Report to the Department of Health

REVIEW OF THE FETAL EFFECTS OF PRENATAL ALCOHOL EXPOSURE

May 2006

Ron Gray

Jane Henderson

National Perinatal Epidemiology Unit,

University of Oxford

Contents

Executive summary 1

General introduction 5

Part a 7

Prenatal Alcohol Exposure 9

1 Background and Context 9

2 Methods 9

3 Outline of document 9

4 Alcohol consumption during the periconceptional period and during pregnancy 9

4.1 Measuring alcohol consumption 10

4.2 Alcohol consumption in women of childbearing age: recent trends in the UK 12

4.3 Biomarkers 13

5 Terminology of fetal alcohol syndrome and fetal alcohol spectrum disorders 14

6 Diagnosis and classification of fetal alcohol spectrum disorders 14

7 Overview of human observational studies investigating fetal alcohol effects 16

7.1 Descriptive studies of prevalence 17

7.2 Analytical studies 18

7.3 Methodological issues in design and analysis 19

8 Key points from animal experiments 21

8.1 Advantages of experimental studies 22

8.2 Main findings 22

8.3 Limitations 23

9 Paternal contribution to fetal alcohol effects 23

10 Prevention 24

11 New developments 26

Part b 27

Systematic Review of Fetal Effects of Low-to-Moderate Prenatal Alcohol Exposure and Binge Drinking 29

1 Introduction 29

1.1 Aim 29

1.2 Preliminary search of the literature 29

1.3 Measurement of alcohol consumption and definition of low-to-moderate and bingeing 30

2 Methods 30

2.1 Study inclusion and exclusion 30

2.2 Outcomes (as defined by the authors) 31

2.3 Search strategy (appendix 5) 31

2.4 Study selection criteria and procedures 31
2.5 Quality assessment 32
2.6 Data extraction 32
2.7 Presentation and synthesis of extracted data 32
3 Results 32
3.1 Abbreviations used in tables 33
Fig 1 - Results of search 33
3.2 Low-to-moderate alcohol consumption 34
3.2.1 Spontaneous abortion 34
3.2.2 Stillbirth 38
3.2.3 Antepartum haemorrhage 41
3.2.4 Intrauterine growth restriction 42
3.2.5 Birth weight 46
3.2.6 Preterm birth 56
3.2.7 Malformations 63
3.2.8 Postnatal growth 67
3.2.9 Head circumference and length at birth 69
3.2.10 Neurodevelopmental outcomes 73
3.3 Binge drinking 79
3.3.1 Birthweight, gestational age and growth 79
3.3.2 Birth defects 79
3.3.3 Neurodevelopmental outcomes 79
3.4 Quality of papers included in systematic review 89
4 Discussion and Conclusions 91
4.1 Principal conclusions from the systematic review 91
4.2 Issues arising from expert group meeting (see appendix 2) 93
5 Acknowledgments 93
6 References 94
Part c **103**
Appendix 1: Members of advisory group on fetal effects of alcohol 105
Appendix 2: Report of meeting of advisory group – 8th Dec 2005 108
Appendix 3: UK and USA standard measurement and quantity of alcohol 111
Appendix 4: Neurodevelopmental outcomes 113
Appendix 5: Medline search strategy 116
Appendix 6: Newcastle - Ottawa quality assessment scale 119
Appendix 7: Data extraction form 121
Appendix 8: Summary of outcomes by first author 125

Executive summary

In May 2005 the Department of Health for England commissioned the National Perinatal Epidemiology Unit (NPEU) to undertake a review of the existing evidence on the effects of alcohol on the developing embryo, fetus and child. The purpose of the review was first, to present the existing evidence on the effects of prenatal alcohol exposure (with a particular focus on the effects of low-to-moderate exposure and binge drinking), second, to identify research gaps and priorities.

The review was informed by input from an advisory group (members listed in appendix 1). Following a background section on prenatal alcohol exposure (Part a), the methods and results of a systematic review of the effects of low-to-moderate prenatal alcohol exposure and binge drinking are reported (Part b) and the results discussed along with the research priorities identified by the advisory group in the area of prenatal alcohol exposure.

Summary of methods and results of systematic review of the fetal effects of low-to-moderate prenatal alcohol exposure and binge drinking

This review was carried out between June and December 2005. It was only feasible to carry out a systematic review on a particular area of policy interest within the time available and hence the choice was made to focus on the effects of low-to-moderate prenatal alcohol exposure and binge drinking on the embryo, fetus and developing child.

1 Exposure to alcohol - definitions used

Low-to-moderate prenatal alcohol exposure - This was defined as less than one drink per day (equivalent to maximum 1.5 UK units or 12 grams of alcohol daily). This was compared to no alcohol consumption or very small amounts.

Binge drinking - Authors' definitions were used. These definitions varied between studies but a 'binge' was mainly defined as 5 or more drinks on any one occasion.

2 Methods

The bibliographic databases Medline, Embase, PsychInfo and Cinahl were searched using a search strategy developed and piloted in Medline. In addition a number of reviews and books, as well as material from the advisory group, were used to supplement the search. For pragmatic reasons the search was restricted to material in English during 1970-2005. Review articles, commentaries, case series and editorials were excluded. The final search resulted in 3630 papers. The titles and abstracts, where available, were independently scanned by two researchers. Using predefined inclusion and exclusion criteria these were narrowed down to 395 potentially relevant papers. These were obtained and independently read in full by two researchers. Of these, 74 were included (8 of which were unobtainable) inclusion and exclusion being based on the same criteria as before. The papers were assessed for quality using the Newcastle-Ottawa Quality Assessment Scales and data were extracted on the pre-specified range of outcomes detailed below.

3 Results

3.1 General Conclusion

In summary, for most outcomes, there was no consistent evidence of adverse effects from low-to-moderate prenatal alcohol consumption. Nevertheless, the evidence is probably not strong enough to rule out any risk. There was some evidence of adverse effects on neurodevelopment of binge drinking during pregnancy.

3.2 Low-to-moderate consumption

3.2.1 Spontaneous abortion - There were eight studies which examined the effect of low-to-moderate alcohol consumption on this outcome. Although five of these reported a significant effect, two had significant limitations, and in one paper the only significant result was amongst heavy smokers.The remaining two studies reported results of borderline statistical significance.

3.2.2 Stillbirth - None of the five studies which examined this outcome found a significant effect of low-to-moderate drinking in pregnancy. Three studies reported higher rates of stillbirth in women who abstained but these were not statistically significant differences and were unadjusted for potential confounders.

3.2.3 Antepartum haemorrhage - There was only one study which examined this outcome and no significant differences were found.

3.2.4 Intrauterine growth restriction - Only one of the seven studies which examined this found a significant association and that was unadjusted for potential confounders. Three studies found low-to-moderate alcohol consumption to be mildly protective but, although of borderline statistical significance, two may have been subject to recall bias.

3.2.5 Birth weight - Of the 20 studies which included birth weight as an outcome, only one reported a significant excess of low birth weight associated with low-to-moderate alcohol consumption in pregnancy. This result was inconsistent in that higher levels of consumption were not associated with increased risk. Small amounts of alcohol appeared to exert a mildly protective effect.

3.2.6 Preterm birth - As with birth weight, only one study out of 16 reported a significantly increased risk of preterm birth. This study may be subject to residual confounding as it was unadjusted for socioeconomic status.

3.2.7 Malformations - None of the seven studies that examined this outcome found a significant association.

3.2.8 Head circumference and birth length - Of the five studies reporting on these outcomes, one found a higher proportion of low birth weight babies among those whose mothers drank low-to-moderate amounts in pregnancy. However, the tests of statistical significance were across the whole range of exposure so interpretation of this difference was problematic. Moreover, there was no adjustment for potential confounders in this analysis. None of the other studies reported any differences at these levels of consumption.

3.2.9 Postnatal growth - There were only two studies which examined the association between alcohol exposure and growth as measured in childhood. One of these studies, which followed children up to age 14, found that children of women who drank small amounts in pregnancy were consistently lighter with smaller head circumference. However, the other study found the opposite,

that children of abstainers tended to be lighter with smaller head circumference. However, neither of these studies reported the statistical significance of these findings and there were significant other problems with the second one.

3.2.10 Neurodevelopmental outcomes - Of the seven studies which looked at this outcome, one was conducted at birth, the others were later in childhood. Only one study found small but significantly poorer results in children of low-to-moderate drinkers. However, this analysis was unadjusted for potential confounders.

3.3 Binge drinking

There were 11 separate studies which examined the effect of binge drinking on the outcomes above. Only the four studies that looked at neurodevelopmental outcomes showed consistently poorer results in children exposed to binge drinking in pregnancy. Effects, which were generally quite small, included an increase in 'disinhibited behaviour', a reduction in verbal IQ and increase in delinquent behaviour, and more learning problems and poorer performance. The studies which considered these issues were not without problems, including possible overlap between binge drinkers who otherwise drink little and binge drinkers who generally drink substantial amounts. However, they represent the most consistent evidence of a possible effect suggesting that binge drinking in pregnancy may be associated with poor neurodevelopmental outcomes.

3.4 Quality of included studies

Although the studies generally scored quite high on the Newcastle-Ottawa Quality Assessment Scale, there were some areas in which there were problems. Many of the papers assessed alcohol consumption postnatally when the outcomes were apparent and may therefore be subject to recall bias. Although most studies did adjust for potential confounders, it was often done at a later stage in the analysis than those describing results for low-to-moderate alcohol consumption. Alternatively, there may have been residual confounding. Similarly, statistical significance was often reported for the whole range of consumption including high levels, not specifically for low-to-moderate consumption.

Summary of meeting of advisory group, Dec 2005 (see section part b, 4.2 and appendix 2)

This group (see appendix 1 for list of members) was brought together at a one-day meeting in London to hear the preliminary results of the systematic review and to identify research priorities for the field of prenatal alcohol exposure. Short presentations were given by five members of the group on measuring alcohol consumption by maternal self report; diagnosis of fetal alcohol syndrome; neuroimaging studies; study designs for prenatal alcohol exposure; and treatment of alcohol problems in women. The preliminary results of the systematic review were then presented and discussed in plenary.

Key research topics prioritised by the three groups included:

- What are the effects of low-to-moderate prenatal exposure on IQ, socio-emotional development and behaviour?
- What is the prevalence of alcohol consumption in UK pregnant women?
- Are the risks of fetal alcohol exposure at levels below dysmorphology contingent upon other prenatal risks and/or postnatal risk environment?
- Are the behavioural and cognitive sequelae of overt FAS modifiable? Are treatment implications different from non-FAS?

- What are the reasons for the large differences between the UK and USA in rates of FAS and FASD?
- Are preventive and treatment interventions effective?
- What is the contribution of prenatal alcohol exposure to neurodevelopmental disorders and neurobehavioural functions?
- What is the prevalence of FAS in the UK?

General introduction

Context of the review

In May 2005, the Department of Health (DH) commissioned the National Perinatal Epidemiology Unit (NPEU) to undertake a review of the existing evidence on the effects of alcohol on a developing embryo, fetus and child with a particular focus on evidence regarding low to moderate alcohol consumption. This review is part of the DH funded programme of work at NPEU. Funding for this comes from the Policy Research Programme (PRP), which is part of the Department of Health's Research and Development Directorate. Through funding high quality research, the PRP aims to provide a knowledge base for policy covering health, social services and public health.

The main purpose of the review is to update what we know, from existing evidence, about the effects of prenatal alcohol exposure - particularly focusing on low-to-moderate levels of alcohol and binge drinking. This is important for strengthening the evidence base for health promotion messages about alcohol consumption for women who are planning to get pregnant and during pregnancy. The review is also expected to be a valuable mechanism for identifying research gaps and priorities, and for articulating research questions in this complex area. In so doing, it provides guidance to DH and to other research funders about future priorities for research, should funding be available.

Structure of review

Work started in June 2005 with the objective of submitting a final report to the DH by late January 2006. Given this timescale and the considerable volume of research on the fetal effects of alcohol, we had to be selective about the evidence that we were able to review. It was agreed to concentrate on the evidence for harmful effects at low-to-moderate levels of prenatal alcohol consumption and on binge drinking since these represent areas which have been recently highlighted for public health and policy relevance. We carried out a systematic review of this area.

In order to carry out this review, we formed an advisory group who reviewed the protocol for the systematic review, and the draft final report. The advisory group met once in London in December 2005 to consider the future UK research needs in this area.

Therefore, this final report to the DH consists of the following:-

a) A background section on the context of prenatal alcohol exposure which introduces the reader to the key issues around prenatal alcohol exposure.
b) A systematic review of the available evidence from human observational studies on the effects of low-to-moderate level alcohol consumption and binge drinking during pregnancy on the developing embryo, fetus and child. This includes the conclusions of the systematic review and the key research questions prioritised by the advisory group.
c) A series of appendices to the report.

Part a

Prenatal Alcohol Exposure

1 Background and Context

The material presented in this background section provides the broad context for the ensuing systematic review. It aims to introduce the reader to the key issues around alcohol and pregnancy and includes sections on fetal alcohol syndrome and fetal alcohol spectrum disorders. It also identifies some of the gaps in knowledge about the effects of low-to-moderate prenatal alcohol exposure.

This background material is adapted from a briefing paper originally prepared for the the advisory group. The material raises and summarises important issues but it is not a systematic review.

2 Methods

A computerized search on PubMed using the terms ‘fetal’ and ‘alcohol’ was run to identify review papers published between 2000 and 2005. This was supplemented by Abel’s book length review of Fetal Alcohol Abuse Syndrome (Abel 1998); by the recently produced guidelines for referral and diagnosis of FAS from the Centers for Disease control and Prevention (Bertrand et al 2004); and by relevant review articles from the alcohol review journal Alcohol Research and Health. Further reviews were located from reference lists in retrieved papers and from suggestions by members of the advisory group. Data on alcohol consumption were extracted from survey reports on three UK cohorts (Rickards et al 2004, Sproston et al 2005, Dex and Joshi 2005). This paper is a narrative review of these findings.

3 Outline of document

This paper is structured as follows:

Review of findings on alcohol consumption during the periconceptional period and during pregnancy.

- Description of the terminology and classification of fetal alcohol syndrome and fetal alcohol spectrum disorders and the diagnostic criteria for the different subtypes.
- Summary of the epidemiology including prevalence of fetal alcohol syndrome and the associations between prenatal alcohol exposure and putative fetal alcohol effects from human observational studies.
- Review of the limitations of human observational studies.
- Summary of the evidence from experimental studies in animals.
- Outline of the paternal contribution to fetal alcohol effects.
- Summary of strategies for the prevention of prenatal alcohol exposure.
- Finally, there is a section on new developments in neuroimaging and ultrasound.

4 Alcohol consumption during the periconceptional period and during pregnancy

Most women who drink before they become pregnant either stop drinking or reduce the amount of alcohol they consume substantially once they know they are pregnant. For example, in a Scottish study, 70% of women reduced their alcohol consumption after becoming aware they were pregnant (Plant 1984). However, the ‘at risk’ group for fetal alcohol effects includes not only women

who know they are pregnant but also women who do not know they are pregnant. Unintended pregnancy may not be apparent until four or more weeks after conception, by which time damage from prenatal alcohol exposure may have already occurred. In the United Kingdom Millennium Cohort Study, a sample of 18000 pregnancies resulting in liveborn children in 2001-2002, 58% of mothers said their pregnancy was unplanned. The proportion of unplanned pregnancies varied by age with 84% of those mothers under 20 reported as having unplanned pregnancies (Dex and Joshi 2005).

Therefore, the drinking trends of interest are amongst fertile women. As we have no knowledge of which women are fertile, the best proxy for the 'at risk' group is women in the childbearing age group i.e. 15-45 years of age.

4.1 Measuring alcohol consumption

A key methodological issue is that measurement of alcohol consumption is inherently difficult and therefore findings are by nature imprecise. The difficulties involved in measurement may lead to misclassification of consumption level which may result in bias in observational studies which associate alcohol with poor health outcome (Dufour 1999). Most measurements of alcohol consumption are based on maternal self report. This is often unreliable due to biases resulting from poor estimation, poor recollection and the social undesirability of heavy drinking during pregnancy.

4.1.1 The standard unit and standard drink

Most countries define standard 'drinks' (International Center for Alcohol Policies 1998); however, in the UK we refer to 'units'. A UK unit contains 10 ml of ethanol. Since the containers of all alcoholic drinks indicate both the volume and the concentration of alcohol by volume, a straightforward calculation yields the number of units contained. For example, a 750 ml bottle of white wine at a concentration of 12% contains 0.12 x 750 = 90ml alcohol = 9 units.

Using this formula a pint of beer is approximately equal to two units, and a glass of spirits or a small glass of wine to one unit of alcohol.

However, the situation is not so straightforward. First, the concentration of alcohol in different types of beer and wines varies considerably. Second, the size of a 'glass' of wine in a restaurant or bar varies enormously. Third, standard measures are only used in bars and restaurants - measures poured in the home are likely to differ (Gill 2004).

These variations affect the conversion of self reported drinking measures to units. When combined with the imprecision and bias involved when women try to estimate and recall their alcohol consumption these variations can lead to considerable underestimation of the amount of alcohol actually consumed (Stockwell et al 2004).

When making international comparisons there is further variation. Different countries use different methods to define the standard 'drink' none of which equate to the UK unit. For example, the standard 'drink', size in the US approximates to 1.5 UK units (International Center for Alcohol Policies 1998). Nevertheless in most countries where studies into prenatal alcohol exposure have been conducted, a 'drink' is approximately 12g of alcohol.

4.1.2 Pattern, volume, timing and duration

In addition to the volume of alcohol consumed during pregnancy, the pattern, timing and duration of consumption are also critical to the study of fetal alcohol effects. For example, animal experiments (reviewed in section 8) suggest that brain development is particularly vulnerable during the first and third trimesters. Similarly, facial development is vulnerable in early pregnancy and growth restriction is a feature of exposure at later stages in pregnancy.

Other evidence (including the findings from our systematic review) suggests that, for some effects at least, binge-type exposure, that is, consumption of six UK units or more on one or more occasions, may be more harmful than chronic lower level exposure.

Therefore, as well as collecting data on volume of alcohol consumed it is important to evaluate pattern, timing and duration. Although these aspects of exposure are usually highly correlated, it is extremely useful to disaggregate them where possible.

4.1.3 Data collection methods

The method used to collect data on alcohol consumption will depend on the purpose for which the data are being collected. This affects both the content of the questions asked and the way in which they are asked. In most general population surveys, respondents are asked about consumption in the past year (the reference period) which is taken to reflect current consumption.

Studies during pregnancy tend to use a shorter reference period. For example, in the Seattle longitudinal prospective study on alcohol and pregnancy (Streissguth et al 1981) mothers were interviewed in the fifth month of pregnancy and asked about the quantity, frequency and variability of their drinking during the month prior to recognition of pregnancy and in the month preceding the interview. In a study from Dundee (Florey et al 1992), women were asked on two occasions during pregnancy about their drinking during the previous seven days.

The use of one week as a reference period appears to increase reliability and if asked about recent consumption respondents tend to report higher levels of daily drinking than when asked how much they drink on a typical drinking occasion (Lemmens et al 1992). However, shorter reference periods have the drawback of potentially misclassifying episodic drinkers as abstainers if they have consumed no alcohol during the reference period but had been drinking before this (Kesmodel and Olsen 2001, Dawson 2003).

There has been some research comparing data about current alcohol intake collected by self-administered questionnaire, interview and diary. The findings indicate little to choose between the different methods for determining self reported average intake. However, when the interest is in binge drinking, diaries might be more accurate (Kesmodel and Olsen 2001). The researchers also concluded that the diaries should be completed for a two-week period to minimize misclassification of abstainers as mentioned above. To the extent that diaries may not be feasible in large scale studies, personal interviews have been shown to perform consistently better than questionnaires with respect to binge drinking (Kesmodel 2001, Kesmodel and Frydenberg 2004).

Since researchers may be interested in both binge drinking and average consumption, it seems reasonable to suggest that diaries should be used more often for data collection in pregnancy. However, it is still important to obtain information for a reference period of the four weeks prior to knowledge of pregnancy as the early first trimester may involve a higher level of drinking with particular implications for early effects on development. Whatever methods are used, it is important to maximize the response rate.

4.1.4 Defining levels of consumption

Quantitative definitions of 'heavy', 'moderate', 'low' and 'light' drinking vary between studies of the general population and there is little standardisation. There is even less agreement on what these levels mean in the context of drinking during pregnancy. A review of the definition of 'moderate' drinking concluded that readers of scientific articles need to pay particular attention to the definitions used by researchers but did not suggest any standard definition to be used (Dufour 1999).

To try and avoid the use of vague terms such as 'moderate' the Centers for Disease Control classify 'risky drinking' for childbearing age women as the average consumption of seven or more drinks per week in the past month or five or more drinks on one occasion (Centers for Disease Control and Prevention 2004). This corresponds to 10 UK units per week or 8 UK units on one occasion.

Since it is not yet clear what level, if any, of drinking in pregnancy can be regarded as 'safe' a better classification might examine the empirical data on fetal alcohol effects and then define levels of 'high risk' and 'low risk' drinking. However, variation between individuals in the population means that for some particularly susceptible individuals, perhaps those with a certain genotype, there may be no 'low risk' level.

4.2 Alcohol consumption in women of childbearing age: recent trends in the UK

Two general population surveys, the General Household Survey (GHS) (Rickards et al 2004) and the Health Survey for England (HSE) (Sproston and Primatesta 2005), provide up to date information on UK trends in women's drinking although they do not have specific information on drinking trends in pregnant women. The most recently reported figures are to 2002 and 2003 respectively. As the questions asked about alcohol consumption and drinking patterns have changed over the years some comparisons have only been possible more recently.

The GHS is a multi-purpose continuous survey carried out by the Social Survey Division of the Office for National Statistics (ONS).The survey collects information on a range of topics, including smoking and drinking, from people living in private households in the United Kingdom. Face-to-face interviews are used with a set sample size of 13,250 and an average response rate of 72%.

The GHS uses two measures of alcohol consumption: the average weekly consumption and the maximum daily amount consumed in the last week. The proportion of women drinking over 14 units per week on average increased from 10% in 1988 to 17% in 2002. The increase was most marked in the 16-24 year age group. In this particular age group, the proportion drinking over 14 units a week has increased from 17% in 1992 to 33% in 2002. There has been a 25% increase in the proportion of women drinking more than six units on at least one day in the previous week (from 8% in 1998 to 10% in 2002). Most of this increase has again been in younger women. The proportion of women aged 16 to 24 drinking more than 6 units on one occasion rose from 24% in 1998 to 28% in 2002.

The HSE is a series of annual surveys carried out by the Joint Survey Unit of the National Centre of Social Research and the Department of Epidemiology and Public Health at University College London. It collects information on a range of topics in a core questionnaire including smoking and drinking. The current sample size is around 16,000 adults and 4,000 children.

The HSE also reports on average alcohol consumption and alcohol consumption on the heaviest drinking day for the previous week. The proportion of women who consumed 21 units or more in the previous week increased from 2% in 1993 to 6% in 2003. In women aged 16-24 years,

the corresponding increase was from 9% in 1993 to 21% in 2002. The proportion of women who consumed 6 units or more on the heaviest drinking day increased from 11% in 1998 to 13% in 2003.

Both surveys highlight the same trend: more women are drinking heavily and binge drinking has increased. This trend is most marked in the 16-24 age group. Table 1 opposite illustrates the trends in consumption by age. However, these trends are for all women and not specifically for pregnant women.

Age Group	16-24	25-4	35-44	45-54	55-64	65-74	75+	All ages
1993	9	8	8	8	5	4	3	7
2002	21	8	9	9	7	3	3	9

Table 1. Proportion (%) of women who exceed 21 units per week in the previous week by age in 1993 and 2002. [From Health Survey for England –Trends, (Sproston and Primatesta 2005)]

In the United Kingdom Millennium Cohort Study, around 18,000 mothers were asked to recall their alcohol consumption during pregnancy when the index children were on average nine months old (Dex and Joshi 2005). Around a third of mothers said they consumed alcohol during pregnancy and 82% said they were consuming alcohol at nine months postnatally (Dex and Joshi 2005).

4.3 Biomarkers

A biomarker is an indicator of exposure in a biological sample. To be useful in research or clinical practice a marker should be easily obtainable with minimal discomfort and be both a sensitive and specific measure of the exposure.

