported more detailed analysis for low birth weight and were included only in our analysis of this outcome.

Ten studies were excluded from the analysis. Five studies in which the results were not adjusted for cigarette smoking (Greenland et al., 1983; Witter & Niebyl, 1990; Hayes et al., 1991; Castro et al., 1993; Yawn et al., 1994) were excluded on the grounds that their results were likely to be affected by confounding. In another study (Singer et al., 1994), the data presented were not adjusted for smoking. Another analysis involved stepwise regression analysis; cigarette smoking was included in this model, but cannabis use did not enter the model and so no results adjusted for smoking were reported (Singer et al., 1994). Three excluded reports were from the Ottawa Prenatal Prospective Study; in two of these (Fried, 1980; Fried et al., 1983), the subjects were subsets of those in another report from this study which was included in the meta-analysis (Fried et al., 1984), and the other report contained no data on the relationship between cannabis use and birth weight (Fried & O'Connell, 1987). The tenth excluded study reported a p-value but no measure of effect (Gibson, Baghurst & Colley, 1983).

**Statistical analysis**

Precision-based weighting was used to obtain pooled estimates for both types of outcome. The pooled estimate is given by:

\[
X_p = \sum_{i=1}^{n} W_i \frac{X_i}{\sum_{i=1}^{n} W_i}
\]

where \(X_i\) is the effect of cannabis consumption within study \(i\), \(W_i\) the reciprocal of its variance and \(n\) is the number of individual results. The variance of the pooled estimate is \(1/\sum W_i\); and the 95% confidence interval is given by \(X_p \pm 1.96/\sqrt{\sum W_i}\). This model is referred to as the fixed effects model. We also used the random effects model to analyse the data (Fleiss, 1993) but in all cases, the two models gave identical point estimates and confidence intervals.

In the case of studies of mean birth weight, \(X_i\) is the difference between the mean birth weights of infants whose mothers used cannabis and the mean birth weight of infants whose mothers did not use cannabis. \(W_i\) is the reciprocal of the square of the standard error of the difference. In the case of studies of low birth weight, \(X_i\) is the natural logarithm of the relative risk for use of cannabis and \(W_i\) the reciprocal of its variance. The pooled estimate of the relative risk is \(\exp(X_p)\).

A test statistic for the hypothesis that the results of the individual studies are homogeneous is

\[
Q = \sum_{i=1}^{n} W_i (X_i - X_p)^2
\]

\(Q\) has a \(\chi^2\) distribution with \(n - 1\) degrees of freedom.

If the results of a study were presented in terms of regression coefficients for difference in birth weight and standard errors, the quantities required for the meta-analysis were available directly. However, for some studies, the mean difference and its standard error had to be calculated from the mean birth weights of exposed and non-exposed infants and the standard devia-
Table 1. Characteristics of studies included in analysis

<table>
<thead>
<tr>
<th>Reference</th>
<th>Period</th>
<th>Subjects</th>
<th>Country</th>
<th>Time collected</th>
<th>Type of data</th>
<th>Outcome*</th>
<th>Adjustments†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fried et al., 1984</td>
<td>1980-82</td>
<td>583</td>
<td>Canada</td>
<td>Prenatal</td>
<td>Self-report</td>
<td>BW</td>
<td>MA, Wt, S, T</td>
</tr>
<tr>
<td>Hatch &amp; Bracken, 1986</td>
<td>1980-82</td>
<td>3857</td>
<td>USA</td>
<td>Prenatal</td>
<td>Self-report</td>
<td>BW &amp; LBW</td>
<td>Race, T</td>
</tr>
<tr>
<td>Hingson et al., 1982</td>
<td>1977-79</td>
<td>1690</td>
<td>USA</td>
<td>Postnatal</td>
<td>Self-report</td>
<td>BW</td>
<td>GA, MA, WG, Race, T</td>
</tr>
<tr>
<td>Kliegman et al. 1994</td>
<td>1990-91</td>
<td>425</td>
<td>USA</td>
<td>Postnatal</td>
<td>Self-report</td>
<td>LBW</td>
<td>A, MA, race, cocaine, T, Par</td>
</tr>
<tr>
<td>Kline et al., 1987</td>
<td>1975-83</td>
<td>3443</td>
<td>USA</td>
<td>Prenatal</td>
<td>Self-report</td>
<td>BW</td>
<td>T</td>
</tr>
<tr>
<td>Linn et al., 1983</td>
<td>1977-80</td>
<td>12424</td>
<td>USA</td>
<td>Postnatal</td>
<td>Self-report</td>
<td>LBW</td>
<td>Race, MA, T, A, Par, education, previous stillbirths and miscarriages</td>
</tr>
<tr>
<td>Shiono et al., 1995</td>
<td>1984-89</td>
<td>7470</td>
<td>USA</td>
<td>Prenatal</td>
<td>Self-report</td>
<td>LBW</td>
<td>T, cocaine, A, previous LBW baby, race, BMI</td>
</tr>
<tr>
<td>Teitelman et al., 1990</td>
<td>1980-82</td>
<td>1206</td>
<td>USA</td>
<td>Prenatal</td>
<td>Self-report</td>
<td>BW &amp; LBW</td>
<td>GA, Par, T, education, race</td>
</tr>
<tr>
<td>Zuckerman et al., 1989</td>
<td>1984-87</td>
<td>1226</td>
<td>USA</td>
<td>Prenatal</td>
<td>Urine assay</td>
<td>BW</td>
<td>GA, race, T, cocaine, A, opiates</td>
</tr>
</tbody>
</table>