For prenatal alcohol consumption, most of the biomarkers available are qualitatively associated with heavy alcohol use or abuse and are not likely to be useful for quantifying the levels of consumption in populations of women. Examples are elevated liver enzymes and red cell mean corpuscular volume greater than 98 fL, which are both obtainable from a venous blood sample. These biomarkers may be very useful in the clinical situation, particularly in assisting with the identification of women who are abusing alcohol.

In contrast to single markers, the use of a series of markers appears more promising with the potential for greatly improved sensitivity and specificity. A recent review described some potentially useful biomarkers that are in development, but concluded that more research was needed to validate them (Bearer 2001). These markers included fatty acid ethyl esters available from neonatal hair and meconium.

Although the development of biomarkers is intuitively appealing (given the shortcomings of maternal self report) much more developmental work is required before they can usefully be applied in epidemiological studies. When and if such biomarkers do become available, it is much more likely they will augment rather than replace maternal self report as a measure of exposure.

In the meantime, the use of blood samples to detect levels of gamma glutamyltransferase, carbohydrate-deficient transferrin and mean corpuscular volume may be useful in the confirmation of problem drinkers but are unlikely to be of much value in universal screening of pregnant women (Bearer 2001) to identify lower levels of alcohol exposure.

5 Terminology of fetal alcohol syndrome and fetal alcohol spectrum disorders

The term ‘fetal alcohol syndrome’ (FAS) was coined in 1973 by Jones and Smith (Jones and Smith 1973) to describe their clinical findings in a series of eight children born to chronic alcoholic mothers in the United States. The principal clinical features included minor but characteristic abnormalities of the face, pre- and postnatal growth restriction/retardation, and severe neurodevelopmental problems. Following this initial publication, FAS was reported in a wide variety of countries including the United Kingdom (Beattie et al 1983, Halliday et al 1982).

There is now general acceptance that FAS is a complex multi-factorial disorder in which teratogenic exposure to heavy alcohol consumption interacts with other environmental factors and genetic predisposition. However, two developments have complicated this initial picture: first, the existence of partial forms of the syndrome and second, evidence that harm may occur with levels of prenatal alcohol exposure within a more ‘moderate’ range. These developments have led to changes in terminology that have caused much debate (Abel 1998).

Following the original clinical description, it soon became apparent that only 4-5% of the children born to women who consumed large amounts of alcohol in pregnancy showed the ‘full-blown’ syndrome. However, many more showed partial features (Abel 1998) of FAS. The initial term used to describe these partial effects was ‘possible fetal alcohol effects’. This term indicated that the effects were observed more often in children exposed to heavy prenatal alcohol consumption than those not exposed. In addition, the qualification ‘possible’ reflected suitable caution about attributing causality given that similar effects could be seen in children who were not prenatally exposed to alcohol.

Further research showed that these ‘possible fetal alcohol effects’ also occurred more frequently than expected in children whose mothers were neither alcoholic nor heavy drinkers but who were more ‘moderate’ in their consumption (Streissguth et al 1981).

Because of these developments, the range of adverse effects of prenatal alcohol exposure on the developing embryo, fetus and child were construed as a spectrum of structural abnormalities and growth and neurodevelopmental impairments. The current preferred term to encompass all these effects is ‘fetal alcohol spectrum disorder’ (FASD) (Sokol et al 2003).

US data suggest that while FAS seems comparatively uncommon with a prevalence of between 0.5 - 2 per thousand live births (May and Gossage 2001), FASD may be relatively common with a prevalence of 9 to 10 per thousand live births (Sampson et al 1997). Around 80% of these cases are children showing neurodevelopmental disorder only. This subtype of FASD is known as ‘alcohol-related neurodevelopmental disorder’ (ARND). It is this disorder (ARND) that has the most public health importance, given its prevalence, and the suggestion that it may be associated with moderate levels of prenatal alcohol exposure.

6 Diagnosis and classification of fetal alcohol spectrum disorders

There is not yet a diagnostic test for FASD such as a blood test or MRI scan. Therefore, diagnosis remains based on history and clinical examination. Making a diagnosis of FASD involves applying a set of clinical criteria and eliciting or assuming a history of prenatal alcohol exposure. Different sets of criteria have been proposed but all concentrate on the triad of signs (specific facial features, growth restriction/retardation and neurodevelopmental disorder), not all of which need to be present.

Broadly speaking, the diagnostic criteria enable the physician to assign the child to one of a set of categories within FASD. These include FAS (all three signs), partial FAS, alcohol-related birth defects (ARBD), and alcohol-related neurodevelopmental disorder (ARND).

The two main diagnostic classification systems in current use in the US are known as the Institute of Medicine (IoM) Criteria (Stratton et al 1996) and the Centers for Disease Control and Prevention (CDC) Criteria (Bertrand et al 2004). Recently a revision of the IoM Criteria has been proposed (Hoyme et al 2005). These two sets of criteria are intended for use in the clinical setting and may require modification for other activities such as research, screening and surveillance.

The main elements of the revised IoM system are:

- Maternal alcohol exposure should be either confirmed or unknown for FAS and partial FAS
- For ARBD and ARND confirmed alcohol exposure must be present
- Confirmed maternal alcohol exposure is defined as substantial regular intake or heavy episodic drinking
- The diagnosis assumes that assessment has ruled out syndromes with similar features e.g. Williams, DeLange and velocardiofacial syndromes.

In addition to a confirmed history of heavy maternal alcohol use, the revised IoM criteria indicate that to diagnose ARND the following are necessary:-

- Either evidence of structural brain abnormality or microcephaly (head circumference <10th centile), OR
- Presence of a complex pattern of behavioural or cognitive abnormalities which are inconsistent with developmental stage and which cannot be explained by genetic predisposition or postnatal environment.
- This complex pattern is characterised by impairment in performance on complex tasks combined with higher-level receptive and expressive language deficits and disordered behaviour.
- Therefore, IQ tests alone are insufficient to make an assessment. Tests of executive functioning, communication and behaviour are also necessary.

The CDC criteria are similar to the revised IoM criteria, differing mainly in the requirements for ARND. CDC does not use the term ARND but instead uses the phrase ‘central nervous system (CNS) abnormalities’. Furthermore, rather than formulating strict criteria for these CNS abnormalities, CDC proposes more general guidelines. CNS abnormalities can be structural, neurological or functional and an individual can present with one or more of the following:-

- Structural: the presence of either
  - Microcephaly (head circumference <10th centile)
  - Neuroimaging evidence of structural brain abnormality
- Neurological: the presence of either
  - Seizures not due to postnatal insult or fever
  - Soft neurological signs outside normal limits
- Functional: the presence of either
  - IQ<70
  - Functional deficits: one standard deviation below the mean on standardised tests in three domains out of six. These domains are cognitive/developmental; executive functioning; motor functioning; attention or hyperactivity; social skills; others including sensory problems, pragmatic language problems, memory deficits etc.

The CDC criteria also quantify 'significant prenatal alcohol use' as seven or more drinks per week (10.5 UK units) or three or more drinks (>4.5 UK units) on multiple occasions or both.

Although diagnostic services for FASD commonly provide assessment by a trained paediatrician or dysmorphologist, it is clear from the above that a comprehensive neuropsychological evaluation will be necessary as well. This is not only to establish whether diagnostic criteria are met but also to provide information relevant to the future management of the child. It has therefore been argued, in a United States context, that FASD diagnosis should move away from the dysmorphology services to neurobehavioural clinics (Burd et al 2003).

These two sets of criteria above represent an advance on previous diagnostic systems, but the particular criteria for the ARND/CNS abnormalities are problematic because of a lack of specificity.

The available evidence suggests that if other phenocopies can be ruled out, 'full-blown' FAS is fairly specific to prenatal alcohol exposure. However, neurodevelopmental disorders are seen both in children of drinkers and non-drinkers. FAS seems to be specifically associated with prenatal alcohol exposure but neurodevelopmental disorders are associated with a wide range of exposures such as preterm birth, genotype, psychosocial adversity and so on. Therefore, the presence of a neurodevelopmental disorder (without the facial features or growth restriction/retardation) in a child exposed to heavy prenatal alcohol does not imply causation. The alcohol may well have been a factor but it could have been incidental.

There is currently considerable debate about whether a syndrome-specific neurobehavioural profile can be associated with prenatal alcohol exposure. On this point, the IoM criteria are much more specific. They construe ARND as specifically comprising executive dysfunction, higher level receptive and expressive language deficits and disordered behaviour. The CDC criteria, on the other hand, are much less restrictive reflecting the degree of uncertainty in current knowledge. The matter remains to be settled. The results seem to depend on the range of domains assessed. Many studies focus on a narrow range of outcomes. However, when a broad range of outcomes is studied in detail, specific impairments seem less likely to be identified. For example, a recent longitudinal study by Korkman and colleagues using a comprehensive system of neuropsychological testing (the NEPSY) testing a broad range of domains revealed non-specific widespread and generalized impairments in children prenatally exposed to greater than 140 grams (17.5 UK units) of alcohol per week for varying durations (Riley et al 2003).

One implication of the above is a need for caution in attributing causation at an individual level for a child with neurodevelopmental disorder. It is possible that making a diagnosis of alcohol-related neurodevelopmental disorder might produce lifelong anxiety and guilt in the parents, stigmatize the child, and do little to help prevention.

Therefore, from a UK perspective, it could be argued that identifying those children with 'full blown' FAS would be a useful first step, while continuing to study the role of prenatal alcohol amongst other factors in the causal pathways to neurodevelopmental disorders. Concentrating efforts on identifying the 'full blown' FAS would allow researchers to better characterize the neurodevelopmental problems these children face and allow clinicians to offer appropriate services. One could then investigate the extent to which similar features are seen in children prenatally exposed to high levels of alcohol but without any accompanying physical features.

7 Overview of human observational studies investigating fetal alcohol effects

Both descriptive studies of FAS prevalence and analytical studies including case control and cohort studies have been used to investigate the epidemiology of fetal alcohol effects.

These studies have mainly taken place in the US with a few studies in the UK, continental Europe, Canada, Australia and South Africa. The findings suggest that alcohol affects the fetus in a dose-dependent fashion with structural abnormalities predominating at high doses and more subtle neurodevelopmental problems predominating at lower doses. Drinking pattern as well as volume, the duration of drinking and the timing during pregnancy are all important. The populations most at risk appear to be those living in multiply deprived inner cities in the US, who may also have poor nutrition and concomitant illicit drug use. In general, at risk women are older mothers, have severe alcohol problems and smoke cigarettes. However, fetal alcohol effects also occur in women outside of this high-risk group. Genotype appears to be important in susceptibility to the effects. The epidemiological studies in this area are subject to the same problems (bias, confounding and chance) as all observational studies but measurement error and residual confounding are particularly important.

7.1 Descriptive studies of prevalence

Following the convention for other congenital anomalies, we describe the measure of disease frequency for FAS as prevalence rather than incidence.

Measuring prevalence relies on accurately identifying those with FAS - the ‘cases’. However, the features of FAS may not be initially apparent at birth and may change as the child develops. The neurodevelopmental problems are best assessed after infancy and appear to persist into adult life. Similarly, the facial features become more apparent from around age eight months but then tend to become much less apparent in later childhood. Therefore, the most accurate period for diagnosis is between three and twelve years of age. However, by age three, the child may have been adopted, the natural mother may have died and therefore it may be difficult to obtain details on alcohol and other exposures (see presentation by Margaret Barrow).

Reported prevalence estimates of FAS vary widely depending on the population studied and the surveillance methods used. Three methods have been employed. First, passive surveillance examines birth certificates, case records or congenital anomaly registers for geographically defined populations. Second, active case ascertainment involves actively finding, recruiting and examining children in a geographically defined population. Third, clinic based samples are recruited, usually at antenatal clinics, and the offspring followed up prospectively. The most accurate (unbiased) estimates are likely to come from active surveillance methods (May and Gossage 2001).

Best estimates of prevalence for the US suggest that FAS has a prevalence of between 0.5 and 2 per 1000 live births (95% CI 0.2 - 7.2) with estimates of alcohol-related birth defects and alcohol neurodevelopmental disorder each around 10 (95% CI 4.8 - 18.3) per 1000 live births (May and Gossage 2001).

The highest prevalence in the world has been reported from a winemaking region of South Africa with FAS prevalence of 50 per 1000 (95% CI 37.3 -65.4) live births (May et al 2000).

In the UK, there have been no active case ascertainment studies. There have been two passive studies one of which calculated a prevalence rate and five clinic-based studies.

Beattie and colleagues (Beattie et al 1983) in Glasgow reported on 40 children with FAS born in the West of Scotland between 1971 and 1981. In the year 1980-81 alone there were 22 cases diagnosed. A prevalence estimate was not reported. Halliday and colleagues (Halliday et al 1982) in Belfast identified 23 babies with varying features of FAS. They estimated a prevalence of 1.7 per 1000 births (95% CI 1.0 -2.7).

Five antenatal clinic-based studies from the UK (Wright et al 1983, Plant 1984, Waterson and Murray-Lyon 1993, Primatesta et al 1989, Sulaiman et al 1988) have collectively examined a sample of 5771 children born to representative samples of women. No cases of FAS were found (Abel 1998). This is surprising in view of the US findings. Based on a US estimate of prevalence of 2 per 1000 births we might expect to have found around 11 children with FAS. The reasons for this finding are unclear.

We do not know the current UK prevalence of FAS. We know that the most accurate time for diagnosis is between three and twelve years of age, and the most accurate method for estimating prevalence is active case ascertainment. If we could be sure of obtaining accurate data on exposure, a study of all children entering primary school at ages 4-5 in a defined geographical area might be helpful. Studies with this design have been successfully used in South Africa (May et al 2000) and Washington State (Clarren et al 2001). This design also has the advantage of a good sampling frame (school registers) and of yielding an age specific prevalence. The disadvantage is that many children with severe impairment who are in special schools or institutional care might be overlooked. It may be useful to investigate the methods used by those estimating the prevalence of autistic spectrum disorders where a considerable number of children will not be in the mainstream school system (Chakrabarti and Fombonne 2001).

7.2 Analytical studies

Although a few case control studies have been conducted, most of the evidence on fetal alcohol effects has come from several prospective longitudinal cohort studies (mainly conducted in the US) within the past 30 years. These researchers have asked women in antenatal clinics to recall their alcohol consumption levels both before they knew that they were pregnant and at various times throughout pregnancy. They have collected data on a number of perinatal and childhood outcomes at different ages and also collected data on potential confounders and effect modifiers. Unfortunately, there has been little uniformity in exposure measurement, outcomes studied or confounders measured. This means that it has been difficult to compare study results. Many of these studies were reviewed by the Institute of Medicine and by Abel in two lengthy reports (Abel 1998, Stratton et al 1996). These reports were published 8 -10 years ago and so they do not include the more recent findings. There is a need to update these reviews to take into account the more recent findings.

Many findings have not been replicated. Those which seem either fairly well established as well as some which seem important but in need of further clarification are summarized below.

1. Heavy maternal alcohol use, especially that associated with alcohol dependence or severe alcohol problems, is associated with FAS, that is, the triad of specific facial features, growth restriction/retardation, and neurodevelopmental problems.
2. Women most at risk of having a child with FAS are those living in deprivation, aboriginal women, women who smoke or use illicit drugs during pregnancy, older mothers and women with poor nutrition during pregnancy.
3. The risk of having a subsequent FAS child, having had one, is 800 times the baseline risk of having a FAS child (Abel 1988).
4. Around 25% of those diagnosed with FAS have an IQ score below 70 and virtually all have problems with attention and behaviour (Streissguth 1991). An IQ score below 70 is more than two standard deviations below the mean for the population and 70 is the upper limit of the range for mental retardation (also known as generalized learning disability).
5. Spontaneous abortion, stillbirth, preterm birth and small for gestational age at birth have all been associated with heavy maternal alcohol use but not consistently.

6. The effects of prenatal alcohol consumption at moderate and low doses have been more controversial. An expert committee from the US Institute of Medicine in 1996 was inconclusive. They investigated "…data on the relation between low or moderate levels of prenatal alcohol exposure and more subtle abnormalities associated with such exposure, but it was unable to conclude that these subtle abnormalities, as detected by statistical calculations from epidemiologic studies of defined populations, do or do not represent a distinct clinical entity" (Stratton et al 1996). Abel, one of the leading experts in the field of fetal alcohol syndrome clearly believes that effects are only apparent after heavy or abusive drinking and that effects seen at low and moderate doses are artefacts due to confounding and underestimation of exposure (Abel 1998). Evidence from two meta analyses on moderate drinking (defined as between two drinks per day and two drinks per week) (Polygenis et al 1998, Makarechian et al 1998) and an expert review of moderate drinking by the National Institute on Alcohol and Alcohol Abuse (Gunzerath et al 2004) also suggest that moderate drinking is not associated with deficits in growth or birth defects. One of these meta analyses showed an increased risk of spontaneous abortion and no effect on preterm birth. The risk of stillbirth was actually reduced by moderate drinking. However, a more recent study has found a 3-fold increase in risk of stillbirth in women drinking five or more drinks weekly (Kesmodel et al 2002).
7. In contrast to point 6 above, more recently conducted studies of the effect of prenatal alcohol exposure of 1-2 drinks per day have found impairments in attention, speed of information processing, intelligence, learning and memory. These effects seem to persist through adolescence (Jacobson and Jacobson 1999). Furthermore, there is some evidence that even low (up to 3-4 drinks per week) amounts of alcohol may have adverse effects on neurodevelopment (Sood et al 2001). However, the amount of the variance that is uniquely explained by the prenatal alcohol exposure at this dose is between 0.6 and 2% (Huizink and Mulder 2006).
8. Drinking pattern as well as volume, the duration of drinking and the timing during pregnancy are important. Most important of all may be binge drinking. There is strong plausibility for the effects of binge drinking in heavy drinkers (Streissguth et al 1993), but it is difficult to disentangle the effects of the bingeing from the effects of a high consumption level. There is insufficient evidence to quantify any potential harm from episodic bingeing in otherwise infrequent drinkers.
9. Polymorphisms of the gene coding for the alcohol dehydrogenase enzyme ADH1B contribute to susceptibility to FAS (Warren and Ting-Kai 2005).

These findings should be interpreted in the light of the methodological problems outlined below.

7.3 Methodological issues in design and analysis

In common with epidemiological studies in general, several methodological problems can affect interpretation of observational studies on fetal alcohol effects. Two are of particular importance: ascertainment of degree of exposure and residual confounding. However, other issues which may cause problems such as lack of power and simultaneously investigating multiple associations between exposure and outcome are also common.

7.3.1 Ascertainment of degree of exposure to alcohol during pregnancy

The measurement of alcohol consumption is still rather imprecise with numerous methodological pitfalls involved in data collection (see section 4.1). There has been a focus on trying to establish an average or typical volume of consumption and a relative neglect of the variation in consumption from day to day in an individual. Therefore women may be asked to estimate how much they drink on average or else their reported consumption is averaged as a daily or weekly amount. This is

understandable, as it is easier to analyse and report averaged results rather than variability in pattern. However, these averages can conceal the true consumption pattern. Abel gives a good example of this from a study by the Jacobson group (cited in Abel 1998). The Jacobson group placed a threshold for alcohol related fetal damage at an average consumption level of one US drink per day (12g). However, the group acknowledged that very few women actually drank in this way. They tended to concentrate their alcohol consumption into a small number of occasions. Those women who drank above this threshold consumed a median of six drinks per occasion (72g). This is equivalent to drinking a whole bottle of wine at one sitting.

Furthermore, many women underestimate how much they actually drink, in some cases intentionally (Abel 1998). The net effect of this averaging and underestimation is misclassification of exposure level. A study using mathematical modeling showed that this misclassification could lead to a bias in the estimate of effect either towards or away from the null value (Verkerk 1992).

7.3.2 Residual confounding

Confounding represents the mixing together of the effects of (a) an exposure with (b) a factor statistically associated with exposure and causally associated with outcome. For example, if there is an apparent effect of alcohol consumption in pregnancy on growth restriction in the fetus, this apparent association may be caused by another factor (the confounder), for example smoking. Smoking is associated with alcohol consumption and smoking is associated with intrauterine growth restriction. Controlling for the effect of smoking may therefore remove any apparent association between alcohol exposure and intrauterine growth restriction.

Residual confounding is the confounding that remains after attempts to adjust completely for confounding. These attempts are usually unsuccessful for two reasons. First, an important confounder may have been omitted from the analysis or data collection. An example of this would be failing to adjust for smoking when measuring the effect of prenatal alcohol exposure and birth weight. Second, there may have been failure to measure the confounder accurately leading to misclassification. This is a particular problem when trying to adjust for psychosocial factors. For example, psychosocial factors are often poorly measured in studies leading to residual confounding after apparent adjustment (Macleod et al 2001).

The effects of residual confounding are generally not predictable and can either mask a true association or else create a spurious association.

The two most important types of confounding of effects on neurodevelopmental outcome are failure to control for the postnatal environment and failure to control for factors which are strongly genetically influenced and which may be related to both prenatal alcohol exposure and the outcome. For example attention deficit hyperactivity disorder (ADHD) is a disorder which has high heritability and is associated with both increased alcohol consumption and FAS. This is not usually measured or controlled for in studies and it may therefore confound effects. In fact, Riley, a behavioural teratologist has noted:-

“Currently, it is unclear to what degree the effects seen in children prenatally exposed to alcohol are influenced not by the teratogenic effect of alcohol, but rather by genetic factors related to alcohol problems in either parent” (Riley 2004).

The same could be said about many aspects of the postnatal environment.

7.3.3 Alternatives to cohort and case control studies

The findings from the human observational studies have suggested an association between prenatal alcohol exposure and a number of adverse effects. Taken together with the animal experiments they provide convergent evidence.

Nevertheless, we may have reached the limits of what we can determine from the standard case control and cohort designs. Teasing apart the relative contributions of exposure and confounding variables and trying to adjust for genetic influences is likely to require the application of study designs that are new for this particular research area. These include, for example, the use of populations where heavy alcohol use is not strongly correlated with adverse postnatal environment and the use of twin study designs which can take account of genetic influences.

An example of the latter is a recently published study (Knopik et al 2005) on 1,936 Missouri adolescent female twin pairs born between 1975 and 1985. Knopik and her colleagues investigated the contribution of prenatal cigarette and alcohol exposure to the risk of developing attention deficit hyperactivity disorder. Twin studies essentially compare monozygotic twins who have 100% of their genes in common with monozygotic twins who share 50%. Making certain assumptions about the environment which the twins share and do not share, it is possible to apportion variation between twins to shared environmental factors, environmental factors which are not shared and genetic factors. In this study the researchers were able to use the twin design to determine the relative contributions of prenatal exposures including alcohol and smoking, parental alcoholism and genetic factors. The effect of the prenatal exposures was relatively small in comparison with genetic effects.

However, the likely aetiological heterogeneity of many neurodevelopmental disorders means that although the overall effect of prenatal alcohol exposure may be small, in certain subgroups it may play a very important role. For example, an association between attention deficit hyperactivity disorder and polymorphism of the dopamine transporter gene has been known for some time. However, a recent study of children with attention deficit hyperactivity disorder found that this association only held for the subgroup of children who were prenatally exposed to alcohol (Asherson et al 2005).

There has only been one small twin study on FAS (Streissguth 1993) in which five out of five of the monozygotic twin pairs were concordant for diagnosis compared with seven out of eleven dizygotic twin pairs. These twins were ascertained from FAS clinic populations. The twin design would seem to have considerable potential, but perhaps more to look at the contribution of prenatal alcohol exposure to common childhood psychiatric disorders or measure traits such as IQ rather than FAS, which is likely to be rare even in large twin studies.

Other designs which are being increasingly used to test for the effects of environmental exposures in psychiatric disorders may also be of value (Rutter et al 2001).

8 Key points from animal experiments

As indicated in the previous section, the evidence from human observational studies is consistent but prone to bias and confounding. On its own, it would be persuasive but not compelling evidence. However, when combined with the evidence from animal experiments there is then a strong case for alcohol as a human teratogen (Abel 1998). In particular:

- Animal experiments have strong face validity: virtually all the human anomalies associated with fetal alcohol exposure have been reproduced in experimental animal models.
- Comparable behavioural effects to those seen in humans have been demonstrated in animals including motor overactivity, impaired learning, and attentional problems.
- Fetal alcohol effects have been produced in a number of different species.

- Dose-response relationships have been established for most of the effects. For all outcomes studied, peak blood alcohol levels appear to be more important than length of alcohol exposure.
- Timing of exposure during pregnancy appears to be important. In particular, the periods corresponding to the first and third trimesters in humans appear to be sensitive periods for inducing CNS abnormalities.
- Results have been achieved using experimental control of both exposure and confounders.

8.1 Advantages of experimental studies

In experimental studies, exposure and potential confounding factors can be controlled by matching and random assignment.

Exposure to alcohol is usually achieved by injection of predefined doses into a tube inserted into the animal's stomach. Each animal in an experiment can receive the same dose at the same time and control of the dose means that high blood-alcohol levels can be sustained or else dose-response relations can be studied.

Confounders such as calorific value of alcohol can be adjusted for by matching the diet of alcohol-exposed mothers to controls. This is called pair-feeding. Even confounders such as the postnatal environment can be controlled by cross-fostering the offspring of exposed and control mothers.

8.2 Main findings

Findings using several species of mammal including primates and rodents as well as non-mammalian models show consistent effects of high doses of alcohol on the fetus. These effects occur through multiple mechanisms depending on the dose, pattern, and timing of the exposure. Research conducted more recently suggests that effects on the CNS may occur at levels of exposure that are much lower than previously thought and which correspond with low consumption in humans.

Several species have been used in the study of fetal alcohol effects. Rodent models have been most often studied as there is such a lot known about their physiology, genetics and behaviour. One of the earliest experiments involved the demonstration that alcohol affected craniofacial development in the mouse (Sulik et al 1981). The craniofacial anomalies produced included microcephaly, microphthalmia, short palpebral fissures and a long upper lip with deficient philtrum. These effects can be produced with maternal blood alcohol levels around 200 mg/dl which, in a human, can be achieved even at 'moderate' consumption levels. Studies are now looking at agents that can potentiate or ameliorate these craniofacial effects of alcohol (Sulik 2005).

The period of greatest brain growth in the human pregnancy is the third trimester. This is equivalent to the early postnatal period in rats. Therefore, a number of studies have investigated alcohol exposure during the early neonatal period in this species. Numerous studies have reported that heavy alcohol exposure at this time reduces brain weight and volume, particularly the forebrain, brainstem, corpus callosum and cerebellum (Chen et al 2003). Much of this reduction in volume is explained by neuronal loss in areas such as the cerebellum, cortex and hippocampus.

Non-human primates have not been studied as extensively as rodents. However, Schneider and colleagues in Wisconsin have conducted studies on Rhesus Monkeys exposed to the equivalent human dose of alcohol of between 1-2 drinks daily throughout pregnancy. This dose affected attention and neuromotor functioning in the early postnatal period but not birthweight, gestational length, and facial dimensions (Schneider et al 1997).

Apart from other mammals such as pigs, non-mammalian species such as chick embryos and tadpoles have also been used. For example, recent studies on Xenopus embryos are providing useful information into the mechanisms of alcohol teratogenesis at a molecular level (Peng et al 2004).

These mechanisms are of interest, not just because they may explain the causal pathways, but also because they may suggest potential interventions. It has become clear that alcohol is non-specific in its actions: it operates via a number of mechanisms and more than one mechanism may be at work in producing a given effect. For example, alcohol is able to disrupt processes involved in cell adhesion, gene expression, growth factor activity and neurotransmitter systems. However, it has recently become clear that alcohol may exert many of its effects by a final common pathway of increasing apoptosis, that is, programmed cell death. Apoptosis happens as an essential part of normal development, for example in the nervous system a certain amount of 'pruning' seems to be essential. Yet alcohol may upset the balance leading to increased or unnecessary cell death. The effect of this may depend critically on the timing of exposure, that is, on developmental stage (Goodlett and Horn 2001).

These different mechanisms may also be activated at different blood alcohol levels. There has been a great deal of interest in some recent studies in neonatal rats and mice that appear to show a significant apoptotic response in the brain after only a transient and relatively low dose of alcohol. Transient blood alcohol levels in the range of 80 mg/dl for around 60 minutes were sufficient to produce neuroapoptosis at a higher rate than saline treated controls (Ikonomidou et al 2000, Farber and Olney 2003). The significance of these results in the longer term and their importance for human development remains to be determined.

8.3 Limitations

Although animal experiments have contributed greatly to this field, they do have their limitations. The complex postnatal environment in which humans are raised and the complexity of human behaviour are poorly approximated in animal models. Therefore, the strength of the animal studies is in the support they give to the human observational studies: the convergence of evidence. There is an extensive literature on animal experiments in this field but no systematic review. A systematic review may be difficult to perform, however, since negative results in animal research are generally not prepared for publication (Lemon and Dunnett 2005).

9 Paternal contribution to fetal alcohol effects

There is a great deal of evidence to show that paternal alcoholism and heavy alcohol consumption affect subsequent child development and behaviour, but few studies have investigated the effects of alcohol consumption in the father on the risk of FAS.

It would be interesting to find out whether similar abnormalities to FAS are seen in children where the father consumes alcohol heavily but the mother does not drink during pregnancy (Rutter 2005).

If such abnormalities were present, then this would suggest three possibilities. First, that factors other than prenatal alcohol exposure, for example genetic influences transmitted via the father, are involved. Second, that the effects are due to the postnatal environment: that is the psychosocial adversity associated with the father's continued drinking after pregnancy. Third, that the alcohol affected the father's sperm to produce a non-genetic heritable abnormality, that is, epigenetic transmission (Holliday 1998).

As far as we are aware, no study has compared outcomes in different groups depending on the presence or absence of heavy drinking in mothers and fathers. However, this type of study would be difficult to conduct because of assortive mating: heavy drinking women and men tend to associate with one another. A further problem to consider is that the assumed father may not be the actual father, that is, non-paternity.

However, some studies have investigated paternal exposure while adjusting or stratifying for maternal consumption in the analyses. These studies have addressed perinatal outcomes, chiefly birthweight, preterm birth and spontaneous abortion. The conclusion overall is that there is no evidence of a significant paternal contribution. There have been no studies focussing on neurodevelopmental outcome.

In contrast to the human studies, animal studies have provided some evidence of effect. For example, rats sired by males exposed to alcohol before the pregnancy have greater difficulties than controls in learning certain tasks, deficits in spatial memory and increased motor activity (Abel 2004).

In a study to investigate epigenetic effects, male rats treated with alcohol for nine weeks before breeding had decreased cytosine methyltransferase mRNA levels in paternal sperm, compared with controls. The researchers suggested that the alcohol consumption might have resulted in reduced DNA methylation leading to the expression of normally silent paternal alleles (Bielawski et al 2002).

10 Prevention

As we are not certain either how much prenatal alcohol exposure results in FAS or which particular pregnancies are going to be affected, the only certain way to achieve complete prevention of FAS is to ensure no exposure to alcohol in pregnancy at all. Thus abstinence during pregnancy has been recommended by a number of authorities including the US Surgeon General (Surgeon General 2005).

However, others have pointed out that recommending abstinence goes beyond our current evidence base and may have its own adverse effects (Abel 1998) such as producing prenatal anxiety and guilt which in turn may have a negative impact on subsequent child development (O'Connor 2002). It also stigmatizes mothers whose drinking may not have harmed their child. The issue is a contentious one and will not be further addressed here.

There are three complementary strategies for preventing fetal alcohol effects: universal, selective and indicated prevention (Hankin 2002). Universal prevention strategies are aimed at the general population and usually consist of educational messages about fetal alcohol effects. Selective prevention strategies are aimed at women of childbearing age and include screening of pregnant women for alcohol use followed by advice or treatment if required. Indicated prevention strategies are aimed at defined high-risk group such as women with alcohol problems or a previous birth with FASD.

The universal approach has two main strands, education and policy. Educational strategies include advertising campaigns, labeling alcoholic beverages with warning statements and education in schools. These strategies have been popular but ineffective: they have improved knowledge and influenced attitudes but have failed to change behaviour (Room et al 2005). That is not to say that educational strategies should be discontinued. On the contrary, giving health advice and education to the population might arguably be a public good in itself, even if it does not seem to alter behaviour. Furthermore, educational strategies might be designed differently to target the behaviours more effectively.

The main policy approach is to control the price and availability of alcohol. This is effective at reducing population mean alcohol consumption, the number of problem drinkers and the incidence of liver cirrhosis. It could be expected to have an effect on drinking during pregnancy as well.

There are already some measures in place impacting on price and availability of alcohol, but furthur use of such levers although potentially effective is currently unpopular (Room et al 2005).

Targeting women at risk of (usually heavy or problem) drinking during pregnancy has used screening tools followed by brief counselling interventions. Screening alone seems to be related to reduction of prenatal alcohol consumption (Chang et al 2000) perhaps by simply raising awareness.

Two similar screening questionnaires have been specifically designed for the prenatal setting; the T-ACE (Sokol et al 1989) and the TWEAK (Chan et al 1993). The T-ACE questions are:

How many drinks does it take to make you feel high? (Tolerance)

- Have people annoyed you by criticizing your drinking? (Annoyance)
- Have you ever felt you ought to cut down on your drinking? (Cut down)
- Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover? (Eye-opener)

The TWEAK questions are similar and include a question on alcohol induced amnesia. For the identification of problem drinkers, both tests have similar sensitivities at around 75% but the T-ACE seems to be more specific at around 90%. However, there is some evidence that the TWEAK may be better at identifying those women who are not problem drinkers but who are nevertheless drinking at a level that may still pose a risk to the fetus (O'Connor and Whaley 2003).

Indicated prevention strategies have targeted women with histories of heavy drinking during pregnancy or who have given birth to a child affected with FAS. While most women with severe alcohol problems require specialist input from an alcohol treatment team, those women whose alcohol problems are not severe can benefit from brief intervention. In brief intervention, the woman is offered assessment, feedback and goal setting using techniques derived from motivational interviewing and cognitive-behavioural psychotherapy (Sokol et al 2003). There have now been several randomized controlled trials using brief intervention that show an effect on reduction or cessation of drinking (Chang 2005). In addition to showing reduction in consumption, a few of these trials also demonstrated improved birth outcomes (Handmaker and Wilbourne 2001) and, in one study, better neurobehavioural outcomes at 13 months of age in children of mothers assigned to the intervention compared to the control group (Hankin 2002). These trials are particularly interesting because they provide some evidence of a causal effect of reduction in prenatal alcohol exposure and improved birth and neurodevelopmental outcome.

Following birth, for many children with FAS the effects from prenatal exposure may be compounded by the multiple risks associated with living in a family where one or both parents may be alcoholic. Therefore, early postnatal intervention may be indicated. This might include treatment of the alcoholism and other comorbid disorders, use of supportive services, provision of an enriched environment, measures to alleviate poverty or sometimes adoption. The use of pharmacological agents such as antioxidants and nutritional supplementation to reduce toxicity or enhance neuroplasticity and recovery is the subject of much new research in animal models but has not yet been tested in humans (Goodlett et al 2005).

11 New developments

Two recent developments using neuroimaging and fetal ultrasound seem promising avenues for future research.

Riley and colleagues in San Diego have been carrying out neuroimaging studies since 1992 in those exposed to high prenatal alcohol consumption both with and without full-blown FAS. The studies of alcohol-exposed individuals consistently show overall reduction in brain size with specific volume reductions or anomalies in frontal lobes, cerebellum, corpus callosum and the basal ganglia (affecting the caudate nucleus) in particular (Riley and McGee 2005). The neuroimaging abnormalities are consistent with the performances on neuropsychological tests. Even more interesting is that the brain changes seen and the test scores are qualitatively the same in exposed children with and without the physical features of fetal alcohol syndrome (Riley and McGee 2005).

Hepper and colleagues in Belfast have used analysis of real time fetal ultrasound images to study the behaviour, perceptual and learning abilities of the fetus. They have examined mouth movements, startle, and habituation to sound in fetuses of non-smokers exposed to relatively low amounts of prenatal alcohol compared to non-smoking abstainers. Their findings indicate significant differences between the exposed and unexposed fetuses on these three behaviours and suggest possible fetal neurotoxicity with as few as three drinks per week during pregnancy on average (Hepper et al 2005, Little et al 2002). These findings are exciting and provide the best direct evidence to date for effects of low dose prenatal alcohol consumption on the human fetus. However, the implications of these findings for significant or enduring problems in the infant and child remains uncertain.

Part b

Systematic Review of Fetal Effects of Low-to-Moderate Prenatal Alcohol Exposure and Binge Drinking

1 Introduction

There is now a vast literature on the effects of maternal alcohol consumption during pregnancy on the developing embryo, fetus and child. It is generally accepted that both abusive and heavy drinking is associated with fetal alcohol syndrome (FAS) and fetal alcohol effects such as growth retardation, birth defects, and neurodevelopmental problems (Abel, 1998). Furthermore drinking on a daily basis (seven or more drinks per week) has been linked to an increased risk of both perinatal and neurodevelopmental problems (Sampson et al, 1994).

The current UK Department of Health guidelines recommend that women who are trying to become pregnant or are at any stage of pregnancy, should not drink more than 1 or 2 units of alcohol once or twice a week, and should avoid episodes of intoxication (Department of Health, 2005). Therefore the focus of research interest has now moved to the evaluation of drinking at low-to-moderate levels and drinking in binges. This is an area of increasing scientific interest and controversy. Since most pregnant women either abstain or drink at low-to-moderate levels, this area is also the one of most practical concern to clinicians, public health practitioners and women themselves.

To address these practical concerns, for this review we focused on studies which evaluated intake of less than seven drinks per week i.e. less than one drink per day. We considered whether an intake of up to six drinks a week is associated with more risk than total abstention; whether there appears to be a ‘safe’ level; and whether binge drinking by low-to-moderate drinkers is associated with harm.

1.1 Aim

To systematically review the published literature on the effects of low-to-moderate prenatal alcohol exposure and binge drinking on the embryo, fetus and developing child.

1.2 Preliminary search of the literature

A preliminary search of the literature was carried out to identify previous systematic reviews and meta-analyses, and to develop the definitions and search strategy. This yielded one systematic review of the effects of prenatal alcohol exposure on infant mental development (Testa et al, 2003), one systematic review of the effects of moderate alcohol consumption on birth defects (Polygenis et al, 1998) and one systematic review of the effects of moderate alcohol consumption on spontaneous abortion, stillbirth and premature birth (Makarechian et al, 1998). These reviews considered exposures up to much higher levels than 6 drinks per week. A number of useful review articles and a highly cited textbook on the fetal effects of alcohol were located and their reference lists searched to complement the computerised search (Abel, 1998; Makarechian et al, 1998; Polygenis et al, 1998; Jacobson & Jacobson, 1999; Kesmodel, 2001; National Institute on Alcohol Abuse and Alcoholism (Gunzerath et al, 2004); Testa et al, 2003; Huizink & Mulder, 2005).

1.3 Measurement of alcohol consumption and definition of low-to-moderate and bingeing

The measurement of alcohol consumption in pregnancy used in human observational studies is usually expressed as average daily, weekly or monthly consumption and then categorised into abuse, heavy, moderate, low, light, social etc. Unfortunately, there has been little uniformity in definitions of low and moderate levels of alcohol consumption during pregnancy in the literature on alcohol and pregnancy (Kalter, 2003). As concerns mount about the potential effects of low-to-moderate exposure, investigators have started to look in more detail at weekly rather than daily consumption, which corresponds to much more prevalent levels of exposure in the general population.

This review evaluated studies about two measures of consumption: (1) average alcohol intake of less than 7 drinks per week (or less than one drink per day) and (2) bingeing. We decided to use the authors' definition of binge drinking since there was such variability in definition.

For the purposes of comparison, the national alcohol harm reduction strategy for England defines a binge episode (for women) as six or more units in a single session, which is equivalent to the US definition of four drinks or more on one occasion.

2 Methods

2.1 Study inclusion and exclusion

The following inclusion criteria were applied:-

1. Human studies only: Pregnant women, stillborn and live children (up to age 16).
2. Case-control, cohort or cross-sectional studies.
3. Studies published between January 1970 and July 2005 in the English language in a peer-reviewed journal.
4. Average weekly alcohol consumption level grouped into two or more categories. The ranges of at least two of these categories must be contained within the range of less than seven drinks per week (or equivalent) and may include an abstainer or infrequent drinker (<1 drink per week) group if available.

   OR
5. Studies reporting an effect measure for binge drinking during pregnancy.

   AND
6. Outcome data on any of the outcomes below.

Studies were excluded for any of the following reasons:-

1. No quantitative measure of alcohol consumption which could be converted to UK standard units and grams of alcohol.
2. Average alcohol consumption treated only as a continuous (and not as a grouped) variable and not limited to the low-to-moderate range.
3. Insufficient data for an (adjusted and/or crude) effect measure of low-to-moderate consumption and/or binge drinking to be extracted.
4. Duplicate publication.
5. Study available in abstract form only.

2.2 Outcomes (as defined by the authors)

Miscarriage
Antepartum haemorrhage
Stillbirth
Intrauterine growth restriction
Preterm birth
(Low) birth weight
Small for gestational age at birth
Small for age in childhood
Birth defects
Microcephaly and head circumference
Fetal alcohol syndrome
Neurodevelopmental outcomes (see appendix 4)

2.3 Search strategy (appendix 5)

A computerised literature search was undertaken using the WebSpirs 5 software and the following databases:

Medline on WebSpirs (1970-2005)
Embase on WebSpirs (1980-2005)
Cinahl on WebSpirs (1982-2005)
PsychInfo on WebSpirs (1972-2005)

MeSH headings and free text terms were used for the exposure and outcomes. The search was limited to studies where either 'low', 'light', 'social', 'moderate', 'dose' or 'binge' appeared in the text in relation to the exposure. The results were then 'filtered' using the 'high sensitivity' filter for aetiological studies developed by Wilczynski et al (2003). The results were then limited to human studies published in the English language in peer reviewed journals from 1970 - 2005. Review articles, commentaries, case series and editorials were excluded.

All located records with available abstracts (where available) were downloaded to a Procite database and stored with the search strategies for each database. Duplicate records were then identified and deleted.

2.4 Study selection criteria and procedures

1. Title and abstract (if present) of all studies identified by computerised literature search were reviewed independently by two members of the research team to identify potentially relevant papers.
2. Differences were resolved by discussion.
3. Papers deemed relevant, or of uncertain relevance were obtained and read in full.
4. All selected papers were reviewed against inclusion/exclusion criteria independently by two members of the research team to identify relevant papers. Differences were resolved by discussion and a third party if necessary. Reasons for exclusion were identified.

Progress was quantified at all stages of study selection using a flow diagram.

2.5 Quality assessment

This was performed using the Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies in meta-analyses. The scale is recommended by the Cochrane Non-Randomized Studies Methods Working Group and is reproduced in appendix 6 (Wells et al, 2005). It uses a system in which a study is judged on three areas: the selection of the study groups; the comparability of the groups; and the ascertainment of either the exposure or outcome of interest for case-control or cohort studies respectively.

2.6 Data extraction

A data extraction form was designed, piloted and revised (see appendix 7). Each included article was read and data extracted by a member of the study team; a second member checked table entries for accuracy against the original article.

2.7 Presentation and synthesis of extracted data

Data were synthesised in tables giving descriptive information for each study included. This was performed separately for each of the outcomes within the low-to-moderate range and separately for binge drinking. Where authors had not presented effect measures with confidence intervals or tested statistical significance for their findings we calculated these from the summary statistics presented where possible using Stata 8. These calculated values are marked in the tables in bold italic type to distinguish them from the results presented by the authors. For a few studies this could not be done as it would have required contacting authors for further data.

3 Results

Searches of Medline, Cinahl, Embase and PsychInfo resulted in 3543 papers (see Fig 1). Exact duplicates were deleted on merging. Of these, 308 papers were marked as either relevant (121) or uncertain relevance (187) on the basis of title and abstract (where available). Full text of these papers was obtained. (A further 268 papers were duplicates although differences in punctuation, etc, meant that they were not deleted earlier). In addition, 87 papers from bibliographies were also obtained (395 in total).

Final result of search:

66 papers were included
321 excluded
8 unobtainable

A summary of which outcomes were considered in which studies is given in appendix 8.

Reasons why individual studies were excluded are available from the authors.

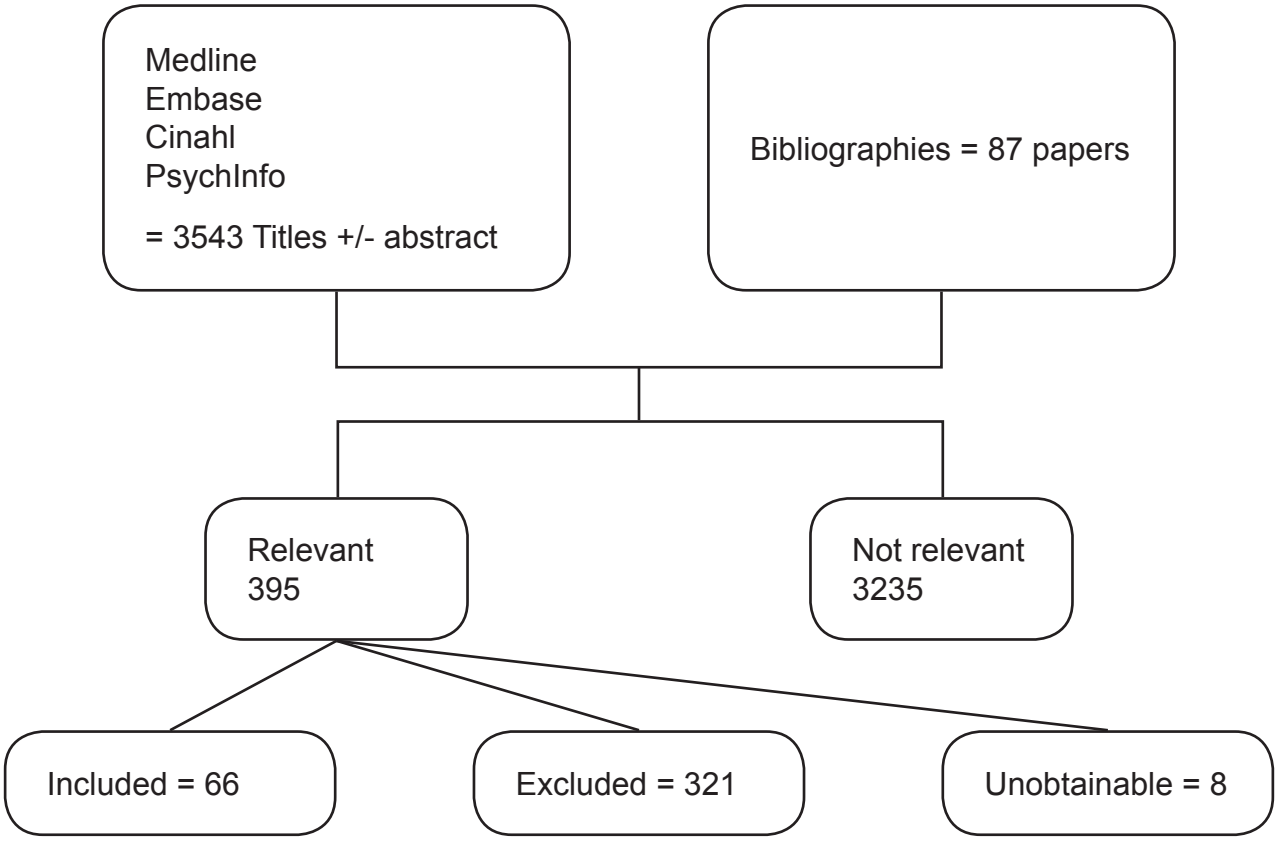
The next section considers each outcome in turn in relation in low-to-moderate alcohol consumption; for each outcome we present a summary table of the papers included and data extracted, and describe the results in text.

We conclude with a discussion of the quality of the evidence and the strengths and weaknesses of findings.