*BW, birth weight (g); LBW, low birth weight (< 2500 g versus ≥ 2500 g).
†A, alcohol consumption; GA, gestational age; WG, weight gain during pregnancy; Ht, maternal height; Par, parity; S, sex of infant; Inc, income; MA, maternal age; T, tobacco smoking.
tions and sample sizes of the two groups (see Table 2). For one report (Hingson et al., 1982), the standard errors of the regression coefficients had to be estimated from the p-values given in the paper. In one study (Day et al., 1991), separate results were presented for cannabis use during each of the trimesters of pregnancy. Because these results involved the same subjects, to include them all in the same meta-analysis would be inappropriate. Therefore, we calculated three pooled estimates, corresponding to each of the separate results for this study. The results of the individual studies are shown in Table 2.

For studies of mean birth weight, we examined any use of cannabis during pregnancy and we also attempted to examine dose–response effects by grouping the data wherever possible into infrequent use, defined arbitrarily as use no greater than once per week and frequent use, defined as use at least four times per week. We were concerned that apparent differences in effect according to frequency of consumption might actually be due to heterogeneity among studies. Thus, we restricted these analyses to studies which had data in both our designated categories; the only two studies meeting these criteria were Fried et al. (1984) and Kline et al. (1987). The analysis of infrequent use included Fried et al. (≤ 1 per week) and Kline et al. (< 1/month and 2-4/month). The analysis of frequent use included Fried et al. (> 5/week) and Kline et al. (4-6/week and daily). We did, however, repeat the analysis of frequent use after including the studies of Day et al. (1991), in which consumption was daily, and Zuckerman et al. (1989), in which cannabis use was defined as a positive urine assay. The latter was included on the grounds that a positive urine test represents use in the immediate past, and probably indicates frequent use.

All studies in which low birth weight was the outcome reported odds ratios and corresponding confidence intervals; the variances were calculated from the confidence intervals given in the reports (Greenland, 1987). The results of the individual studies are shown in Table 3. Because four of the five studies reported results only for