3.1 Abbreviations used in tables

AA	absolute alcohol	lbw	low birth weight
adj	adjusted	OR	odds ratio
AN	antenatal	PN	postnatal
bw	birth weight	RR	relative risk
CI	confidence interval	SA	spontaneous abortion
cr	crude	SD	standard deviation
GA	gestational age	SE	standard error
g/wk	grams per week	SES	socioeconomic status
g/days	grams per 24 hours	SGA	small for gestational age
HMO	Health Maintenance Organisation	*	$p < 0.05$
IUGR	intrauterine growth restriction		

Fig 1 - Results of search



3.2 Low-to-moderate alcohol consumption

3.2.1 Spontaneous abortion

There were eight studies describing the effects of low-to-moderate dose alcohol consumption on spontaneous abortion, three from the USA, two each from Denmark and the UK, and one from Canada. Two were case control studies, the rest were cohort studies. Information about spontaneous abortion was from maternal self report (Armstrong et al, 1992), hospital records or chart review (Harlap & Shiono 1980; Long et al, 1994; Windham et al, 1992; Windham et al, 1997), record linkage (Kesmodel et al, 1987) and in one case daily urine specimens were used to establish pregnancy and pregnancy loss (Henriksen et al, 2004). Spontaneous abortion was limited to the first trimester in one study (Long et al, 1994), 20 weeks in two (Windham et al, 1992, 1997) and 27/28 weeks in the rest. All of these studies except Henriksen et al (2004) used hospital or antenatal populations and thus would not have been able to include very early pregnancy loss.

The case control study by Long et al (1994) limited to 1st trimester miscarriages reported that women drinking 1-10 units of alcohol per week (up to 80g, equivalent to 6.7 drinks) had a relative risk of 3.79 (95% CI 1.18 to 12.17). This was adjusted for confounders but exactly what confounders were not stated. Moreover, alcohol exposure was 'pre-pregnancy' but it was not clear how long prior to pregnancy. The other case control study (Windham et al, 1992) reported an adjusted odds ratio of 1.2 (95% CI 0.91 to 1.90) for up to 6 drinks per week compared to less than half a drink per week.

The cohort studies reported rates of spontaneous abortion in non-drinkers (or <0.5 drinks per week) ranging from 1.1% (Davis et al, 1982) to 20.5% (Armstrong et al, 1992). In those who consumed up to 4 drinks per week, rates of spontaneous abortion ranged from 0.6% (Davis et al, 1982) to 32% (Henriksen et al, 2004). Three studies used survival analysis to calculate hazard ratios (Kesmodel et al, 1987; Windham et al, 1997; Henriksen et al, 2004). Adjusted for the major confounders, these ranged from 0.8 (Kesmodel et al, 1987, 1st trimester 3-4 drinks per week) to 2.1 (Henriksen et al, 2004, 1-4 drinks per week). This latter result was of borderline statistical significance. The wide range in rates of SA reported in these studies may be due to a number of factors. Studies which monitored urine (such as Henriksen et al, 2004) tend to report higher rates. Conversely, some studies only count SA after a fetal heart beat has been seen on ultrasound.

Of the eight studies, five found that women who consumed less than 7 drinks per week were at significantly increased risk of spontaneous abortion. The highest reported risk (Long et al, 1994) was a relative risk of 3.79 (95% CI 1.18 to 12.17) associated with consuming up to 10 units (equivalent to 6.7 drinks). However, this study had significant limitations as previously described. Three other studies reported adjusted relative risks of 2.0, 1.9 and 2.1 (Harlap & Shiono, 1980; Windham et al, 1997; Henriksen et al, 2004). However, the first of these was among women who also smoked one and a half packets of cigarettes per day, and the second and third were of only borderline statistical significance. A further study by Armstrong et al (1992) reported odds ratios of 1.11 (95% CI 1.05 to 1.18) and 1.23 (95% CI 1.13 to 1.34) associated with drinking 1-2 and 3-4 drinks per week respectively. However, this study was entirely based on maternal recall and likely to be subject to bias.

In summary, there were eight studies which examined the effect of low-to-moderate alcohol consumption on spontaneous abortion. Although five of these reported a significant effect, two had significant limitations, in one paper the only significant result was amongst heavy smokers and the remaining two were of only borderline statistical significance.

Table 1 - Spontaneous abortion (SA)

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Armstrong, 1992; Canada, cohort study	1982-84, 47,146 women	Interview following delivery or SA; SA at <28 wks by maternal self report	g/wk %SA OR (95%CI) None 20.5 1.00 1-24 22.7 1.11 (1.05, 1.18) 25-72 25.3 1.23 (1.13, 1.34)	OR adjusted for age, number of previous live births and SA, ethnic group, education, and employment status. Cigarette smoking and coffee consumption given separately. Survey response not stated – may be subject to recall bias
Davis, 1982; UK, cohort study	1980 973 women	Questionnaire at booking visit; outcomes from routine exam-ination and notes	g/day %SA 0 2.3 1-8 1.7 ***RR 0.73 (0.27, 2.00)***	White women only. Study underpowered and unadjusted for potential confounders.
Harlap, 1980; USA, cohort study	1974-77 32,019 women	Written self report at 1st AN visit; SA up to 27 wks from surveillance of hospital admissions and chart review	RR <12g/day compared to 0 (95% CI) 1st trimester 1.12 (0.59, 2.13) 2nd trimester 1.03 (0.57, 1.86) *2nd trimester miscarriages* *Non-smokers* *g/day Adjusted rate Adjusted RR* 0 2.3% 1.00 <12 2.4% 1.05 *<1.5 packs cigarettes/day* 0 2.4% 1.05 <12 3.0% 1.27 *1.5+ packs cigarettes/day* 0 2.6% 1.13 <12 4.7% 2.00*	Kaiser Foundation population. RR adjusted for age and gestational age at study entry.

Henriksen, 2004; Denmark, cohort study	1992-94 430 couples	Weekly written self report; SA within 28 wks, conception and early loss by daily urine sample	g/wk SA RR of SA (95% CI) 0 15% 1.0 1-48 32% 2.1 (1.0, 4.3)* Crude and adjusted hazard ratios: crude 2.2 adj. 2.1 (0.9, 4.9) 1-4 drinks compared to none	Limited to planned pregnancies. Hazard ratio adjusted for caffeine, smoking, age, menstrual cycle length. If also adjusted for partner alcohol intake, hazard ratio falls to 1.8.
Kesmodel, 1987; Denmark, cohort study	1989-96 24,679 women	Questionnaire in early pregnancy; SA before 28 wks through record linkage to Danish National Patient Registry	g/wk 1st trimester SA 2nd trimester SA <12 1.4% 0.8% 12-24 2.3% 1.0% 25-48 2.0% 1.2% Crude & adjusted hazard ratios (95% CI) compared to <12g /wk: 1st trimester 12-24 *cr* 1.5 (1.0, 2.4) *adj* 1.3 (0.8, 2.0) 25-48 *cr* 1.2 (0.6, 2.5) *adj* 0.8 (0.4, 1.7) 2nd trimester 12-24 *cr* 1.3 (0.9, 1.7) *adj* 1.2 (0.9, 1.7) 25-48 *cr* 1.3 (0.8, 2.1) *adj* 1.1 (0.7, 1.9)	Questions about alcohol consumption validated against other methods. Hazard ratios adjusted for smoking, caffeine, age, pre-pregnant body mass index, marital status, occupational status, education & parity. Additional adjustment for previous SA made little difference. Similarly, stratifying by week of study entry made no difference.
Long, 1994; UK, case-control study	95 cases, 3348 controls (recruitment period not stated)	Interview at first AN visit (controls), on admission (cases); women presenting with 1st trimester SA	Pre-pregnancy g/wk Cases Controls 0 6.3% 12.8% 1-80 58.9% 68.5% RR of SA (95% CI) 1-80g 3.79 (1.18, 12.17) compared to non-drinkers	Study limited to 1st trimester miscarriage. Exposure ‘pre-pregnancy’ - not clear how much before pregnancy. Reported that RR adjusted for confounders but which confounders were not stated.

Windham, 1997; USA, cohort study	1990-91 5342 women	Telelphone interview at <13 wks; SA up to 20 wks, 75% from hospital records	g/wk %SA Risk ratio (95% CI) 0 9.4 1.0 1-6 17.5 1.9 (1.1, 3.2) 7-12 10.2 1.1 (0.7, 1.7) 13-36 12.9 1.3 (0.7, 2.3) 37-48 23.1 2.5 (1.2, 5.0) Adjusted OR Hazard ratio 1-6g/wk 1.9 (1.0, 3.8) 1.9 (1.0, 3.5) 7-36g/wk 1.0 (0.7, 1.5) 1.1 (0.75, 1.6)	Kaiser Permanente population. Telephone interviews conducted at <13 wks to reduce recall bias. Adjusted for age, prior SA, gestational age at interview, cigarettes and caffeine in week before interview.
Windham, 1992; USA, case-control study	June 1986-Feb 1987 626 cases, 1300 controls	Interview after delivery or SA; SA up to 20 wks for which pathology specimen examined, data from chart review	g/wk Cases Controls <6 62.0% 67.1% 7-36 27.3% 25.2% 37-72 6.4% 5.6% g/wk Adjusted OR (95% CI) <6 referent 7-36 1.2 (0.92, 1.5) 37-72 1.2 (0.81, 1.9)	Controls randomly selected live births matched by hospital and week of last menstrual period. Cases were women presenting with SA. ORs adjusted for active and passive smoking, age and nausea.

3.2.2 Stillbirth

The association between low-to-moderate levels of alcohol consumption in pregnancy and stillbirth have been examined in five studies, three cohort, two case control. All five studies used large hospital or maternity data sets. In the three cohort studies rates of stillbirth were between 3-6 per thousand births (Davis et al, 1982; Marbury et al 1983). Only one study reported significantly increased rates of stillbirth in babies of women who drank up to 25-60g per week in pregnancy (Faden et al, 1997). However, this finding was based on only 20 cases and 11 controls in the exposed group. Three of the other studies reported higher rates of stillbirth in women who did not drink at all (Davis et al, 1982; Marbury et al, 1983; Little & Weinberg, 1993).

However, in two out of these three studies, alcohol consumption was only asked about after delivery. Results are therefore likely to be subject to recall bias. None of the three studies adjusted for confounders in the analyses reporting low-to-moderate levels of alcohol consumption (although adjustments were made in other analyses). The only paper reporting results not subject to recall bias and adjusting for confounders (Kesmodel et al, 2002) also used a validated questionnaire to ask about alcohol consumption. They found that, although low-to-moderate level alcohol consumption in pregnancy was associated with slightly higher rates of stillbirth, it was not statistically significant. They also reported no interaction or effect modification with smoking.

In summary, only one of the five studies which examined this outcome found a significant effect of low-to-moderate drinking in pregnancy and that was based on small numbers. Three studies reported higher rates of stillbirth in women who abstained but these were not statistically significant differences and were unadjusted for potential confounders.

Table 2 - Stillbirth

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Davis, 1982; UK, cohort study	1980 973 babies	Questionnaire at booking visit; outcomes from routine examination and notes	g/24 hrs % stillbirth 0 0.4 1-8 0.3 ***RR = 0.99 (0.22, 4.41)***	White women only. Study underpowered and unadjusted for potential confounders.
Faden, 1997; USA, nested case control study	1988 2019 stillbirths 8996 live births	Written self report postpartum; outcome from National Maternal & Infant Health Survey	g alcohol: live births / Weighted % stillbirths None 79.3 / 84.5 < 12/mth 10.8 / 7.8 12/mth 3.5 / 1.7 13-36/mth 2.5 / 2.1 12/wk 1.3 / 1.3 13-24/wk 1.3 / 0.8 25-60/wk 0.8 / 1.0	Reported as a representative sample but authors do not state how sample was taken. Response rates were 74% for live births and 69% for stillbirths. Weighted % reported to correct for over sampling of low birthweight and black infants. However, not adjusted for confounders in this analysis.
Kesmodel, 2002; Denmark, cohort study	1989-96 24,768 including 116 stillbirths	Questionnaire in early pregnancy; outcome data from Birth Registration forms and Danish Medical Birth Register through record linkage	g/wk % stillbirth <12 4.2 12-24 5.1 25-48 5.6 g/wk Crude (95%CI) OR / Adj. (95%CI) OR <12 1.00 / 1.00 12-24 1.23 (0.8, 1.9) / 1.27 (0.8, 2.0) 25-48 1.35 (0.7, 2.6) / 1.25 (0.6, 2.5)	Questions about alcohol consumption validated against other methods. OR adjusted for maternal smoking, caffeine, age, pre-pregnancy body mass index, marital and occupational status, education, parity and sex of child. No interaction or effect modification with smoking. Analysis excluding women with missing data (N=16,010; 74 stillbirths) increased OR for 25-48g but still not significant.

Little, 1993; USA, case control study	1980 1835 cases, 2832 live birth controls	Postpartum questionnaire; outcome data from birth / death certificates and hospital records. Included antepartum and intrapartum deaths.	*Alcohol use prior to pregnancy (g/day)* % live births / %AP deaths / % IP deaths 0: 40.9 / 45.5 / 48.7 0.2-2.4: 37.6 / 33.6 / 36.0 2.5-6.0: 11.4 / 11.4 / 11.6 6.1-12.0: 5.8 / 5.2 / 1.6 *Alcohol use during pregnancy (g/day)* 0: 51.8 / 57.4 / 63.2 0.2-2.4: 42.8 / 39.2 / 35.3 2.5-6.0: 3.7 / 2.2 / 1.0 6.1-12.0: 0.7 / 0.7 / 0.0 ***Alcohol use prior to pregnancy (g/day)*** ***OR (95% CI) AP/IP*** ***0: 1.00*** ***0.2-2.4: 0.80 (0.68, 0.94)*** ***2.5-6.0: 0.89 (0.70, 1.13)*** ***6.1-12.0: 0.71 (0.49, 1.00)*** ***Alcohol use during pregnancy (g/day)*** ***0: 1.00*** ***0.2-2.4: 0.80 (0.69, 0.93)*** ***2.5-6.0: 0.48 (0.28, 0.77)*** ***6.1-12.0: 0.69 (0.22, 1.82)***	'Before' and 'during' pregnancy not defined. Response rate to questionnaire 78% and further 22% excluded due to incomplete telephone interviews. Differential response rate not mentioned. Only 39% of cases included in analysis due to exclusion of multiple births and women with serious medical conditions. Drinking asked about after delivery – may be subject to recall bias. No adjustment for confounders in this analysis. Protective effect.
Marbury, 1983; USA, cohort study	about 1982 12,440 births included 73 stillbirths	Retrospectively by PN interview; outcome from medical records	g/wk / % stillborn / ***RR (95% CI)*** 0: 0.6 / ***1.00*** 1-24: 0.6 / ***0.97 (0.50, 1.89)*** 25-72: 0.4 / ***0.66 (0.21, 2.10)***	Drinking asked about after delivery – may be subject to recall bias. Not adjusted for confounders in this analysis. Adjustment only done on regrouped drinking variable.

3.2.3 Antepartum haemorrhage

There was the only study which included antepartum haemorrhage (APH) as an outcome. No increase in risk of APH was apparent in this study.

Table 3 - Antepartum haemorrhage

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Raymond, 1993; USA, cohort study	30,681 including 307 placental abruptions	Questionnaire in early pregnancy; outcome data from hospital records and autopsy reports	Alcohol % abruption 0 g/day 0.99 <12 g/day 0.99 RR (95% CI) 1.00 1.00 (0.80, 1.26)	Kaiser Permanente may not be a truly representative population. Diagnosis of placental abruption dependent upon clinical judgement. No adjustment for confounding in this analysis.

3.2.4 Intrauterine growth restriction

There were seven studies in which intrauterine growth restriction (IUGR) or small for gestational age (SGA) were reported outcomes. They were all cohort studies except for two case control studies (Windham et al, 1995; Yang et al, 2001). IUGR is an intrauterine diagnosis based on failing growth on serial ultrasound; SGA is diagnosed at birth. Definitions of IUGR and SGA varied. Some studies used the 5th or 10th percentiles of birthweight for gestational age (McDonald et al, 1992; Lundsberg et al, 1997 respectively), some corrected for race and sex (Mills et al, 1984) and parity (Whitehead & Lipscomb, 2003). One (Verkerk et al, 1994) used a ratio of observed to expected birthweight for gestational age, sex and parity.

Only one of the studies found a significant positive association between low-to-moderate levels of alcohol consumption and IUGR (Windham et al, 1995). However, the relevant analysis in this paper was not adjusted for potential confounders and therefore may to be misleading. This was also true of the study by Whitehead & Lipscomb (2003).

Information about alcohol exposure was collected retrospectively in all these studies and may therefore be inaccurate. Moreover, in five studies (McDonald et al, 1992; Verkerk et al, 1994; Windham et al, 1995; Yang et al 2001; Whitehead & Lipscomb, 2003) data about the outcomes were also collected retrospectively. There is, therefore, the potential for recall bias. Three studies found that low-to-moderate levels of alcohol consumption appeared to be mildly protective against IUGR (McDonald et al, 1992; Whitehead & Lipscomb, 2003; Lundsberg et al, 1997). However, the McDonald et al (1992) study adjusted for previous low birthweight baby which may represent over-adjustment if the previous baby was also exposed to alcohol in utero.

In summary, only one of the seven studies which examined intrauterine growth restriction found a significant association and that was unadjusted for potential confounders. Three studies found low-to-moderate alcohol consumption to be mildly protective but, although of borderline statistical significance, two may have been subject to recall bias.

Table 4 - Intrauterine Growth Restriction

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Lundsberg, 1997; USA, cohort study	1988-92 2714 babies	Interview before 16th wk of gestation; outcomes from examination at birth; IUGR defined as lowest 10th percentile of birthweight for gestational age	g/day RR (95% CI) adj OR (95% CI) *Alcohol at mth 1* abstinent 1.00 1.00 <2.4 0.74 (0.53, 1.05) 0.85 (0.57, 1.26) 2.4-6 0.37 (0.20, 0.68) 0.39 (0.20, 0.76) *Alcohol at mth 7* abstinent 1.00 1.00 <2.4 0.90 (0.62, 1.30 1.12 (0.73, 1.69) 2.4-6 0.44 (0.18, 1.05) 0.48 (0.18, 1.23)	Adjusted for smoking in 7th month of pregnancy, height, weight, ethnicity, infant sex, parity, bleeding during pregnancy, high blood pressure, pre-/eclampsia. Participation rate 75%
McDonald, 1992; Canada, cohort study	1982-84 40,445 pregnanc-ies	Postpartum interview; clinical data from medical records; SGA defined as bottom 5% of birthweight for gestational age.	g/wk % SGA OR (95% CI) None 5.5 1.00 1-24 4.1 0.80 (0.70, 0.91) 25-72 4.5 0.85 (0.72, 1.00)	OR adjusted for age, pregnancy order, previous spontaneous abortion, previous lbw infant, pre-pregnancy weight and employment, ethnic group, education, cigarette and coffee consumption. This may represent over-adjustment if previous lbw baby was also associated with alcohol consumption.
Mills, 1984; USA, cohort study	1974-77 31,503 women	Retrospectively by questionnaire in 1st trimester; outcomes from medical records. SGA defined as <10th centile for race, sex and GA	g/day % SGA None 5.8 <12 6.9 Adj OR for SGA 1.11 (1.00, 1.23) <12g/day compared to 0	Kaiser Permanente population. Alcohol consumption only collected for 1st trimester. OR for SGA adjusted for education, race, smoking and parity (age, hypertension, spontaneous abortions and a smoking-alcohol interaction were not significant and were dropped).

Verkerk, 1994; Netherlands, cohort study	1988-89 2027 babies	Interview at about 3 wks postpartum; IUGR by birthweight ratio (ratio of observed bw to expected mean bw for GA, sex and parity).	Regression coefficients for birthweight ratio None were significant. g/wk Unadjusted (95% CI) Adjusted (95% CI) 0 0 0 1-12 .006 (-.008, .021) .001 (-.014, .015) 13-84 .009 (-.020, .038) .003 (-.025, .015)	Large population based sample with 97% response rate. However, interviews done approx. 3 months postpartum so potential for recall bias. Adjusted for smoking, social class, ethnicity, occupation, maternal age and height.
Whitehead, 2003; USA, cohort study	1996-99 50,461 women	Questionnaire shortly after delivery with telephone follow-up of non-responders; outcomes from birth certificate data	*3 mths before pregnancy* g/wk %SGA RR (95% CI) None 7.8 1-36 7.1 0.92 (0.84, 1.00) *last 3 mths of pregnancy* g/wk %SGA RR (95% CI) None 7.9 1-36 7.5 0.95 (0.80, 1.14)	SGA defined as birthweight <10th percentile for GA according to race and parity specific standards. 76% response rate but over-representation of SGA babies. No adjustment in this analysis.
Windham, 1995; USA, cohort study	1986-87 1252 women	Interview 8-9 mths after delivery; outcomes from chart review	IUGR 10th percentile g/wk % OR (95% CI) None 8.9 1.0 1-24 8.4 0.9 (0.6, 1.4) 25-60 20.3 2.6 (1.4, 5.0) IUGR 5th percentile g/wk % OR (95% CI) None 3.3 1.0 1-24 4.6 1.4 (0.8, 2.5) 25-60 7.3 2.3 (1.3, 17.9)	Principally case control study of spontaneous abortion. However, this study focused only on the controls, which were generated from a random sample of birth certificates. ORs unadjusted in this analysis.

Yang, 2001; USA, case-control study	1993-95 701 cases, 336 controls	Postpartum interview; small for gestational age defined as <10th percentile for gestational age; <5th percentile = severe, 5th to 10th = moderate IUGR	>0 <36 g/wk 3 mths prior 1st trimester 2nd/3rd trim'r compared to 0 g per week	Moderate OR (95%CI) 1.2 (0.8, 1.8) 1.4 (0.9, 2.2) 0.9 (0.6, 1.4)	Severe IUGR OR (95%CI) 0.9 (0.6, 1.4) 1.2 (0.8, 1.9) 0.9 (0.6, 1.4)	Participation rate 47%. ORs adjusted for maternal age, weight gain during pregnancy, educational attainment, race, cigarettes. Potential for recall bias due to asking about alcohol consumption after delivery.

3.2.5 Birth weight

There were 20 studies which examined the association between alcohol consumption and birthweight (mainly low birthweight). All were cohort studies ranging in size from 412 (Jacobson et al, 1994) to 40,445 (McDonald et al, 1992). Birthweight and alcohol consumption are both strongly associated with cigarette smoking making this a potential confounding factor. Six studies did not adjust for this factor in their analyses looking at the association with low-to-moderate alcohol (Marbury et al, 1983; Lumley et al, 1985; Sulaiman et al, 1988; Virji et al, 1990 and 1992; O'Callaghan et al, 2003) although some did in other analyses. Another important variable in this context is ethnicity, since both birthweight and alcohol consumption are associated with this. Of the 12 studies that carried out some adjustment, only five either adjusted for ethnicity (Mills et al, 1984; McDonald et al, 1992; Lundsberg et al, 1997) or included only white or only black women (Brooke et al, 1989; Jacobson et al, 1994).

A further potential problem with several studies was that of differential recall bias where women were asked after birth about their drinking habits in pregnancy (Marbury et al, 1983; Virji et al, 1990; McDonald et al, 1992; Lazzaroni et al, 1993; Primatesta et al, 1994). In these studies the women may have inaccurately recalled consumption therefore introducing bias. However, the paper by Primatesta (1994) cites a separate study in which women were asked about alcohol consumption whilst pregnant and this was compared with the level of consumption they recalled in pregnancy when asked about it after birth. The results showed that women reported higher consumption during pregnancy when asked about it after birth. This suggests that women may feel at greater pressure to under-report alcohol consumption when asked about it during pregnancy.

The four studies that adjusted for relevant confounders and asked about alcohol consumption prior to delivery were either not truly representative (Mills et al, 1984; Brooke et al, 1989; Day et al, 1990; Jacobson et al, 1994) or had a relatively low participation rate (76% Lundsberg et al, 1997) and poor outcomes may, therefore, have been under-represented.

Nevertheless, there were few statistically significant results. Only the study by Lundsberg et al 1997) reported a significant increase in the risk of low birthweight with consumption of <0.1 oz alcohol per day (adjusted RR 3.20, 95% CI 1.87 to 5.46). However, at 0.1 - 0.25 oz per day, the RR was lower at 1.36 (95% CI 0.48 to 3.88). McDonald et al (1992) found an OR of 0.79 (95% CI 0.70 to 0.90) associated with consuming 1-2 drinks per week; in other studies there was a tendency for mean birthweight to be slightly higher in light drinkers (Mills et al, 1984; Bell et al, 1989; Verkerk et al, 1993; Primatesta et al, 1994; Shu et al, 1995; Passaro et al, 1996).

Table 5 - Birth weight

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Bell, 1989; Australia, cohort study	1985 8884 women	All data from the Victorian Perinatal Morbidity Statistics	Mean birthweight (sd) Non-drinkers 3420 (556) <36g /wk 3459 (497) *p<0.01*	45% of hospitals took part, of these 65% recorded smoking and alcohol details. Some of the analyses adjusted for smoking but no other potential confounders were included in these analyses. Possible bias due to smoking and drinking details being less well recorded in cases of poor outcome.
Brooke, 1989; UK, cohort study	Not stated when but over a 20 mth period; 1513 women	Interviews at booking, 28 and 36 wks gestation regarding consumption in preceeding week; outcome data from obstetric record	Alcohol (g/wk) / Adj birthweight ratio (95% CI) Non-smokers / Smokers 0: 1.05 (1.04, 1.06) / 1.02 (1.00, 1.04) 1-19: 1.05 (1.04, 1.07) / 1.01 (0.98, 1.03) 20-49: 1.06 (1.04, 1.08) / 0.99 (0.96, 1.01)	White women only, 81% response rate. Birthweight ratio, defined as the ratio of observed birthweight to expected mean birthweight for GA from an external standard, adjusted to maternal height, parity and gender of baby.
Day, 1990; USA, cohort study	1983-86 595 babies	Interview in 4th and 7th month of gestation; outcome by routine postnatal examination	g/day / Adjusted mean weight (kg) +/- SE None: 3.25 +/- 0.02 >0 to 11g: 3.23 +/- 0.04	Population predominantly low socioeconomic status. Methods for estimating alcohol exposure validated. Adjusted for maternal height, GA, weight gain during pregnancy, smoking, race & gender.

Jacobson, 1994; USA, cohort study	Neither year nor period of recruitment stated; 412 babies	Alcohol exposure collected for fortnight preceding each antenatal visit by interview; growth at 6.5 months after birth blind to exposure	g/day	% infants in bottom 10th centile weight	***RR (95% CI)***	Black infants only; inner city deprived population. Excluded if baby <1500g or <32 wks GA, malformed or a multiple. Oversampled drinkers of at least 0.5oz AA/day. Participation rate and follow-up rates not stated. Adjusted for smoking, opiate and cocaine use, maternal age, pre-pregnancy weight, height, education, marital status, welfare status, parity, infant's age, sex, number of antenatal visits.
			0	6.1	***1.00***	
			1-6	8.9	***1.48 (0.53, 4.13)***	
			7-12	11.8	***1.94 (0.52, 7.29)***	
Lazzaroni, 1993; Italy, cohort study	Feb 1989 - July 1990 2145 live births	Retrospectively by postpartum interview; outcomes from clinical notes and paediatrician examination	1-10g alcohol/day	OR (95% CI) of lbw		OR adjusted for maternal age, sex and gestational age of baby. Low birth weight defined as <2500g.
			smokers	1.08 (0.33, 3.38)		
			non-smokers	0.58 (0.19, 2.64)		
			compared to abstainers			
Lumley, 1985; Australia, cohort study	1981-82 10,319 births	All data from the Tasmanian perinatal statistics. lbw defined as <2500g.	*% low birthweight*	g/wk		Study based on all Tasmanian births 1981-82. However, only 69% response on alcohol consumption. Non-respondents similar in terms of parity and social class but significantly lower birthweight. No adjustment for confounders in this analysis other than social class. Difference between abstainers and light drinkers not seen consistently within parity or social class.
			Social class	None	1-24	25-72
			Professional	3.6	3.0	-
			Intermediate	2.7	3.8	4.9
			Skilled	3.7	4.6	6.2
			Semiskilled	5.2	5.5	12.5
			Unskilled	5.4	4.6	3.2
			ALL	4.7	4.5	5.6
			RR (95% CI)	***1.0***	***0.9 (0.8, 1.1)***	***1.2 (0.8, 1.7)***

Lundsberg, 1997; USA, cohort study	1988-92 2714 singleton live births	Interview before 16 wks gestation; examination at birth. lbw defined as <2500g.	*Alcohol at mth 1* *g/day* *RR lbw (95% CI)* *adj OR (95% CI)* abstinent 1.00 1.00 <2.4 0.79 (0.48, 1.28) 1.05 (0.58, 1.89) 2.4-6 0.88 (0.48, 1.59) 1.29 (0.63, 2.64) *Alcohol at mth 7* abstinent 1.00 1.00 <2.4 1.89 (1.21, 2.94) 3.20 (1.87, 5.46) 2.4-6 0.94 (0.38, 2.30) 1.36 (0.48, 3.88)	Adjusted for smoking in mth 7, height, weight, ethnicity, infant sex, parity, coffee in mth 7, exercise in 3rd trimester, employment, bleeding during pregnancy, high blood pressure, pre-/eclampsa, anomalies and placental problems. Participation rate 76%
Marbury, 1983; USA, cohort study	about 1982 12,440 births	Postpartum interview; outcome from medical records. lbw defined as <2500g.	g/wk % lbw ***RR (95% CI)*** 0 8.1 ***1.00*** 1-24 7.2 ***0.89 (0.74, 1.06)*** 25-72 6.9 ***0.80 (0.61, 1.05)***	% lbw unadjusted in this analysis. Adjustment only done on alcohol grouped <1, 1-10 and >10 drinks/wk.
McDonald, 1992; Canada, cohort study	1982-84 40,445 pregnanc-ies	Postpartum interview; clinical data from medical records. lbw defined as <2500g.	g/wk % lbw OR (95% CI) None 6.0 1.00 1-24 4.4 0.79 (0.70, 0.90) 25-72 5.0 0.87 (0.74, 1.01)	OR adjusted for age, pregnancy order, previous spontaneous abortion, previous lbw infant, pre-pregnancy weight and employment, ethnic group, education, cigarette and coffee consumption. This may represent over-adjustment if previous lbw baby was also associated with alcohol consumption.

Mills, 1984; USA, cohort study	1974-77 31,503 women	Questionnaire in 1st trimester; outcomes from medical records; lbw defined as <2500g.	g/day% lbw RR (95% CI) None 4.0 **1.00** <12 4.1 **1.03 (0.92, 1.14)** Mean birthweight (g) (+/- sd) *smoking 0 <.5 1 >1.5 packs/day* No 3468 3330 3256 3274 Alcohol (534) (548) (533) (519) <12g 3500 3386 3276 3255 per day (527) (522) (532) (548) Adj mean difference in birthweight <12g /day compared to 0: -14g (-26, -2.8)	Kaiser Permanente population. Alcohol consumption only collected for 1st trimester. Mean difference in birthweight adjusted for education, race, smoking and parity (age, hypertension, spontaneous abortions and a smoking-alcohol interaction were not significant and were dropped).
O'Callaghan, 2003; Australia, cohort study	1981-84 6320 at birth, 4038 at 5 yrs	Alcohol exposure by interview at first AN visit and postpartum; outcomes from medical staff at delivery	*Alcohol in early pregnancy* Birthweight % <3rd 3-10th percentile Nil 2.9 6.9 <6g/day 2.9 7.1 6-12g/day 4.3 5.3 *Alcohol in late pregnancy* % <3rd 3-10th percentile Nil 2.8 7.2 <6g/day 2.5 6.7 6-12g/day 2.8 7.6 ***Alcohol in early pregnancy RR (95% CI)*** ***<3rd 3-10th percentile*** ***Nil 1.00 1.00*** ***<6g/day 1.00 (0.77, 1.31) 1.03 (0.87, 1.22)*** ***6-12g/day 1.49 (0.77, 2.89) 0.77 (0.43, 1.38)*** ***Alcohol in late pregnancy <3rd 3-10th percentile*** ***Nil 1.00 1.00*** ***<6g/day 0.89 (0.65, 1.21) 0.93 (0.77, 1.12)*** ***6-12g/day 0.99 (0.47, 2.10) 1.05 (0.67, 1.64)***	No adjustment for confounders in this analysis. Participation at birth 74%. Women asked about alcohol consumption rather vaguely in terms of 'glasses', frequency measured as 'daily, occasionally, a few times a week, a few times a month', etc then converted to oz AA. Early pregnancy was around 20 weeks, late pregnancy was not defined.

Ogston, 1992; Europe multicentre, cohort study	Period of recruitment not stated; 8469 women	Interview at 12-18 wks and 28-32 wks; outcomes from hospital record	Mean birthweight *Alcohol intake* *(g/wk) Before pre.g. Early pregnancy* 0 3325 3314 >0-29 3326 3347 30-59 3351 3332 *Alcohol intake in early pregnancy* *(g/wk) Non-smokers Smokers* 0 3363 3225 >0-29 3414 3225 30-59 3419 3201 *Parameter estimates (SE) in multiple regression* >0-29 7.039 (10.57) 30-59 7.554 (17.53)	Representative of European population but different centres used slightly different methods. No information about participation rate. Unadjusted in first two analyses. Multiple regression includes gestation, centre, smoking, sex, age and parity of mother. No evidence of an effect of alcohol at these levels.
Ørskou, 2003; Denmark, cohort study	1990-99 36,265 babies in unadjusted analysis, 24,093 in adjusted analysis	Alcohol exposure at 16 weeks gestation; outcomes from birth registration forms	g/wk % *High birth weight (>4000g)* <12 9.7 12-24 19.4 25-48 19.0 unadj OR (95% CI) adj OR (95% CI) <12 1.00 1.00 12-24 0.98 (0.92, 1.05) 0.89 (0.82, 0.97) 25-48 0.96 (0.86, 1.08) 0.91 (0.79, 1.06)	Study included all singleton non-diabetic women. >80% returned both medical and research questionnaires. Inclusion of all who returned medical questionnaire only did not change results. Adjusted OR controls for maternal pre-pregnancy weight, height, age, parity, smoking, caffeine, living with partner, education, GA (estimated from ultrasound or date of last period), gender of child and year of birth. Reanalysis focussing on babies >4500g did not change results.

Passaro, 1996; UK, cohort study	1991-92 10,539 women	Alcohol exposure at 18 weeks gestation; outcomes from hospital delivery records	<2500g >4000g Pre-pregnancy abstainers 7% 11% <12g /day 4% 11% Early pregnancy abstainers 4% 11% <12g /day 4% 10% Adjusted mean differences in birthweight *Non-smokers Smokers* *Abstainers* -50 (-89, -11) -11 (-84, 62) *Drank before,* - abstained in preg ref ref - <12g/wk -9 (-31, 13) -21 (-66, 24) - 12-72g/wk 5 (-26, 36) 20 (-35, 75)	Large population based study with high response rate. Mean differences in birthweight adjusted for GA, infant sex, primiparity, maternal Body Mass Index. Only significant difference was lower birthweight in babies of women who abstained. Type of alcoholic drink made no difference to results. Limiting analysis to exclude users of marijuana, crack and cocaine, and excluding women with a history of alcoholism made no difference.
Primatesta, 1994; Italy, cohort study	1986-87 1516 women	Alcohol exposure by postpartum interview retrospectively after birth; outcome from clinical record	*Alcohol intake before pregnancy* Mean birthweight Unadjusted Adjusted (g/wk) Males Females Males Females 0 3376 3134 3413 3141 1-20 3406 3219 3413 3253 *Alcohol intake during pregnancy* Mean birthweight Unadjusted Adjusted (g/wk) Males Females Males Females 0 3402 3110 3434 3126 1-20 3372 3227 3401 3238	Study took place in 2 Milan hospitals; representativeness of sample not described but high risk pregnancies excluded. Potential for recall bias due to postpartum interview about. alcohol exposure. Adjusted for smoking, age, GA, maternal weight, height and parity, paternal weight. Type of alcohol consumed made no difference.

Shu, 1995; USA, cohort study	1987-89 876 women	Interview at 13 wks gestation then by telephone at 28 and 36 wks; outcomes from chart review	Smokers *1st trimester* g/wk none <12 <24 *2nd trimester* none <12 <24 *3rd trimester* none <12 <24 Non-smokers *1st trimester* drinks/wk none <12 <24 *2nd trimester* none <12 <24 *3rd trimester* none <12 <24	mean birthweight 3241 3333 3225 3247 3199 3255 3290 3320 3086 mean birthweight 3460 3489 3526 3447 3531 3554 3498 3489 3540	regression coeffic. ref 56 -36 ref 44 66 ref 152 -89 regression coeffic. ref 31 60 ref 48 -29 ref 8 -4	(95% CI) (-69, 180) (-258, 185) (-134, 222) (-311, 443) (-121, 425) (-441, 262) (95% CI) (-50, 112) (-97, 218) (-40, 137) (-200, 143) (-100, 115) (-221, 212)	Population based study in rural Pennsylvania. Predominantly white middle class. 74% participation rate. Regression coefficient adjusted for GA, log of pre-pregnant weight, per capita income, bleeding during 1st trimester and stratified for smoking. Type of alcoholic drink made no difference to results.

Sulaiman, 1988; UK, cohort study	1985-86 952 babies	Interview at 1st antenatal visit and in 3rd trimester; outcomes from hospital records	Birthweight (g) Abstainers 3277 1-50g/wk 3282	Population based study with 90% participation rate. Primiparous women only. Unadjusted in this analysis. No significant effects at 1-50 g/wk compared with abstainers.
Verkerk, 1993; Netherlands, cohort study	1978-79 3447 women	Interview in mid-pregnancy and immediately after birth; source of outcome data not stated	Alcohol Coefficient (birthweight ratio) *1st trimester* unadj adj (95% CI) abstainers 0 0 1-50g/wk 0.002 -0.005 (-0.016, 0.007) *2nd trimester* abstainers 0 0 ex-drinkers -0.001 -0.005 (-0.018, 0.008) 1-50g/wk 0.006 -0.002 (-0.014, 0.010) *3rd trimester* abstainers 0 0 ex-drinkers 0.005 0.002 (-0.012, 0.016) 1-50g/wk 0.006 0.003 (-0.010, 0.016)	Not clear how women selected, Dutch speaking only. Participants in midwife care, at low risk of complications; 15% lost to follow-up mainly due to referral to obstetrician. However, among women referred, early alcohol consumption was similar to that in women not referred. Information about 3rd trimester drinking obtained after birth, potential for recall bias. Reference group were women who had abstained from alcohol both prior and during pregnancy. OR adjusted for smoking, partner's smoking, education, employment, age and marital status. No significant differences found.
Virji, 1990; USA, cohort study	1980 3300 babies	Written self report postpartum; birthweight from National Natality Survey	g/mth % <2500g None 13.7 1-36 13.8 ***OR = 1.02 (0.78, 1.31)***	Study mainly focused on occupational influences on birthweight. Low birthweight babies oversampled. Study limited to white, married women delivering a singleton who were employed at some time in 12 months prior to delivery. No adjustment for confounders in this analysis.

Virji, 1991; USA, cohort study	1980 5400 babies	Written self report postpartum; birthweight from National Natality Survey	g/mth % <2500g None 14.0 1-36 14.4 ***RR = 0.96 (0.61, 1.53)*** g/mth Mean birth weight (sd) None 3301.14 (698.0) 1-36 3299.59 (680.9) ***t = 0.0618 p = 0.9507***	Data relate to white, married mothers giving birth to singletons for whom complete information on sociodemographic and risk factors were available i.e. 5400 out of 9941 (54%). No data on non-respondents. Results also stratified by smoking (not shown).

3.2.6 Preterm birth

There were 16 studies meeting the inclusion criteria which considered preterm birth as an outcome. Of these, two were poor quality case control studies (Berkowitz et al, 1982; Parazzini et al, 2003). They suffered from a possible lack of blinding to case/control status. Also, there was potential for recall bias due to information on alcohol consumption being collected after birth - women may have differentially over- or under-reported consumption in light of having a preterm baby. Moreover, recall of alcohol consumption over such a long period is unlikely to be accurate. In addition, the earlier paper (Berkowitz et al, 1982) failed to control for potential confounders.

Of the 14 cohort studies, half were based on very large datasets ranging in size from 8469 (Ogston & Parry, 1992) to 40,892 (Albertson et al, 2004), the other half were on a smaller scale ranging from 952 (Sulaiman et al, 1988) to 4111 (Wisborg et al, 1996). Many of these studies had the potential for recall bias (as described above) due to asking about alcohol consumption after birth (Marbury et al, 1983; McDonald et al, 1992; Verkerk et al, 1993 (partly); Verkerk et al, 1994). Half of the papers estimated gestational age from the date of last menstrual period and/or by ultrasound. However, two papers used the Dubowitz examination which is based on specific physiological and neurological characteristics of the neonate (Berkowitz et al, 1982; Sulaiman et al, 1988), three presumably used the information from the birth registration form or hospital records (Marbury et al, 1983; McDonald et al, 1992; Kesmodel et al, 2000) and two did not state how gestational age was estimated (Verkerk et al, 1993; Peacock et al, 1995).

Adjustment for confounding was done in all studies except seven (Berkowitz et al, 1982; Marbury et al, 1983; Sulaiman et al, 1988; Ogston & Parry, 1992; Peacock et al, 1995; Passaro et al, 1996; Wisborg et al, 1996) although in some of these studies adjustment was done in other analyses in which alcohol consumption was more broadly grouped, or when examining other associations. A further two studies did not make any attempt to control for socioeconomic status (through social class, education or occupation) (Lazzaroni et al, 1993; Lundsberg et al, 1997). However, some studies may have over-adjusted, controlling for previous low birth weight or preterm birth or spontaneous abortion, which may themselves have been associated with alcohol consumption.

Despite the different methods and different limitations of these studies, all except one study found either no effect or a reduction in risk of prematurity with consumption of up to 6 drinks per week. The exception was a US study based on two Health Maintenance Organisations and 11 private practices (Lundsberg et al, 1997). They found relative risks of 2.11 and 2.15 in women consuming <0.1 oz and 0.1-0.25 oz respectively of absolute alcohol per day at 7 months gestation. However, as mentioned above, they did not control for socioeconomic status.

Three studies found a significant protective effect of low-to-moderate alcohol consumption (Shiono & Klebanoff, 1986; McDonald et al, 1992; Kesmodel et al, 2000). This occurred at up to two drinks per week. Bell & Lumley (1989) (see Binge drinking section) postulated a ‘healthy drinker effect’ to account for this in parous women. It may be that if a woman is aware of potential problems, perhaps due to obstetric history, she may be more likely to abstain from alcohol.

In summary, as with birth weight, only one study out of 16 reported a significantly increased risk of preterm birth. This study may be subject to residual confounding as it was unadjusted for socioeconomic status.

Table 6 - Preterm birth

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results			Comments
Albertsen, 2004; Denmark, cohort study	1997-2000 40,892 women, 1880 preterm babies	Telephone interview at 12-16 wks gestation; outcome through record linkage to Danish Discharge Register	g/wk	% 32-36 wks	% <32 wks	Nationwide study including 35% of pregnant population. However, only about 60% of invited women participated (no information about non-respondents). GA estimated from date of last period or by ultrasound. Adjusted for type 1 diabetes, age, previous preterm birth, parity, smoking, coffee and occupational status. Controlling for previous preterm birth may represent over adjustment. Analysis by type of drink found no difference.
			0	4.3	0.5	
			1-6	4.0	0.5	
			7-18	3.6	0.6	
			19-42	3.3	0.5	
			43-78	4.7	0.8	
			32-36 wks	unadj RR	adj RR (95% CI)	
			0	1.00	1.00	
			1-6	0.91	0.93 (0.81, 1.06)	
			7-18	0.83	0.87 (0.76, 1.00)	
			19-42	0.75	0.77 (0.64, 0.93)	
			43-78	1.08	1.10 (0.79, 1.54)	
			<32 wks	unadj RR	adj RR (95% CI)	
			0	1.00	1.00	
			1-6	0.89	0.91 (0.61, 1.35)	
			7-18	1.17	1.24 (0.87, 1.76)	
			19-42	1.02	1.06 (0.66, 1.69)	
			43-78	1.52	1.53 (0.67, 3.49)	

Berkowitz, 1982; USA, case control study	1977 166 cases 299 controls	Postpartum interview; GA by Dubowitz examination	*1st trimester* OR (95% CI) Never/<12g/wk 1.0 12-72g/wk 1.0 (0.7, 1.5) *2nd trimester* Never/<12g/wk 1.0 12-72g/wk 1.0 (0.6, 1.4) *3rd trimester* Never/<12g/wk 1.0 12-72g/wk 0.8 (0.5, 1.2)	Not clear how representative Yale New-Haven hospital is. Only included whites and blacks. Alcohol intake only rough estimate. Postpartum interview may not have been blind to outcome, also potential for recall bias. Reclassification of cases and controls according to date of last period made no difference to results. Results are unadjusted for confounders.
Kesmodel, 2000; Denmark, cohort study	1989-91 & 1992-96 18,228 women	Questionnaire at 16 wks and 30 wks gestation; outcome from birth registration form	*g/wk at 16 weeks* % Risk ratio of <37 wks (95% CI) <12 4.3 1.00 12-24 3.9 0.91 (0.64, 1.15) 25-48 3.7 0.86 (0.52, 1.52) Mean GA (days) adj OR of <37 wks (95% CI) <12 281.3 1.00 12-24 281.8 0.90 (0.74, 1.10) 25-48 281.5 0.72 (0.51, 1.01) *g/wk at 30 weeks* % Risk ratio of <37 wks (95% CI) <12 3.9 1.00 12-24 2.7 0.69 (0.56, 0.86) 25-48 3.2 0.82 (0.60, 1.13) Mean GA (days) adj OR of <37 wks (95% CI) <12 281.7 1.00 12-24 281.4 0.70 (0.55, 0.88) 25-48 281.3 0.76 (0.55, 1.08)	Large population based study but only 73% returned questionnaires at 16 wks and only 63% returned 30 wk questionnaire. However, non-responders did not differ significantly from responders. Not clear over what period alcohol intake assessed. OR adjusted for smoking, caffeine, maternal age, height, pre-pregnant weight, marital and occupational status, education, parity, chronic diseases, previous preterm birth, parity and sex of child. Controlling for previous preterm birth may represent over-adjustment if previous event was also associated with alcohol consumption.

Lazzaroni, 1993; Italy, cohort study	Feb 1989 - July 1990 2145 live births	Postpartum interview; outcomes from clinical notes and paediatrician examination	OR of <37 wks (95% CI) 1-10g alcohol/day 0.64 (0.29, 1.15) compared to abstainers	GA estimated from date of last menstrual period. OR adjusted for maternal age and smoking habits, sex and birthweight of baby. Potential for recall bias due to information on drinking being gathered postpartum.
Lundsberg, 1997; USA, cohort study	1988-92 2714 singleton live births	Interview before 16 wks gestation; outcome from neonatal examination	*Alcohol at mth 1* *g/day* RR (95% CI) adj OR (95% CI) abstinent 1.00 1.00 <2.4 0.83 (0.49, 1.42) 1.02 (0.55, 1.90) 2.4-6 0.91 (0.48, 1.75) 1.09 (0.51, 2.33) *Alcohol at mth 7* abstinent 1.00 1.00 <2.4 2.11 (1.26, 3.54) 2.88 (1.64, 5.05) 2.4-6 2.15 (1.03, 4.52) 2.96 (1.32, 6.67)	Not clear how representative the HMOs and private practices are. GA estimated from date of last menstrual period. Adjusted for smoking in mth 7, height, bleeding during pregnancy, pre-/eclampsia, anomalies. Participation rate 76%, no information about non-respondents.
Marbury, 1983; USA, cohort study	about 1982 12,440 births	Postpartrum intervie; outcome from medical records	g/wk % preterm ***RR (95% CI)*** 0 7.7 ***1.00*** 1-24 6.8 ***0.88 (0.73, 1.06)*** 25-72 6.0 ***0.78 (0.58, 1.04)***	Method of estimation of GA not stated. % preterm unadjusted in this analysis. Adjustment not carried out in this analysis. Potential for recall bias due to timing of interview.
McDonald, 1992; Canada, cohort study	1982-84 40,445 pregnancies	Postpartum interview after birth; clinical data from medical records	g/wk % premature OR of <37 wks (95% CI) None 7.3 1.00 1-24 5.8 0.86 (0.77, 0.96) 25-72 6.2 0.90 (0.79, 1.04)	Method of estimation of GA not stated. OR adjusted for age, pregnancy order, previous spontaneous abortion & lbw infant, pre-pregnancy weight and employment, ethnic group, education, cigarette and coffee consumption. May represent over-adjustment if previous lbw baby was also associated with alcohol consumption. Potential for recall bias due to timing of interview.