<table>
<thead>
<tr>
<th>Reference</th>
<th>Type of use</th>
<th>Difference in mean birth weight (g)</th>
<th>SE (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day et al., 1991</td>
<td>1 + /day in 1st trimester</td>
<td>[− 45]*</td>
<td>46.14</td>
</tr>
<tr>
<td></td>
<td>1 + /day in 2nd trimester</td>
<td>[93]</td>
<td>70.63</td>
</tr>
<tr>
<td></td>
<td>1 + /day in 3rd trimester</td>
<td>[142]</td>
<td>68.46</td>
</tr>
<tr>
<td>Fried et al., 1984</td>
<td>&lt; = 1/week</td>
<td>[67]</td>
<td>77.88</td>
</tr>
<tr>
<td></td>
<td>2-5/week</td>
<td>[117]</td>
<td>124.73</td>
</tr>
<tr>
<td></td>
<td>&gt; 5/week</td>
<td>[− 52]</td>
<td>124.11</td>
</tr>
<tr>
<td>Hingson et al., 1982</td>
<td>&lt; 3/week</td>
<td>− 95</td>
<td>40.84†</td>
</tr>
<tr>
<td></td>
<td>3 + /week</td>
<td>− 139</td>
<td>59.75†</td>
</tr>
<tr>
<td>Kline et al., 1987</td>
<td>&lt; 1/month (Phase I)</td>
<td>36.9</td>
<td>60.8</td>
</tr>
<tr>
<td></td>
<td>2-4/month (Phase I)</td>
<td>199.1</td>
<td>92.8</td>
</tr>
<tr>
<td></td>
<td>2-3/week (Phase I)</td>
<td>− 6.9</td>
<td>112.0</td>
</tr>
<tr>
<td></td>
<td>4-6/week (Phase I)</td>
<td>− 222.5</td>
<td>204.3</td>
</tr>
<tr>
<td></td>
<td>Daily (Phase I)</td>
<td>84.9</td>
<td>119.7</td>
</tr>
<tr>
<td></td>
<td>&lt; 1/month (Phase II)</td>
<td>30.4</td>
<td>84.8</td>
</tr>
<tr>
<td></td>
<td>2-4/month (Phase II)</td>
<td>6.6</td>
<td>102.0</td>
</tr>
<tr>
<td></td>
<td>2-3/week (Phase II)</td>
<td>− 126.9</td>
<td>96.9</td>
</tr>
<tr>
<td></td>
<td>4-6/week (Phase II)</td>
<td>− 143.3</td>
<td>199.3</td>
</tr>
<tr>
<td></td>
<td>Daily (Phase II)</td>
<td>− 229.6</td>
<td>79.3</td>
</tr>
<tr>
<td>Zuckerman et al., 1989</td>
<td>Positive urine assay</td>
<td>− 79</td>
<td>45.13</td>
</tr>
</tbody>
</table>

*Square brackets indicate that mean change in birth weight and its standard error were calculated from mean birth weights, sample sizes and standard deviations in exposed and non-exposed groups. †Standard error calculated from p-value. ‡Phase I and Phase II refer to different study populations.
‘any’ use, only one analysis, for any use during pregnancy, was possible.

For both types of outcome, we also reanalyzed the data after excluding those studies in which the information on cannabis use during pregnancy was collected postnatally, since these studies might be subject to recall bias.

**Results**

The pooled estimates for reduction in mean birth weight associated with any cannabis use during pregnancy varied from 35 g to 48 g (Table 4). The confidence intervals for the estimates including the first and second trimester results of Day et al. (1991) did not include zero. There was substantial heterogeneity among studies and all p-values for tests of homogeneity were less than 0.05. After removing the results of Hingson et al. (1982), the only study involving postnatal collection of data on cannabis, the estimated reduction in mean birth weight was 26 g (95% CI - 66 to 15 g) for the analysis involving the first trimester results from Day et al. The mean reductions were lower for the other two analyses.

Use no more than once per week was associated with an increased mean birth weight, although the confidence interval included zero. Frequent consumption was associated with a reduction in birth weight of 131 g, for which the confidence interval was (-209 to -52 g) (Table 4). When we included the results from Zuckerman in the analysis of frequent use, the mean reduction in birth weight was 108 g (95% CI -139 to -50 g). Adding the results for each trimester from the study by Day et al. gave estimates and confidence intervals of -90 g.
(−139−40 g) for the first trimester, 78 g (−132−24 g) for the second trimester results and 68 g (−122−14 g) for the third trimester data.

The analysis of studies reporting results for low birth weight gave a pooled estimate of the odds ratio of 1.09 (95% CI 0.94–1.27) for any cannabis use during pregnancy. The test of homogeneity among studies had a p-value of 0.19. When we removed the data from the two studies in which information on consumption was collected postnatally (Linn et al., 1983; Kliegman et al., 1994), the pooled estimate was 1.11, with 95% CI 0.89–1.40.