Ogston, 1992; European multicentre cohort study	Period of recruitment not stated; 8469 women	Interview at 12-18 wks and 28-32 wks gestation; outcomes from hospital record	Alcohol (g/wk) <37 wks Before pregnancy: % ***RR (95% CI)*** Early pregnancy: % ***RR (95% CI)*** 0: 4.0 ***1.0*** / 3.9 ***1.0*** >0-29: 3.7 ***0.9 (0.7, 1.2)*** / 3.7 ***0.9 (0.7, 1.2)*** 30-59: 3.5 ***0.9 (0.6, 1.2)*** / 5.1 ***1.3 (0.9, 1.8)***	Representative of European population but different centres used slightly different methods. Unadjusted in this analysis. No information about participation rate
Passaro, 1996; UK, cohort study	1991-92 10,539 women	Alcohol exposure by questionnaire at 18 weeks; outcomes from hospital delivery records	<33 wks: % ***RR (95% CI)***; 33-36 wks: % ***RR (95% CI)*** Pre-pregnancy abstainers: 1 ***1.0*** / 6 ***1.0*** <12g /day: 1 ***0.9 (0.4, 2.1)*** / 5 ***0.8 (0.6, 1.0)*** Early pregnancy abstainers: 1 ***1.0*** / 5 ***1.0*** <12g /day: 1 ***1.1 (0.7, 1.6)*** / 5 ***1.0 (0.8, 1.1)***	Large population based study with high response rate. No adjustment for potential confounders in this analysis.
Parazzini, 2003; Italy, case control study	Period of recruitment not stated; 502 cases, 1966 controls	Both exposure and outcome data retrospectively by interview whilst women in postnatal ward	12g/day vs. 0: OR (95% CI) Before conception: 1.0 (0.7, 1.4) 1st trimester: 1.3 (0.9, 1.8) 2nd trimester: 1.2 (0.9, 1.7) 3rd trimester: 1.1 (0.8, 1.5)	No information on public/private mix of patients so generalisability unclear. Method for estimating GA not stated. Information on outcome and exposure all collected postpartum so potential for recall bias. Not clear whether interviewers blind to case/ control status. ORs adjusted for age, centre, education, parity, coffee, smoking, gestational hypertension and previous preterm birth (may be over-adjustment).
Peacock, 1995; UK, cohort study	1982-84 1513 women	Interviews at booking, 17, 28 & 36 wks gestation for alcohol consumption (source of outcome data not stated)	Alcohol consumption (g/wk): % preterm: ***RR (95% CI)*** none: 8.4: ***1.00*** 1-19: 4.6: ***0.55 (0.31, 0.98)*** 20-49: 6.2: ***0.74 (0.40, 1.36)***	Study principally about effects of socioeconomic factors on prematurity. Effects of alcohol unadjusted for confounders. Study restricted to white women so may not be representative.

Shiono, 1986; USA, cohort study	1974-7728, 330 women	Questionnaire at 1st clinic visit re. consumption in first 3 mths; outcome from medical records	<12g/day vs. 0 Preterm (24-36 wks gestation) Very preterm (24-32 wks gestation)	OR (95% CI) 0.89 (0.81, 0.99) 0.76 (0.59, 0.97)	Study mainly about effects of ethnicity on prematurity. Kaiser population largely middle class, not representative. Participation rate not stated. GA estimated from menstrual history and physical examination. OR adjusted for ethnicity, age, education, marital status, previous spontaneous and induced abortions, week prenatal care started and smoking.
Sulaiman, 1988; UK, cohort study	1985-86 952 babies	Interview at 1st antenatal visit and in 3rd trimester; outcomes from hospital records		Gestational age (wks) Abstainers 39.9 1-50g/wk 40.0	Population based study with 90% participation rate. Primiparous women only. Unadjusted in this analysis. GA calculated from Dubowitz score. No significant difference.
Verkerk, 1993; Netherland, cohort study	1978-79 2901 women	Interview in mid-pregnancy and immediately postpartum; source of outcome data not stated	Alcohol *1st trimester* abstainers 1-50g/wk *2nd trimester* abstainers ex-drinkers 1-50g/wk *3rd trimester* abstainers ex-drinkers 1-50g/wk	% preterm 4.9 3.6 4.9 3.5 4.2 4.8 3.8 3.0	unadj OR 1 0.73 1 0.71 0.84 1 0.80 0.62	adj OR (95%CI) 1 0.71 (0.45, 1.11) 1 0.65 (0.38, 1.11) 0.82 (0.52, 1.30) 1 0.76 (0.44, 1.32) 0.60 (0.35, 1.01)	Not clear how women selected, Dutch speaking only, may not be representative. Women in midwife care, at low risk of complications; 15% lost to follow-up mainly due to referral to obstetrician. However, early alcohol consumption was similar between women who were and were not referred. Information about 3rd trimester drinking obtained postpartum, potential for recall bias. Reference group were women who had abstained from alcohol both prior and during pregnancy. OR adjusted for smoking, education, employment, age and marital status.

Verkerk, 1994; Netherlands, cohort study	1988-89 2027 babies	Interview at about 3 wks after birth; outcomes from medical records	g/wk 0 1-12 13-84	% preterm 5.8 4.4 1.1	Unadj OR (95% CI) 1.00 0.75 (0.45, 1.24) 0.17 (0.02, 1.24)	Adj OR (95% CI) 1.00 0.80 (0.47, 1.36) 0.15 (0.02, 1.15)	Large population based sample with 97% response rate. However, interviews done approx. 3 weeks after birth so potential for recall bias. GA estimated from date of last menstrual period. Adjusted for smoking, social class, ethnicity, occupation, maternal age and height.
Wisborg, 1996; Denmark, cohort study	1989-91 4111 women	Questionnaire at 16 wks gestation; outcomes from attending midwife	g/wk <12 12-24 25-48	% preterm 4.3 3.7 4.9	***RR (95% CI)*** ***1.00*** ***0.87 (0.61, 1.27)*** ***1.15 (0.71, 1.88)***		Population based study with 99% participation rate. Focus of study was smoking and prematurity. Restricted to nulliparous women. Unadjusted for potential confounders in this analysis.

3.2.7 Malformations

There were six studies which examined the association between low-to-moderate alcohol consumption and incidence of malformations, including fetal alcohol effects, in the baby. They were all cohort studies, three from USA, one each from the UK, Australia and Denmark. Four studies analysed total malformations (Marbury et al, 1983; Ernhart et al, 1989; Lumley et al, 1985; Mills & Graubard, 1987); two studies included major malformations (Davis et al, 1982; Mills & Graubard, 1987); and two examined anomalies related to fetal alcohol effects (Ernhart et al, 1989; Olsen & Tuntiseranee, 1995). Major malformation was undefined in the study by Davis et al (1982) and defined as causing functional impairment or requiring surgical correction in the other study (Mills & Graubard, 1987). Only two studies adjusted for potential confounders in the relevant analyses (Ernhart et al, 1989; Mills & Graubard, 1987), the latter may have over-adjusted by including previous malformed infant or spontaneous abortions which may have been associated with alcohol exposure. Exposure to alcohol was assessed by interview or questionnaire antenatally in all but one study (Marbury et al, 1983). In four studies the neonatal assessment was done blind to alcohol consumption (Ernhart et al, 1989; Mills & Graubard, 1987; Olsen & Tuntiseranee, 1995; Stoler et al, 2002). In the other studies it was either not stated or data were from routine statistics (Lumley et al, 1985).

None of the studies reported a significant association between low-to-moderate alcohol consumption and malformations although a trend in that direction was apparent in some studies (Davis et al, 1982; Ernhart et al, 1989; Lumley et al, 1985).

Table 7 - Malformations

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Davis, 1982; UK, cohort study	1980 973 babies	Questionnaire at booking visit; outcomes from routine examination and notes	g/24 hrs % major malformations 0 0 1-8 0.6 χ^2 test $p < 0.01$	White women only. Study underpowered and unadjusted for potential confounders. Major malformations undefined.
Ernhart, 1989; USA, cohort study	3 year period (not stated when); 239-873 depending on measure	Alcohol exposure estimated 3 ways: *in pregnancy* by interview relating to 2 wks preceding each AN visit; *retrospective* by interview 5 yrs later about the index pregnancy; *embryonic* equation derived from separate sample to estimate exposure before woman knew she was pregnant. Outcomes from neonatal examination	g/day Mean Anomalies Tally *In pregnancy* / *Retrospective* / *Embryonic* 0 2.53 / 2.71 / 2.18 1-2.4 2.60 / 2.81 / 2.41 2.5-6 2.62 / 2.80 / 2.45 g/day Mean no. craniofacial anomalies *In pregnancy* / *Retrospective* / *Embryonic* 0 1.85 / 1.84 / 1.60 1-2.4 1.92 / 2.10 / 1.81 2.5-6 2.03 / 1.94 / 1.91 Multiple one tailed t-tests all non-significant	Disadvantaged population, oversampled women screening positive for alcoholism. Neonatal examination blind to exposure to alcohol in pregnancy. *Retrospective* data at 5 yrs unlikely to be accurate, potential for recall bias. Matched or adjusted for parity, smoking, race, date of recruitment, drug abuse, pre-pregnancy weight, weeks gestation at registration. Multiple outcomes investigated and multiple testing performed.

Lumley, 1985; Australia, cohort study	1981-82 10,319 births	All data from the Tasmanian perinatal statistics	g/wk % malformed (95% CI) 0 2.3 (1.8, 2.7) 1-24 2.5 (2.1, 2.9) 25-72 2.7 (1.3, 4.1)	Study based on all Tasmanian births 1981-82. However, only 69% data on alcohol consumption. Non-respondents similar in terms of parity and socioeconomic status (SES) but significantly lower birthweight. No adjustment for confounders in this analysis. Difference between abstainers and light drinkers not seen consistently within parity or SES.
Marbury, 1983; USA, cohort study	about 1982 12,440 births	Postpartum interview; outcomes from medical records	g/wk Malformations major: % ***RR (95% CI)***; minor: % ***RR (95% CI)*** 0 2.6 ***1.0*** 6.1 ***1.0*** 1-24 3.0 ***1.1 (0.8, 1.5)*** 6.3 ***1.0 (0.8, 1.3)*** 25-72 3.3 ***1.3 (0.8, 1.9)*** 4.9 ***0.8 (0.6, 1.1)***	% malformed unadjusted in this analysis. Adjustment only done on regrouped drinking variable. Potential for recall bias.
Mills, 1987; USA, cohort study	1974-77 32,870 babies	Questionnaire relating to 1st 3 mths; outcomes from discharge diagnoses, notes and autopsy reports, blind to alcohol exposure	g/day All malformations Rate (%) OR (95% CI) AdjOR (95% CI) 0 7.8 1.00 1.00 <12 7.7 0.99 (0.91, 1.08) 0.96 (0.87, 1.04) Major malformations 0 1.6 1.00 1.00 <12 1.8 1.12 (0.95, 1.33) 1.17 (0.97, 1.40)	Kaiser Permanente population may not be representative. Power to detect 12% increase in total malformations in women consuming <1 drink/day with 80% power. However, some malformations not immediately apparent. Adjusted for maternal age, race, education, parity, previous malformed infant or spontaneous abortion, smoking, diabetes mellitus, exposure to radiation. May be over-adjusted.

Olsen, 1995; Denmark, cohort study	1988-89 323 babies	Questionnaire at 1st antenatal visit and interview subsequently; outcomes measured from photos at birth and 18 months after birth	g/wk	Palpebral fissure	Nose upper lip	Root of nose	Oversampled women who consumed 5+ drinks per week. Measurement from photos blind to alcohol exposure during pregnancy but correlation between observers inconsistent. The data in previous column relate to newborns. A similar pattern was observed at 18 mths. However, features associated with FAS may not become apparent until 2 yrs. Unadjusted for confounders in this analysis.
			Before pregnancy	(standardised units)			
			0	20.8	3.3	.38	
			1-48	22.3	3.1	.41	
			1st trimester				
			0	22.4	3.0	.39	
			1-48	22.3	3.0	.42	
			1st & 2nd trimester				
			0	21.6	3.0	.41	
			1-48	22.5	3.1	.41	

3.2.8 Postnatal growth

There were only two studies which examined the association between alcohol exposure and growth as measured at birth or later in childhood. One of these studies, the Maternal Health Practices and Child Development Project (which has given rise to many papers, three of which fell within our inclusion criteria) followed 565 children up to age 14 (Day et al, 1990, 1999 & 2002). They found that children of women who drank up to 11g of alcohol per day in pregnancy were consistently lighter but height was not affected. The statistical significance of these findings was not reported. This study population was predominantly low socioeconomic status and may, therefore, be particularly susceptible to the effects of alcohol due, for example, to poor nutrition.

The study by O'Callaghan et al (2003) looked at weight of children 5 years after birth according to the amount of alcohol consumed by their mothers in early and late pregnancy. The proportion of children with weight in the bottom 3rd and 3rd to 10th percentiles was highest amongst children of abstainers except when estimated for late pregnancy. However, these differences were not statistically significant. Moreover, there was no adjustment for potential confounders in this analysis and follow-up at 5 years was only 47%.

In summary, there were only two studies which examined the association between alcohol exposure and growth as measured in childhood. One of these studies, which followed children up to age 14, found that children of women who drank small amounts in pregnancy were consistently lighter. However, the other study found the opposite, that children of abstainers tended to be lighter.

Table 8 - Postnatal growth

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Day, 1990, 1999 & 2002; USA, cohort study (Maternal Health Practices and Child Development Project)	1983-86 595 babies at birth, 462 at 8 mths gestation; wider cohort, 610 at 6 yrs, 557 at 10 yrs, 565 at 14 yrs	Interview at 4th and 7th mth gestation; outcome from examination by study nurses in participants' homes	g/day *in 3rd trimester*: None; 1-11 Adjusted mean weight at 8 mths(kg): 8.92 +/- 0.06; 8.81 +/- 0.09 *2nd trimester*: None; 1-11 Adjusted mean weight (lb) at 6 yrs: 51.8; 49.9 Adjusted mean weight (lb) at 10 yrs: 94.4; 87.5 Adjusted mean weight (lb) at 14 yrs: 149.5; 138-141 *1st trimester*: None; 1-11 Adjusted mean height/length (in) at 6 yrs: 47.2; 46.8 Adjusted mean height/length (in) at 10 yrs: 56.8; 56.8 Adjusted mean height/length (in) at 14 yrs: 65.3; 65-65.4	Population predominantly low socioeconomic status, 66% single parents 8 mths after birth. Oversampled women who drank 3 or more drinks per week. Methods for estimating alcohol exposure validated. Adjusted for maternal height, GA, weight gain during pregnancy, smoking, race & gender. 78% follow-up at 8 mths, 76% at 14 yrs but those lost to follow-up not significantly different. Infant examinations by study nurses with 90% reliability, blind to maternal substance use.
O'Callaghan, 2003; Australia, cohort study	1981-84 4038 at 5 yrs	Alcohol exposure retrospectively by interview at first antenatal visit and postpartum; head circumference measured by paper tape at 5 yrs	*Weight at 5 yrs* <3rd: % / ***RR (95% CI)*** 3-10th percentile: % / ***RR (95% CI)*** Alcohol in early pregnancy (g/day) Nil: 3.4 / ***1.0*** ; 7.5 / ***1.0*** 1-6: 2.6 / ***0.8 (0.5, 1.1)*** ; 6.3 / ***0.8 (0.7, 1.1)*** 7-11: 2.2 / ***0.6 (0.2, 2.6)*** ; 6.5 / ***0.9 (0.4, 1.9)*** Alcohol in late pregnancy (g/day) Nil: 3.2 / ***1.0*** ; 6.9 / ***1.0*** 1-6: 2.7 / ***0.8 (0.6, 1.3)*** ; 7.1 / ***1.0 (0.8, 1.3)*** 7-11: 0.8 / ***0.2 (0.0, 1.7)*** ; 9.1 / ***1.3 (0.7, 2.3)***	No adjustment for confounders in this analysis. Participation at birth 74%. Follow-up at 5 yrs only 47%.

3.2.9 Head circumference and length at birth

There were five studies which included these outcomes in an investigation of the effects of low-to-moderate alcohol consumption in pregnancy. One study examined birth length (Jacobson et al, 1994); one examined head circumference (O'Callaghan et al, 2003); and three examined both birth length and head circumference (Sulaiman et al, 1988; Day et al, 1990; Primatesta et al, 1994). All were cohort studies.

O'Callaghan et al (2003) reported a higher proportion of babies in the 3rd - 10th percentile among those whose mothers had consumed 7-11g of alcohol per day in pregnancy compared to those who had consumed less, while the proportion of babies below the 3rd percentile was highest among the abstainers. However, these differences were not statistically significant. Moreover, there was no adjustment for potential confounders in this analysis. None of the other studies reported any differences at these levels of consumption.

In summary, of the five studies reporting on these outcomes, only one found a higher proportion of low birth weight babies among those whose mothers drank low-to-moderate amounts in pregnancy. However, there was no adjustment for potential confounders in this analysis. None of the other studies reported any differences at these levels of consumption.

Table 9 - Head circumference and length at birth

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Day, 1990; USA, cohort study (Maternal Health Practices and Child Development Project)	1983-86 595 babies at birth, 462 at 8 mths; wider cohort, 610 at 6 yrs, 557 at 10 yrs, 565 at 14 yrs	Interview in 4th and 7th mth gestation; outcome from examination by study nurses in participants' homes	g/day — *1st trimester* — Adjusted mean height/length (in) at birth None — 19.4 1-11 — 19.4 *1st trimester* — Adjusted mean head circumference (mm) at birth None — 341.7 1-11 — 340.7	Population predominantly low socioeconomic status, 66% single parents at 8 mths. Oversampled women who drank 3 or more drinks per week. Methods for estimating alcohol exposure validated. Adjusted for maternal height, GA, weight gain during pregnancy, smoking, race & gender.78% follow-up at 8 mths, 76% at 14 yrs but those lost to follow-up not significantly different. Infant examinations by study nurses with 90% reliability, blind to maternal substance use.
Jacobson, 1994; USA, cohort study	Neither year nor period of recruitment stated; 412 babies	Alcohol exposure collected for fortnight preceding each AN visit by interview; growth at 6.5 mths blind to exposure	g/day — % infants in bottom 10th centile length at birth — ***RR (95% CI)*** 0 — 6.1 — ***1.00*** 1-6 — 9.6 — ***1.58 (0.57, 4.41)*** 7-12 — 6.1 — ***1.00 (0.19, 5.18)***	Black infants only; inner city deprived population. Excluded if baby <1500g or <32 wks GA, malformed or a multiple. Oversampled drinkers of at least 0.5oz AA/day. Participation rate and follow-up rates not stated. Adjusted for smoking, opiate and cocaine use, maternal age, pre-pregnancy weight, height, education, marital status, welfare status, parity, infant's age, sex, number of antenatal visits.

O'Callaghan, 2003; Australia, cohort study	1981-84 6320 at birth	Alcohol exposure retrospectively by interview at first AN visit and postpartum; head circumference measured by paper tape at birth	*Head circumference*					No adjustment for confounders in this analysis. Participation at birth 74%. No information about non-responders.
				<3rd		3-10th percentile		
				%	***RR (95% CI)***	%	***RR (95% CI)***	
			Alcohol in early pregnancy (g/day)					
			Nil	3.6	***1.0***	5.7	***1.0***	
			1-6	3.1	***0.9 (0.6, 1.3)***	5.8	***1.0 (0.8, 1.2)***	
			7-11	2.9	***0.8 (0.3, 1.9)***	7.0	***1.2 (0.7, 2.2)***	
			Alcohol in late pregnancy (g/day)					
			Nil	3.5	***1.0***	5.7	***1.0***	
			1-6	2.7	***0.8 (0.6, 1.1)***	5.5	***1.0 (0.8, 1.2)***	
			7-11	2.3	***0.7 (0.3, 1.6)***	6.9	***1.2 (0.7, 2.0)***	

Primatesta, 1994; Italy, cohort study	1986-87 1516 women	Alcohol exposure by postpartum interview; outcome from clinical record	*Alcohol intake before pregnancy* Mean birth length Unadjusted Adjusted (g/wk) Males Females Males Females 0 50.0 49.0 51.2 49.1 1-20 50.2 49.1 50.2 49.3 *Alcohol intake during pregnancy* Mean birth length Unadjusted Adjusted (g/wk) Males Females Males Females 0 50.1 48.8 50.2 49.0 1-20 50.2 49.2 50.2 49.2 *Alcohol intake before pregnancy* Mean head circumference Unadjusted Adjusted (g/wk) Males Females Males Females 0 34.6 33.8 34.8 33.8 1-20 34.7 33.9 34.7 34.0 *Alcohol intake during pregnancy* Mean head circumference Unadjusted Adjusted (g/wk) Males Females Males Females 0 34.7 33.7 34.8 33.8 1-20 34.7 34.0 34.7 34.0	Study took place in 2 Milan hospitals, representativeness was not described but high risk pregnancies excluded. Participation rate not stated. Potential for recall bias due to postpartum interview re. alcohol exposure. Adjusted for smoking, age, GA, maternal weight, height and parity, paternal weight. Type of alcohol consumed made no difference.
Sulaiman, 1988; UK, cohort study	1985-86 952 babies	Interview at 1st AN visit and in 3rd trimester; outcomes from hospital records	Length (mm) Head circum-ference (cm) Abstainers 504.2 46.3 1-50g/wk 504.0 45.9	Population based study with 90% participation rate. Primiparous women only. Unadjusted in this analysis.

3.2.10 Neurodevelopmental outcomes

There was one study which examined neurodevelopmental outcomes at birth (Streissguth et al, 1983) and six studies which examined the longer term neurodevelopmental outcomes associated with low-to-moderate level alcohol exposure. Four of these used the Bayley scales of infant development (Forrest et al, 1991; Parry & Ogston, 1992; Jacobson et al, 1993 and 1999; Olsen, 1994), one used the Wide Range Achievement Test (WRAT), the Peabody Individual Achievement Test (PIAT) and teacher assessment (Goldschmidt et al, 2004), one used the Achenbach Child Behaviour Checklist (CBCL) (Sood et al, 2001) and one used the Brazelton Neonatal Assessment (Streissguth et al, 1983). None of the studies using the Bayley scales reported any significant effect at low-to-moderate levels of alcohol. All but one (Streissguth et al, 1983) of these studies was adjusted for relevant confounders, although the two EuroMac papers (Parry & Ogston, 1992; Olsen, 1994) did not adjust for gestational age or birthweight.

The study using the WRAT, PIAT and teacher assessment (Goldschmidt et al, 2004) carried out multiple comparisons but there were no significant differences at the 5% level.

The study using the Achenbach Child Behaviour Checklist (CBCL) (Sood et al, 2001) was the only one to find consistently poorer results among the children of low level drinkers. However, at low-to-moderate levels of consumption there were no statistically significant differences in scores. Moreover, the analysis comparing no exposure to low exposure was unadjusted for potential confounders.

Of the seven studies which looked at this outcome, one was conducted at birth, the others were later in childhood. Only one study found poorer results in children of low-to-moderate drinkers. However, this analysis was unadjusted for potential confounders and the difference was not statistically significant.

Fuller descriptions of the neurodevelopmenatal tests are given in appendix 4.

Table 10 - Neurodevelopmental outcomes

1st author, year of publication, country, study type	Period & numbers recruited, age at assess-ment	Measures of alcohol exposure and outcome	Results	Comments
Forrest, 1991; UK, cohort study	1985-86 592 children at 18 mths	Interview at 1st AN visit and in 3rd trimester; outcomes assessed by psychologist blind to alcohol exposure in hospital using Bayley scales	*Regression coefficients (95% CI) for consumption of 1-49g compared with abstainers adjusted for confounders* Mental develop-ment index / Psychomotor development index Before pre.g. 4.5 (-1.3, 10.3) / 3.6 (0.6, 6.7) Early pre.g. 0.8 (-2.2, 3.8) / 0.4 (-1.2, 2.0)	Study limited to primiparous women, representative of population but oversampled heavier drinkers. Adjusted for smoking, maternal age, social class, child sex and GA. Follow-up only 70%, no data on those lost to follow-up. When analysis restricted to full term infants, and then full term and not admitted to special care, results were not significantly different.

Goldschmitt, 2004; USA, cohort study	1984-87 606 children at age 10	Alcohol by interview at each trimester; outcomes assessed blind by trained researchers in child's home. Wide Range Achievement Test (WRAT) Peabody Individual Achievement Test (PIAT) Teachers' rating: 1=bottom 10% 3=average 5=top 10% Underachievement= disparity between WRAT and intellectual ability measured by Stanford-Binet Intelligence Scale	*1st trimester* Reading recognition (WRAT) Spelling (WRAT) Maths (WRAT) Reading comprehension (PIAT) Teacher's rating % underachieving *2nd trimester* Reading recognition (WRAT) Spelling (WRAT) Maths (WRAT) Reading comprehension (PIAT) Teacher's rating % underachieving *3rd trimester* Reading recognition (WRAT) Spelling (WRAT) Maths (WRAT) Reading comprehension (PIAT) Teacher's rating % underachieving	*none* 93.4 92.5 88.3 93.8 3.0 11.2 94.7 94.1 90.2 95.2 3.1 9.2 94.3 93.8 88.9 94.2 3.0 9.7	*<12g /day* 95.6 95.3 89.9 94.7 3.1 9.4 95.6 94.7 88.3 93.6 3.1 10.2 95.3 94.3 90.4 94.8 3.1 12.3	Participants predominantly low income, not truly representative. Follow-up rate 79% but those lost to follow-up did not differ significantly in terms of demographics or alcohol exposure. Further, 110 teachers did not complete rating but, again, no significant differences in terms of school achievement. No adjustment for confounders in this analysis. All comparisons were non-significant at the p = 0.05 level.

Jacobson, 1993; USA, cohort study 2 papers, Jacobson, 1993 and Jacobson, 1999	1986-89 382 children at 6.5 and 13 mths	Alcohol by interview at each AN visit relating to previous fortnight; outcomes assessed blind by trained researchers in university lab. Bayley scales – Mental Development Index (MDI), Psychomotor Development Index (PDI); McCall index	g/day 0 1-12 ***RR*** g/day 0 1-6 7-12 g/day 0 1-6 7-12	No (%) scoring 1 sd above/below mean of residual Processing speed: 9/50 (18%); 25/225 (11%); ***0.62 (0.31, 1.24)*** Elicited play: 3/53 (6%); 29/222 (13%); ***2.30 (0.73, 7.29)*** infants in bottom 10% Bayley Scales MDI: % 6.7; ***adj RR 1.0*** / % 9.4; ***1.4 (0.5, 4.0)*** / % 5.4; ***0.8 (0.2, 4.2)*** PDI: % 8.6; ***1.0*** / % 8.8; ***1.0 (0.4, 2.6)*** / % 8.1; ***1.0 (0.2, 3.8)*** infants in bottom 10% McCall index % 6.9; ***adj RR 1.0*** / % 8.8; ***1.3 (0.5, 3.6)*** / % 5.4; ***0.8 (0.2, 4.2)***	Only included black infants 1500g+, >31 wks gestation singletons without chromosomal anomalies or NTD. Area served is predominantly deprived population. Participation rate not stated. Those lost to follow-up significantly older, less likely to be on welfare, higher education. Values adjusted for maternal age, parity, no. AN visits, quality of parenting, opiates, smoking, age of child at visit, examiner. Processing speed based on Visual Recognition Memory using Novelty Preference Test and cross-modal fixation duration.
Olsen, 1994; Denmark, cohort study (EuroMac study)	1988-89 276 children at 18 mths, 217 at 3.5 yrs	Alcohol by questionnaire at 32 wks; Bayley at 18 mths and Griffiths at 3.5 yrs by psychologist blind to exposure and previous reports	*Alcohol intake* g/wk <1 1-48 <1 1-48	Bayley at 18 mths PDI unadj: 107; 106 / adj: 107; 106 / MDI unadj: 107; 106 / adj: 106; 106 Griffiths overall at 3.5 yrs Unadjusted: 106; 107 / adjusted: 105; 107	Included women consuming 5 or more drinks per week matched 1:1 with lower level drinkers. Matched on expected date of delivery and woman's year of birth. Initial sample representative of population but follow-up at 3.5 yrs 76% and only 66% included in analysis. No data about those lost to follow-up. Adjusted for parents' education, type of residence and smoking.

Parry, 1992; Berlin, Germany, Odense, Denmark, Dundee, Scotland, cohort study (EuroMac)	Period and year of recruitment not stated; Dundee 592, Odense 286, Berlin 522	Interview at 1st AN visit and at 28-32 wks relating to drinking in previous wk; outcomes by Bayley assessment in hospital or home by psychologist at 18 mths.	Alcohol g/wk 0 1-29 30-59 Regression coefficients MDI (se) 0 1.80 (1.1) 2.34 (1.6)	Mean MDI 103 105 105 PDI (se) 0 0.81 (0.8) -0.05 (1.1)	Mean PDI 104 105 104	Participation rate not stated. Not clear if assessment done blind to alcohol exposure. Regression coefficients adjusted for smoking, sex, birthweight, GA, mother's age, parity and education show no significant effect. Follow-up at 18 mths was 52% in Berlin, 70% in Dundee and 86% in Odense. No data on those lost to follow-up.
Sood, 2001; USA, cohort study	1989-91 501 children at 6-7 yrs	Alcohol by interview at each antenatal visit relating to previous fortnight; outcomes assessed blind by trained researchers in research facility. Achenbach Child Behaviour Checklist (CBCL) clinical cut-off at 60 or 67 for syndrome/total and problem sub-scales respectively.	Mean CBCL raw scores Externalising Internalising Social problems Attention problems Thought problems Total score % of CBCL above clinical cut-off Externalising Internalising Social problems Attention problems Thought problems Total score ***RR of CBCL above clinical cut-off*** ***Externalising*** ***Internalising*** ***Social problems*** ***Attention problems*** ***Thought problems*** ***Total score***	*Prenatal alcohol* 0 8.5 5.1 1.9 3.2 0.7 23.7 14.8 11.3 7.9 12.5 9.6 15.7 ***1.56 (0.96, 2.52)*** ***1.56 (0.88, 2.75)*** ***1.41 (0.70, 2.84)*** ***0.99 (0.57, 1.72)*** ***0.99 (0.51, 1.91)*** ***1.57 (0.98, 2.50)***	<7g /day 11.1 6.4 2.4 3.7 0.8 29.0 22.5 17.2 10.7 12.8 9.4 24.1	Sample restricted to black women with singletons, oversampled women consuming >12g alcohol/day. Excluded women with HIV and those without antenatal care. Extensive screening for alcohol and drug abuse may have reduced participation. Not clear how 4800 women reduced to 665. Further exclusions and loss to follow-up reduced sample to 501 (75%). Non-respondents were older and of higher parity but children did not differ. Unadjusted for potential confounders in this analysis. Mean GA and birthweight lower in exposed group. Alcohol accounted for 1-2% of the variance.

Streissguth, 1983; USA cohort study (Seattle study)	1974-75 417 babies	Alcohol exposure by interview at 5 mths in their own homes; Brazelton Neonatal Assessment carried out within 35 hrs of birth by examiners blind to exposure	g/day 0 1-2.4	Mean Habituation 17.0 16.0	Mean Low arousal 9.3 8.8	Predominantly white, middle class, married. Participation rate 89%. Oversampled for heavier drinkers and smokers. Excluded multiple births. Unadjusted for potential confounders.

3.3 Binge drinking

The national alcohol harm reduction strategy for England defines a binge episode (for women) as six or more units in a single session, which is equivalent to the US definition of four drinks or more on one occasion. We decided to use the authors' definition of binge drinking since there was such variability in definition.

There were 11 separate studies (counting all the papers from the Seattle study as one) meeting the inclusion criteria for this review that included information on binge drinking during pregnancy. Binge drinking was most commonly defined as consuming 5 or more drinks on a single occasion (2.5 oz or 60g of alcohol), but has also been defined as 10 or more drinks (Plant & Plant, 1988) and 40-45g (equivalent to about 3.5 drinks – Passaro et al, 1996). One study only considered a woman to be a binge drinker if she consumed 5 or more drinks on an occasion at least once in every fortnight of her pregnancy (Bailey et al, 2004). It was sometimes unclear whether the women were otherwise heavy drinkers or not. Many of the studies considered multiple outcomes so they have been ordered alphabetically by first author in the table. However, the outcomes will be considered separately here.

3.3.1 Birthweight, gestational age and growth

Seven of the studies considered these outcomes (Bell & Lumley, 1989; Tolo & Little, 1993; Whitehead & Lipscomb, 2003; Sampson et al, 1994; Passaro et al, 1996; O'Callaghan et al, 2003; Nulman et al, 2004). Only two of these studies found an association between binge drinking and birthweight (Sampson et al, 1994; Passaro et al, 1996). The first of these (part of the Seattle study) reported a modest correlation between bingeing both prior to pregnancy recognition and during pregnancy, and birthweight (-.15 and -.11 respectively). However, the statistical significance of this was not stated. Length, head circumference and subsequent weight (up to 14 years after birth) were not associated with bingeing. Moreover, these results were unadjusted for potential confounders. The other study which reported an association between birthweight and binge drinking (Passaro et al, 1996) only found a significant association in the group who were bingers and/or heavy drinkers (1-2 drinks per day in early pregnancy/binged at least once in mid-pregnancy, or drank 3+ drinks per day in early pregnancy without bingeing in mid-pregnancy). Thus, it is difficult to separate out the effect of binge drinking from heavy drinking. These analyses were also unadjusted for possible confounders. A further study (Bell & Lumley, 1989) reported significantly lower birthweight in abstainers.

3.3.2 Birth defects

There were three studies which considered this outcome (Plant & Plant, 1988; Bell & Lumley, 1989; Olsen & Tuntiseranee, 1995). The first of these counted the mean number of abnormalities at birth and found a significant excess in bingers, particularly if they also smoked 10 or more cigarettes per day. However, a binge was defined as 10 or more drinks on a single occasion, and the analyses were not adjusted for potential confounders other than smoking. The study by Bell & Lumley (1989) had serious problems with participation and completeness of data. They found a slight excess of birth defects but this difference was not statistically significant. The study by Olsen & Tuntiseranee (1995) was a study of the craniofacial features of FAS. They found that newborn children of binge drinkers had slightly shorter palpebral fissures.

3.3.3 Neurodevelopmental outcomes

Four studies considered these outcomes in relation to binge drinking (Streissguth et al, 1983, 1989 & 1990; Olsen, 1994; Bailey et al, 2004; Nulman et al, 2004). Two of these used the Bayley Scales of Infant Development at 18 mths (Olsen, 1994) and up to 36 months after birth (Nulman et

al, 2004) but neither found a statistically significant difference in score in children of women who binged in pregnancy. The only difference found by Nulman et al (2004) was a greater degree of ‘disinhibited behaviour’ as shown in the significantly higher scores for adaptability and approach. However, this study did not collect any data on maternal behaviour which may be a confounder. The study by Bailey et al (2004) reported a significant reduction in verbal IQ and increase in delinquent behaviour in children of women who had binged in pregnancy. However, this study only counted women as bingers if they binged throughout pregnancy, not just a single occasion. The Seattle Longitudinal Prospective Study on Alcohol and Pregnancy (Streissguth et al, 1989 & 1990) followed children up to age 14 using a variety of tests. They reported significantly more learning problems, and poorer performance as assessed by both parents and teachers, in children of bingers. This effect appeared to persist up to age 14. The proportion lost to follow-up was not stated but may have been quite substantial, around 30%, which may have affected the results. Nevertheless, this study represents the strongest evidence yet of a neurodevelopmental effect of binge drinking.

In summary, there were 11 separate studies which examined the effect of binge drinking on the outcomes above. Only the four studies that looked at neurodevelopmental outcomes showed consistently poorer results in children exposed to binge drinking in pregnancy. Effects, which were generally quite small, included an increase in ‘disinhibited behaviour’, a reduction in verbal IQ and increase in delinquent behaviour, and more learning problems and poorer performance.

Table 11 - Binge drinking

1st author, year of publication, country, study type	Period & numbers recruited	Measures of alcohol exposure and outcome	Results	Comments
Bailey, 2004; USA, cohort study	1989-91 499 children at age 6-7 yrs	Alcohol exposure by detailed interview at each antenatal visit regarding previous 2 wks intake; outcomes assessed at research facility by researchers blind to alcohol exposure, teacher data by questionnaire	Pearson correlation / p-value Verbal IQ -.12 .00 Performance IQ -.02 .61 Aggressive behaviour .07 .13 Delinquent behaviour .12 .01 Regression for Verbal IQ -2.47 .014 Regression for Delinquent behaviour 2.26 .024	Study included black mothers only, oversampled alcohol exposed pregnancies. Excluded multiples, malformations, HIV. Bingeing defined as 5+ drinks on a single occasion at least once in every 2 wks during pregnancy. Follow-up only 75%, no data on those lost to follow-up. Regression included quality of home environment, SES, child violence exposure, AN exposures, cigarettes and cocaine. Numerous covariates also examined. Babies of bingers had lower birthweight but birthweight was not included in the regression analysis.
Bell, 1989; Australia, cohort study	1985 8884 women	All data from the Victorian Perinatal Morbidity Statistics	*Mean birthweight (sd)* Binge drinkers 3352 (518) Non-bingers (also non-smokers) 3420 (556) Smoking abstainers 3256 (576) *All differences were significant at p <0.05* *% Congenital malformations* Binge drinkers 1.9 Non-bingers 1.2 Abstainers 1.3 *No differences significant*	45% of hospitals took part of which 65% recorded smoking and alcohol details. Only adjusted for smoking (some of the time). Bingeing defined as any episode of 5+ drinks. Possible selection bias due to smoking and drinking details being less well recorded in cases of poor outcome.

Nulman, 2004; Canada, Cohort study	1987-97 50 women, 51 babies assessed at various ages 0-9 yrs	Alcohol exposure by interview antenatally; outcomes assessed at research facility by psychometrist blind to alcohol exposure Bayley - up to 36 mths (n=26) McCarthy - up to 7 yrs (n=22) Weschler - > 7 yrs (n=3)	*mean ± sd*	Bingers	Non-bingers	p	Participants were women who contacted the Motherisk program out of concern at exposure to alcohol or other substance. In 10 yrs only 92 'eligible' women identified but eligibility undefined. Controls were women consulting about an exposure other than alcohol (teratogens excluded) matched on age, SES, smoking, time of conception and child's age. No data about alcohol consumption in controls. Bingers (defined as any episode of 5+ drinks) were not alcoholics. Numerous neuropsychological tests performed and reported. Significant differences appeared for 3/9 subscales on a temperament/behaviour score only. Multiple regression included maternal IQ, SES, Parenting stress index and GA. Frequent bingers (n=12, defined as >6 binges) numbers were too small to be useful. Similarly there were only 3 children over 7 yrs who were given the Weschler. Study did not collect measures of maternal behaviour.
			Gestational age	39.9 ± 1.8	39.5 ± 1.7	.2	
			Birthweight	3390 ± 609	3376 ± 565	.9	
			Weight (%ile)	60.6 ± 25	57.9 ± 31	.6	
			Height (%ile)	55.2 ± 25	64.0 ± 25	.08	
			Head circum-ference (%ile)	51.8 ± 20	56.6 ± 23	.3	
				Difference (95% CI)			
			Bayley MDI	-2.83 (-13.1, 7.5)			
			Bayley PDI	-4.31 (-10.5, 1.8)			
			McCarthy				
			GCI	-5.92 (-18.9, 7.0)			
			verbal	1.93 (-4.8, 8.7)			
			perceptual	-1.45 (-6.5, 3.6)			
			quantitative	-0.19 (-6.5, 6.2)			
			memory	-0.05 (-5.9, 5.6)			
			motor	0.59 (-5.5, 6.6)			
			Multiple regression of no. binges on outcomes				
				B (95% CI)			
			Adaptability	1.50 (0.52, 2.49)			
			Approach	1.69 (0.74, 2.66)			
			Distractability	0.54 (-0.41, 1.48)			

O'Callaghan, 2003; Australia, cohort study	1981-84 6320 at birth, 4038 at 5 yrs	Alcohol exposure retrospectively by interview at first AN visit and postpartum; head circumference measured by paper tape at birth and 5 yrs	*Head circumference at birth and at 5 yrs* % <3rd 3-10th percentile birth 5yrs birth 5yrs Binge drinking Nil 3.3 2.9 5.4 6.9 < half time 3.9 2.9 7.2 7.8 > half time 2.3 3.4 6.4 5.9 p = 0.11 0.9 *Weight at birth and at 5 yrs* % <3rd 3-10th percentile Binge drinking birth 5yrs birth 5yrs Nil 2.8 2.8 6.6 7.1 < half time 3.6 3.5 8.8 5.8 > half time 3.1 1.7 7.8 10.1 p = 0.02 0.3	No adjustment for confounders in this analysis. Participation at birth 74%. Binge defined as 5+ drinks at one time, frequency of binges categorised rather vaguely as never, <half the time or half the time or more in pregnancy. Only asked at 1st visit. Follow-up at 5 yrs only 47%.
Olsen, 1994; Denmark, cohort study Part of EuroMac study.	1988-89 276 children at 18 mths, 217 at 3.5 yrs	Alcohol by questionnaire at 32 wks gestation; Bayley at 18 mths after birth and Griffiths at 42 months after birth by psychologist blind to exposure and previous reports	*Alcohol intake* Bayley at 18 mths *No. binges* PDI unadj adj MDI unadj adj 0 107 108 106 106 1-4 104 104 104 104 5+ 103 101 104 104 Griffiths overall at 3.5 yrs N Unadjusted adjusted 0 150 107 107 1-4 56 106 106 5+ 11 106 106	Included women consuming 5 or more drinks per week matched 1:1 with lower level drinkers. Matched on expected date of delivery and woman's year of birth. Initial sample representative of population but follow-up at 3.5 yrs 76% and only 66% included in analysis. No data about those lost to follow-up. Adjusted for parents' education, type of residence and smoking. Bingeing defined as 5+ drinks on a single occasion at any time in pregnancy.

Olsen, 1995; Denmark, cohort study	1988-89 323 babies	Questionnaire and interview; outcomes measured from photos at birth and 18 months after birth	*No. of Binges in 1st trimester*	Palpebral fissure	Nose upper lip	Root of nose	Oversampled women who consumed 5+ drinks per week. Measurement from photos blind to alcohol consumption but correlation between observers inconsistent. Unadjusted for confounders in this analysis.
			Newborns				
			0	22.7	3.0	.41	
			1-4	22.0	3.1	.42	
			5+	20.4	2.7	.42	
			18 months				
			0	16.1	3.1	.55	
			1-4	15.7	3.0	.50	
			5+	14.9	3.3	.59	

Passaro, 1996; UK, cohort study	1991-92 10,539 babies	Alcohol exposure by questionnaire at 18 weeks gestation; outcomes from hospital delivery records	Mean birthweight (SD) / Mean GA (SD) Non-bingeing abstainers 3397 (512) 40.1 (2.1) Non-bingeing occ-asional drinkers 3419 (488) 40.1 (1.9) Bingeing occasional drinkers 3408 (498) 40.1 (2.0) Nonbingeing light daily drinkers 3401 (441) 39.8 (1.9) Heavy drinkers/ bingers 3222 (538) 40.1 (2.3) Adjusted mean differences in birthweight (95% CI) compared to prepregnancy drinkers who abstained during pregnancy Non-bingeing abstainers -36 (-71, -1) Non-bingeing occ-asional drinkers -3 (-24, 16) Bingeing occasional drinkers -7 (-33, 177) Non-bingeing light daily drinkers 52 (-73, 177) Heavy drinkers/bingers -152 (-228, -76)	Large population based study with high response rate. Binge defined as 40-45g alcohol on a single occasion in the first 18 weeks of pregnancy. Mean differences in birthweight adjusted for GA, infant sex, parity, maternal smoking and BMI showed no significant differences except for lower birthweight in non-bingeing abstainers and in heavy drinkers/ bingers. Type of alcoholic drink made no difference to results. Limiting analysis to exclude users of marijuana, crack and cocaine, and excluding women with a history of alcoholism made no difference.
Plant, 1988; UK, cohort study	1980-83 1008 women	Interview in 3rd month of pregnancy; outcomes from case notes	Mean no. abnormalities noted at birth Bingers 2.0 Abstainers 1.2 *p<0.05* Bingers and smoked 10+ cigarettes daily 2.4 Non-smoking abstainers 1.0 *p<0.001*	Bingeing defined as drinking 10+ units on a single occasion in first trimester of pregnancy. Not all analyses presented. Participation rate not stated. Not adjusted for other confounders in this analysis.

Sampson, 1994; USA; cohort study (Seattle study)	1974-75 1439 women, follow-up of 500 children	Alcohol exposure retrospectively by interview at 5 months gestation; birth outcomes from medical records, 8 and 18 mths, 4, 7 and 14 yrs after birth by examination	Unadjusted correlations *Bingeing prior to pregnancy recognition* Birth 8mo 18mo 4y 7y 14y Weight -.15 .01 .05 -.03 .02 -.01 Length -.01 .05 .08 .03 .04 .04 Head circum-ference -.07 .01 -.01 -.02 -.01 -.03 *Bingeing during pregnancy* Birth 8mo 18mo 4y 7y 14y Weight -.11 .01 .05 .02 .04 .06 Length -.06 .02 .04 .02 -.01 .05 Head circum-ference -.08 -.02 -.02 -.02 -.04 -.04	Mainly white, married, middle class women. 500 women who were selected for follow-up were selected for high level alcohol consumption. Binge defined as 5+ drinks on a single occasion in first 5 months of pregnancy. Analyses unadjusted for confounders (except at 8 mths after birth where heavier drinkers were inadvertently oversampled). Number lost to follow-up not stated. Strongest correlations were at birth and at 8 months. Possible ‘catch-up’ growth thereafter.

Streissguth, 1981, 1990 and 1994; USA; cohort study (Seattle study)	1974-75 486 children at age 7 yrs 359 children at age 14 yrs	Alcohol exposure by interview at 5 mths gestation in their own homes; Weschler Intelligence Scale for Children (WISC), Wide Range Achievement Test (WRAT), Conners carried out between 6.5-7.5 yrs by examiners blind to exposure	WRAT at age 7	R	Adjusted achievement decrement (p)		Predominantly white, middle class, married. Oversampled for heavier drinkers and smokers. Excluded multiple births. Participation rate 85%. Pre-pregnancy binges defined as 5+ drinks on any one occasion in first 5 months of pregnancy. Follow-up rate not stated. Inter-rater unreliability on WISC and Word Attack scores. WRAT adjusted for maternal and paternal education, no. children in household, household stress, prenatal nutrition, smoking, aspirin and caffeine, sex, race and grade of child, exam conditions. Multiple regression (R) and age 14 years correlations indicates test scores against a binary binge drinking measure.
			Reading	0.55	-3.3 (.038)		
			Spelling	0.58	-1.8 (.227)		
			Arithmetic	0.62	-3.3 (.000)		
			Learning problems		Binge %	No-binge %	p
			Connors				
			learning problems		17	7	.000
			hyperkinesis		15	9	.091
			impulsivity		13	8	.104
			Parental assessment				
			below average		13	4	.000
			Myklebust (teacher)				
			MPRS <65 cutoff		17	10	.037
			School special class		24	15	.011
			<age approp. grade		19	14	.147
			Mean WISC score				
			full scale IQ		104.8	108.7	
			verbal IQ		103.5	107.2	
			performance IQ		105.3	108.8	
			Age 14 correlations				
					Word Attack	Arithmetic	
			Pre-pregnancy		-0.15 p<.01	-0.15 p<.05	
			In pregnancy		-0.12 p<.05	-0.15 p<.05	

Tolo, 1993; USA, cohort study	1982-83 237 1 or more binges, 472 no binges	Questionnaire at 6 months gestation; birth outcomes from medical records; IUGR computed by regressing birthweight on gestational age and adding residuals to sample mean	Outcome	Any binge	No binge	Difference in means (95% CI)	Population source was HMO, consecutive sample overwhelmingly white and well educated, not truly representative. Excluded regular drinkers from sample. Questionnaire was validated in pilot study but had 74% response rate, no information about non-respondents. Results were unadjusted; however, adjusting for confounders made little difference. Similarly, number of binge episodes (binge defined as 2.5oz on a single occasion in first 6 months of pregnancy) and timing (whether in month before pregnancy or in 1st/ 2nd trimester) made no difference to results.
			Birth-weight (g)	3561	3579	-18 (-95, 59)	
			Birth length (cm)	51.2	51.0	0.2 (-0.2, 0.6)	
			Head circum-ference (cm)	34.5	34.6	-0.1 (-0.4, 0.2)	
			Intrauterine growth (g)	3537	3591	-54 (-122, 14)	
Whitehead, 2003; USA, cohort study	1996-99 50,461 women	Postpartum questionnaire with telephone follow-up of non-responders; outcomes from birth certificate data		%SGA		RR (95% CI)	SGA defined as birthweight <10th percentile for gestational age according to race and parity specific standards. Bingeing defined as 5 or more drinks on a single occasion. 76% response rate but over-representation of SGA babies. ORs adjusted for maternal age, education, marital staus, pre-pregnancy weight, public assistance, state of residence, smoking before pregnancy.
			3 mths before pregnancy				
			Bingeing	8.1		1.06 (0.94, 1.21)	
			No bingeing	7.6			
			last 3 mths of pregnancy				
			Bingeing	9.4		1.20 (0.78, 1.87)	
			No bingeing	7.8			
			per binge relative to non-drinkers				
						OR (95% CI)	
			3 mths before pregnancy			0.98 (0.97, 1.00)	
			last 3 mths of pregnancy			0.99 (0.90, 1.09)	

3.4 Quality of papers included in systematic review

The quality of the studies was assessed using the Newcastle-Ottawa Quality Assessment Scale. Although this is recommended by the Cochrane Collaboration for observational studies, there were some specific issues pertaining to studies in this area to which it was not sensitive. These included:

- When was the woman asked about her alcohol consumption? If it was postnatally then, depending when the outcome became manifest, there is the potential for recall bias.
- Recall of alcohol consumption generally – Women are likely to under-report their consumption in pregnancy. The better studies used detailed interview schedules asking about alcohol consumption over the preceding two weeks, tied in to particular activities and times of day.
- Timing of alcohol consumption – Many studies did not ask (or did not report) when in pregnancy or how long prior to pregnancy the alcohol consumption related to.
- Confounding – the Newcastle-Ottawa Quality Assessment Scale asks whether the study controls for the most important confounders. Generally they did, but often the specific analyses relevant to this review were not adjusted. Residual confounding may also have occurred.
- Over-adjustment – Some studies controlled for previous adverse pregnancy outcome because such events may be linked by factors other than alcohol. However, this may also mean that if the previous adverse event was associated with alcohol consumption than controlling for it loses information.
- Many studies reported the statistical significance of their results over the full spectrum of alcohol use as a trend. We were only able to use the information that related to less than one drink per day so the statistical significance of those data was not known.