Discussion

We found weak evidence that any consumption of cannabis during pregnancy was associated with either a reduction in mean birth weight or with low birth weight. Use no more than weekly was associated with increased birth weight, albeit not statistically significant, but use on at least four occasions per week was associated with lower birth weights, although the mean reduction in birth weight was only 131 g. In one study that we excluded, but which adjusted for smoking, the use of cannabis at least once per week was associated with lower birth weights (p = 0.011; Gibson et al., 1983). In another excluded study (Singer et al. 1994), marijuana use was associated with lower birth weight in crude results, but did not enter a stepwise regression analysis, suggesting that the crude association was due to confounding by smoking. There was considerable inconsistency of results in the studies reviewed, as demonstrated by the low p-values of the tests for homogeneity. In some studies, frequent use of cannabis was not associated with any reduction in birth weight, while in other studies in which cannabis use was less frequent, there was a reduction.

Meta-analysis is a convenient way of summarizing the results of multiple studies, but it does not overcome problems associated with differences among studies in subject selection and measurement of exposure. One of its strengths is that it does enable the investigator to test for heterogeneity among studies. When such heterogeneity exists, as it did here, the results should be viewed with caution. Another potential problem when reviewing the literature is publication bias—the tendency to report strong (and statistically significant) associations and not to publish weak (and non-significant effects). For example, in the study by Singer et al. (1994), no results were presented for cannabis use because, after adjustment for cigarette smoking, it did not have a statistically significant association with birth weight. The effect of this bias, presuming it is operating here, would be to increase the apparent effect of cannabis consumption. Thus, the true effect might be even weaker than was observed.

The effect was weaker than the 200 g reduction in mean birth weight or the two-fold increase in the relative risk for low birth weight that has been observed for cigarette smoking (English et al., 1995). Confounding by smoking is likely to be one explanation for any association between cannabis use and low birth weight. Although all the studies included in the pooled analysis were adjusted for smoking, in some the categorization of amount smoked was coarse. Furthermore, confounding by tobacco may arise because of the practice of mixing tobacco with cannabis to assist burning; none of the studies addressed this issue.

If cannabis use does cause low birth weight, then variations in the amount of cannabis used by women in the different studies could have contributed to heterogeneity among studies. The reporting of ‘any’ use by several of the studies lessens the ability to examine dose-response effects and contributes further to heterogeneity. Similarly, varying degrees of under-reporting of cannabis use among studies would also have increased heterogeneity. All studies except one relied on data from self-reported consumption. Several authors have noted the difficulty of obtaining accurate information on use of licit and illicit drugs from pregnant women (Day et al., 1985; Day & Robles, 1989; Elsohly & Elsohly, 1989; Zuckerman et al., 1989). Given an effect, under-reporting of cannabis will weaken it, but differences in the degree of under-reporting according to whether the infant is of low birth weight may weaken or strengthen it. Removing the study in which information was collected postnatally weakened the association with mean birth weight, but had little effect on the odds ratio for low birth weight. Thus, this bias does not appear strong in these studies.

Zuckerman et al. (1989) have argued that the lack of association reported for studies using self-reported exposure is an artefact caused by
poor measurement of cannabis use. They claimed to find an association with cannabis metabolites detected in the urine, but not with self-reported use. In a multiple linear regression, they included two indicators of cannabis use, urine detection with or without self-reported use, and use detected by self report but not in urine. The first variable was statistically significant, but the second was not. However, this analysis does not address the issue of whether self-reported use per se is a predictor of low birth weight. Furthermore, the two methods of measuring cannabis use showed reasonable agreement (kappa = 0.53). Additionally, three-quarters of women who tested positive admitted to use, but only about one-half of those women who reported using cannabis tested positive. Nevertheless, detection of urine metabolites of cannabis may be useful, because it is likely to identify a group of frequent users of the drug.

The amount of cannabis consumed was generally quite low. Does consumption of cannabis at much higher levels—equivalent to the amount of tobacco consumed by smokers—cause low birth weight? It may, but the data are inadequate to provide the answer. We conclude that there is inadequate evidence that maternal cannabis use, at the levels of consumption typically reported, causes low birth weight. More high-quality cohort studies, with adequate documentation of cannabis consumption throughout pregnancy will be required before a more definitive conclusion can be reached.

Acknowledgements

The study was funded by a Research into Drug Abuse grant from the Commonwealth Department of Human Services and Health.

References