Other elements of the Quality Assessment Scale were unhelpful in discriminating between studies. For example, in cohort studies it asks about selection of the non-exposed cohort. In these studies they were almost always drawn from the same community. Similarly, the outcome of interest was never present at the start of the study.

Table 12 – Quality of included papers (see appendix 4 for details)

	Cohort studies			Case-control studies		
	Selection (out of 4)	Comparability (out of 2)	Outcome (out of 3)	Selection (out of 4)	Comparability (out of 2)	Exposure (out of 3)
Albertsen K	4	2	3			
Armstrong BG	1	2	1			
Bailey BN	3	2	2			
Bell R	3	1	3			
Berkowitz GS				4	0	2
Brooke OG	4	2	3			
Davis PJ	2	1	2			
Day NL 1990	3	2	2			
Day NL1999	3	2	2			
Ernhart CB	3	2	1			
Faden VB	4	2	2			
Forrest F	4	2	2			
Goldschmidt L	3	2	2			
Harlap S	3	2	3			
Henriksen TB	2	2	3			
Jacobson JL 1993	3	2	3			

Jacobson JL 1994	3	2	2			
Jacobson SW	3	2	2			
Kesmodel U 2000	3	2	3			
Kesmodel U 2002	3	2	3			
Lazzaroni F	4	2	3			
Little RE				4	2	1
Long MG				2	0	1
Lumley J	3	2	3			
Lundsberg LS	4	2	2			
Marbury MC	5	2	3			
McDonald AD	1	2	2			
Mills JL 1987	2	2	2			
Mills JL 1984	2	2	3			
Nulman	3	2	2			
O'Callaghan	4	2	3			
Ogston S A	4	2	1			
Olsen J 1994	3	2	2			
Olsen J 1995	4	2	2			
Orskou J	3	2	3			
Parazzini F				1	2	1
Parry GJ	4	2	2			
Passaro KT	3	2	3			
Peacock JL 1991	4	2	3			
Peacock JL 1995	4	2	3			
Plant ML	4	2	2			
Primatesta P	2	2	3			
Raymond EG	3	2	2			
Sampson PD	4	2				
Shiono PH	2	2	3			
Shu XO	3	2	2			
Sood B	3	2	2			
Stoler JM	2	2	2			
Streissguth AP 1983	4	2	3			
Streissguth AP 1990	4	2	3			
Streissguth AP 1994	4	2				
Streissguth AP 1989	4	2	3			
Sulaiman ND	4	2	3			
Tolo KA	3	2	3			
Verkerk PH 1993	4	2	2			
Verkerk PH 1994	3	2	3			
Virji SK 1990	3	2	3			
Virji SK 1991	3	2	3			
Whitehead N	4	2	2			
Windham GC 1992				2	2	2
Windham GC 1994				4	2	1
Windham GC 1997	4	2	3			
Wisborg K	3	2	3			
Yang Q				3	2	2

4 Discussion and Conclusions

4.1 Principal conclusions from the systematic review

The principal findings of this systematic review of the fetal effects of low-to-moderate alcohol consumption in pregnancy were that, for most outcomes, there was no consistent evidence of adverse effect across different studies. Two exceptions to this were possible effects of low-to-moderate alcohol exposure on spontaneous abortion, and binge drinking on outcomes. With neurodevelopmental outcomes the effects, which were generally quite small, included an increase in 'disinhibited behaviour' (Nulman et al, 2004), a reduction in verbal IQ and increase in delinquent behaviour (Bailey et al, 2004) and more learning problems and poorer performance (Streissguth et al, 1989, 1990). The studies which considered these issues were not without problems, including possible overlap between binge drinkers who otherwise drink little and binge drinkers who generally drink substantial amounts. However, they represent the most consistent evidence of a possible effect.

Many of the outcomes, including stillbirth, IUGR, birthweight, appeared to have a 'J-shaped' curve with alcohol exposure. This suggests that babies of women who abstained had poorer outcomes than those who drank small amounts in pregnancy. One possible explanation for this may be the 'healthy drinker effect' (Bell & Lumley, 1989) in which women with a poor obstetric history are more likely to abstain from drinking alcohol. It could also be the case theoretically that there might be a beneficial effect of low-to-moderate drinking in pregnancy but we have not established evidence that would confirm this and hence cannot conclude that this is the case.

This systematic review was carried out with limited resources and within a 6 month period; we therefore designed and carried out a pragmatic search strategy. Searches were limited to English language studies in the four main bibliographic databases Medline, Embase, PsychInfo and Cinahl. We did not attempt to access the 'grey' literature nor did we request further data from authors. Also, we were obliged to use a high sensitivity filter to reduce the number of papers to manageable numbers. Nevertheless, we made the inclusion criteria as broad as possible and scanned 3630 titles which were systematically narrowed down to 66 publications. Very few of the retrieved papers specifically addressed low-to-moderate consumption. Most made comparisons across a number of different levels of consumption.

The systematic review may have been affected by publication bias in which studies with positive results are both more likely to be submitted and more likely to be accepted for publication. Although we concentrated on low-to-moderate consumption, the majority of the studies also included higher levels of consumption where positive findings were more common. However, if the results are affected by publication bias then it would imply that low-to-moderate drinking may be safer than it appears from the published literature.

The quality of the included studies was assessed using the Newcastle-Ottawa Quality Assessment Scales (see appendix 6). This scale has been used in other Cochrane reviews of non-randomised studies such as the use of antiepileptic drugs in pregnancy (Adab et al, 2005). Generally the studies included in this review scored quite highly. However, this was not a true reflection of the quality of many of the studies which had problems specific to carrying out research in the area of prenatal alcohol exposure, and which were not covered in the general quality assessment scale. For example, a common problem related to the timing of the questions about alcohol consumption. Women were frequently asked after delivery how much they drank in pregnancy, when the outcome was already apparent. The potential for recall bias is clear. The better studies used validated questionnaires or interviews administered antenatally to ask about specific time periods both prior to pregnancy recognition and during pregnancy. Some of the studies included only small numbers

of children, particularly those examining longer term outcomes. It is therefore possible that they were underpowered to detect small differences. Other studies carried out multiple testing across a range of outcomes without adjustment, increasing the chance of spuriously significant results.

The majority of the included studies were from the USA. The generalisability of these results to the UK may be questionable. Differences in drinking patterns, for example more or less binge type drinking, the extent to which women under-report drinking in pregnancy and ascertainment of outcomes, particularly neurodevelopmental outcomes may all differ between the USA and UK. Therefore, the findings should be treated with caution.

In attempting to be as broad as possible in our inclusion criteria, studies were included if they had at least two categories of consumption within the 12g per day limit (12g being equivalent to one drink in the USA). The majority of these papers also reported outcomes for higher consumption and therefore, tests of statistical significance tended to be tests for trend across a number of different consumption levels. This made interpretation of differences between low-to-moderate levels of consumption and abstinence problematic. When studies did report outcomes within these low-to-moderate categories of consumption, it was often as a first step in the analysis. At this point they were often unadjusted for potential confounders. Therefore, although appropriate adjustment for potential confounders was made, in many cases this could not be related to the low-to-moderate comparisons.

Although we are not aware of any other systematic reviews of fetal effects of alcohol consumption specific to low-to-moderate consumption, we identified three which used a range of moderate consumption, the upper range of which was much higher than we considered, which we consider may be relevant.

Polygenis et al (1998) conducted a meta-analysis of moderate alcohol consumption during pregnancy and the incidence of fetal malformations. Moderate consumption was defined as the range >2 drinks per week and <2 drinks per day (24 - 168g per week). The meta-analysis included 130,810 pregnancy outcomes and reported a relative risk for fetal malformation of 1.01 (95% CI 0.94 to 1.08). Another meta-analysis examined the association between moderate alcohol consumption and spontaneous abortion, stillbirth and premature birth (Makarechian et al, 1998). Definition of ‘moderate consumption’ was the same as Polygenis et al (1998). Odds ratios for spontaneous abortion were 1.35 (95% CI 1.09 to 1.67), stillbirth 0.65 (0.46 to 0.91), and premature birth 0.95 (0.79 to 1.15). However, the result for stillbirth was considered unstable and inconclusive due to the small number of studies, and significant heterogeneity existed among the individual odds ratios for spontaneous abortion. A further meta-analysis was conducted by Testa et al (2003) investigating the association with infant mental development. Exposure was categorised into less than one drink per day, 1-2 drinks per day, and greater than two drinks per day. Outcome was assessed using the Mental Development Index (MDI) of the Bayley Scales of Infant Development. Children aged 6-8 months, 12-13 months and 18-26 months were considered separately. Alcohol consumption at all levels was associated with significantly lower MDI scores in children aged 12-13 months but not other ages. However, there was, again, considerable heterogeneity and the results were not conclusive. These results are broadly in line with those from this systematic review, allowing for the higher consumption. In contrast with these three studies, we did not attempt to conduct meta-analyses in this review due to the considerable heterogeneity in methods of the various studies (Egger et al, 1998).

This systematic review did not find clear and robust evidence of poor outcome amongst women consuming low-to-moderate amounts of alcohol in pregnancy. Nevertheless, the evidence is probably not strong enough to rule out any risk. However, women questioned about their drinking habits in pregnancy tend to under-report. Therefore, actual drinking patterns will be higher and any associations with adverse outcome will be with higher levels of drinking than those reported. Although most studies did not specify when in pregnancy alcohol consumption related to, or if it

was prior to pregnancy recognition, there is no consistent evidence that low-to-moderate drinking in any particular trimester is associated with poor outcome. However, binge drinking in pregnancy may be cause for concern and may be associated with poor neurodevelopmental outcomes.

This systematic review has also revealed significant gaps and weaknesses in the evidence base. There was only one study of antepartum haemorrhage and only two that included growth in childhood among the outcomes. Most of the studies concentrated on birth weight (20 studies) and/or preterm birth (16 studies).

Future research needs to consider the accuracy and validity of estimates of alcohol consumption as described by Ulrik Kesmodel (see appendix 2). Studies concentrating specifically on either low-to-moderate levels of consumption, or binge drinking in women whose average consumption is low-to-moderate would be of benefit. This would allow for more detailed analysis of this area. The specific effects on childhood neurodevelopmenatal outcomes will require long term follow-up studies.

4.2 Issues arising from expert group meeting (see appendix 2)

In order to guide the systematic review, peer-review the protocol and identify research priorities in the area of prenatal alcohol exposure an advisory group was set up. At a meeting in London on 8th December 2005, members of the advisory group were presented with draft findings from the systematic review as well as a background paper (which is now presented as the background section of this report). Using this material as well as their own expert knowledge, the advisory group was asked to identify and prioritise future research needs in the area of prenatal alcohol exposure in general. Thus they were to consider research needs including but not limited to low-to-moderate alcohol consumption. The key research questions highlighted were:

- What are the effects of low-to-moderate prenatal exposure on IQ, socio-emotional development and behaviour?
- What is the prevalence of alcohol consumption in UK pregnant women?
- Are the risks of fetal alcohol exposure at levels below dysmorphology contingent upon other prenatal risks and/or postnatal risk environment?
- Are the behavioural and cognitive sequelae of overt FAS modifiable? Are treatment implications different from non-FAS?
- What are the reasons for the large differences between the UK and USA in rates of FAS and FASD?
- Are preventive and treatment interventions effective?
- What is the contribution of prenatal alcohol exposure to neurodevelopmental disorders and neurobehavioural functions?
- What is the prevalence of FAS in the UK?

5 Acknowledgments

The National Perinatal Epidemiology Unit is funded by the Department of Health. The views expressed in this report are those of the authors and do not necessary reflect those of the Department of Health.

The authors would like to thank the Advisory Group members for the time and effort they put into ensuring this project ran successfully. We would also like to thank the six anonymous referees for their helpful comments on the draft report.

6 References

Abel EL. (1988) Fetal alcohol syndrome in families. *Neurotoxicology and Teratology* 10,1-2.

Abel EL. (1998) *Fetal Alcohol Abuse Syndrome*. New York: Plenum Press .

Abel EL. (2004) Paternal contribution to fetal alcohol syndrome. *Addiction Biology* 9,127-133.

Achenbach, TM, Edelbrock, C. (1983). *Manual for the child behavior checklist and revised child behavior profile.* Burlington, VT: Queen City Printers.

Adab N, Tudur Smith C, Vinten J, Williamson PR, Winterbottem JJ. (2005) Common antiepileptic drugs in pregnancy in women with epilepsy. http://www.thecochranelibrary.com [accessed Jan 2006].

Albertsen K, Andersen AM, Olsen J, Gronback M. (2004) Alcohol consumption during pregnancy and the risk of preterm delivery. *American Journal of Epidemiology* **159** (2)155-61.

Armstrong B, McDonald A, Sloan M. (1992) Cigarette, alcohol, and coffee consumption and spontaneous abortion. *American Journal of Public Health* **82**: 85-87.

Asherson P, Kuntsi J, Taylor E. (2005) Unravelling the complexity of attention deficit hyperactivity disorder: a behavioural genomic approach. *British Journal of Psychiatry* 187,103-105

Bayley N. (1969). *Bayley Scales of Infant Development-manual*. San Antonio, TX: Psychological Corporation.

Bailey BN, Delaney Black V, Covington CY, Ager J, Janisse J, Hannigan JH and Sokol RJ. (2004) Prenatal exposure to binge drinking and cognitive and behavioral outcomes at age 7 years. *American Journal of Obstetrics and Gynecology* **191** (3)1037-43.

Bearer CF. (2001) Markers to detect drinking during pregnancy*. Alcohol Research and Health.* 3,210-218.

Beattie JO, Day RE, Cockburn F, Garg RA. (1983) Alcohol and the fetus in the west of Scotland. *BMJ* 287,17-20.

Bell R, Lumley J (1989) Alcohol consumption, cigarette smoking and fetal outcome in Victoria, 1985. *Community Health Studies* **13**(4)484-91.

Berkowitz G, Holford T, Berkowitz R. (1982) Effects of cigarette smoking, alcohol, coffee and tea consumption on preterm delivery. *Early Human Development* **7**, 239-50.

Bertrand J, Floyd RL, Weber MK, O'Connor M et al. (2004) *National Task Force on FAS/FAE: Guidelines for referral and diagnosis.* Atlanta: CDC.

Bielawski DM, Zaher FM, Svinarich DM, Abel EL. (2002) Paternal alcohol exposure affects sperm cytosine methyltransferase messenger RNA levels. *Alcoholism: Clinical & Experimental Research* 26(3),347-351.

Brazelton TB, Nugent JK. (1995). Neonatal Behavioral Assessment Scale. 3rd ed. *Clinics in Developmental Medicine, No. 137*. London: MacKeith Press.

Brooke OG, Anderson HR, Bland JM, Peacock JL Stewart CM. (1989) Effects on birth weight of smoking, alcohol, caffeine, socioeconomic factors, and psychosocial stress. *BMJ* **298** (6676):795-801.

Burd L, Klug MG, Mertsolf JT, Kerbeshian J. (2003) Fetal alcohol syndrome: neuropsychiatric phenomics. *Neurotoxicology and Teratology* 25,697-705.

Centers for Disease Control and Prevention (CDC) (2004) Alcohol use among women who are pregnant or who might become pregnant - United States 2002 *Morbidity and Mortality Weekly Reports* 53(50), 1178-1181.

Chan AK, Pristach EA, Welte JW, Russell M. (1993) The TWEAK test in screening for alcoholism/ heavy drinking in three populations. *Alcoholism: Clinical and Experimental Research* 6,1188-1192.

Chang G, Goetz MA, Wilkins-Haug L, Berman S. (2000) A brief intervention for prenatal alcohol use: an in depth look. *Journal of Substance Abuse Treatment* 18,365-369.

Chang G. (2005) Screening and brief intervention in prenatal care settings. *Alcohol Research and Health* 28(2),80-84.

Chen WA, Maier SE, Parnell SE, West JR. (2003) Alcohol and the developing brain: neuroanatomical studies. *Alcohol Research and Health* 27(2),174-180.

Clarren SK, Randels SP, Sanderson M, Fineman RM. (2001) Screening for fetal alcohol syndrome in primary schools: a feasibility study. *Teratology* 63,3-10.

Chakrabarti S, Fombonne E. (2001) Pervasive developmental disorders in preschool children. *JAMA* 24,3093-3099.

Conners CK. (2001) *Conners' Rating Scales—Revised*. New York, Multi-Health Systems

Davis P, Partridge J, Storrs C. (1982) Alcohol consumption in pregnancy: how much is safe? *Archives of Disease in Childhood* **57**, 940-43.

Dawson DA. (2003) Methodological issues in Measuring Alcohol Use. *Alcohol Research and Health* 27(1),18-29.

Day NL, Leech SL, Richardson GA, Cornelius MD, Robles N, Larkby C. (2002) Prenatal alcohol exposure predicts continued deficits in offspring size at 14 years of age. *Alcohol: Clinical and Experimental Research* **26** (10) 1584-91.

Day N, Richardson G, Robles N, Sambamoorthi U, Taylor P, Scher M, Stoffer D, Jasperse D, Cornelius M. (1990) Effect of prenatal alcohol exposure on growth and morphology of offspring at 8 months of age. *Pediatrics* **85**, 748-52.

Day N, Zuo Y, Richardson G, Goldschmitt L, Larkby C and Cornelius M. (1999) Prenatal alcohol use and offspring size at 10 years of age. *Alcohol: Clinical and Experimental Research* **23**, 863-69.

Department of Health (2005) *The Pregnancy Book*. London: The Stationery Office.

Dex S, Joshi H (eds.) (2005) *Children of the 21st century: From birth to nine months.* Bristol: Policy Press.

Dufour MC. (1999) What is moderate drinking? Defining "drinks" and drinking levels. *Alcohol Research and Health* 23(1),5-14.

Dunn L, Markwardt F. (1989) *Peabody Individual Achievement Test.* Circle Pines, Minn, American Guidance Services.

Egger M, Schneider M, Davey Smith G. (1998) Meta-analysis - Spurious precision? Meta-analysis of observational studies. *BMJ* **316**, 140-44.

Ernhart C, Sokol R, Ager J, Morrow-Tlucak M, Martier S. (1989) Alcohol-related birth defects: assessing the risk. *Annals of the NewYork Academy of Sciences* **562**, 159-72.

Faden VB, Graubard BI, Dufour M. (1997) The relationship of drinking and birth outcome in a US national sample of expectant mothers. *Paediatric and Perinatal Epidemiology* **11** (2) 167-80.

Farber NB, Olney JW. (2003) Drugs of abuse that cause developing neurons to commit suicide. *Developmental Brain Research* 147,37-45 .

Florey C DuV. (1992) Chapter 6. Methods. *International Journal of Epidemiology* 21(4 Suppl.1),S38-39.

Forrest F, Florey CD, Taylor D, McPherson F, Young JA. (1991) Reported social alcohol consumption during pregnancy and infants' development at 18 months. *BMJ* **303** (6793) 22-6.

Gill JS, Donaghy M. (2004) Variation in the alcohol content of a 'drink' of wine and spirit poured by a sample of the Scottish population. *Health Education Research* 19(5),485-91.

Goldschmidt L, Richardson GA, Cornelius MD, Day NL. (2004) Prenatal marijuana and alcohol exposure and academic achievement at age 10. *Neurotoxicology and Teratology* **26** (4):521-32.

Goodlett CR, Horn KH. (2001) Mechanisms of alcohol - induced damage to the developing nervous system. *Alcohol Research and Health* 25(3),175-184.

Goodlett CR, Horn KH, Zhou FC. (2005) Alcohol teratogenesis: mechanisms of damage and strategies for intervention. *Experimental Biology and Medicine* 230(6),394-406.

Griffiths R. (1984) *The Abilities of Young Children. A Comprehensive System of Mental Measurement for the First Eight Years of Life*. London: The Test Agency Ltd.

Gunzerath L, Faden V, Zakhari S, Warren K. (2004) National Institute on Alcohol Abuse and Alcoholism report on moderate drinking. *Alcoholism: Clinical and Experimental Research* 28(6), 829-847.

Halliday HL, Reid MM, McLure G. (1982) Results of heavy drinking in pregnancy. *British Journal of Obstetrics and Gynaecology* 89,892-895.

Handmaker NS, Wilbourne P. (2001) Motivational interventions in prenatal clinics. *Alcohol Research and Health* 3,219-229.

Hankin JR. (2002) Fetal alcohol syndrome prevention research. *Alcohol Research and Health* 6(1),58-65.

Harlap S, Shiono P. (1980) Alcohol, smoking, and incidence of spontaneous abortions in the first and second trimester. *Lancet* **ii**, 173-5.

Henriksen T, Hjollund N, Jensen T, Bonde J *et al* (2004) Alcohol consumption at the time of conception and spontaneous abortion. *American Journal of Epidemiology* **160**, 661-67.

Hepper PG, Dornan JC, Little JF. (2005) Maternal alcohol consumption during pregnancy may delay the onset of spontaneous fetal startle behaviour. *Physiology and Behavior* 83,711-714.

Holliday R. (1998) The possibility of epigenetic transmission of defects induced by teratogens. *Mutation Research* 422(2), 203-205.

Hoyme HE, May PA, Kalberg WO, Kodituwakku P, Gossage JP, Trujillo PM, Buckley DG, Miller JH, Aragon AS, Khaole N, Viljoen DL, Jones KL, Robinson LK. (2005) A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 Institute of Medicine criteria. Pediatrics 115(1),39-47.

Huizink AC, Mulder EJH. (2006) Maternal smoking, drinking or cannabis use during pregnancy and neurobehavioral and cognitive functioning in human offspring.
Neurosci Biobehav Rev 30(1),24-41.

Ikonomidou C, Bittigau P, Ishimaru M, Wozniak DF et al. (2000) Ethanol-induced apoptotic neurodegeneration and fetal alcohol syndrome. *Science*. 287,1056-1060

International Center for Alcohol Policies. (1998) What is a “Standard Drink”? *ICAP Report No. 5* (Available online at www.icap.org).

Jacobson JL, Jacobson SW. (1999) Drinking moderately and pregnancy: Effects on child development. *Alcohol Research & Health* 23(1),25-30.

Jacobson JL, Jacobson SW, Sokol RJ. (1994) Effects of prenatal exposure to alcohol, smoking, and illicit drugs on postpartum somatic growth. *Alcohol: Clinical and Experimental Research* **18** (2):317-23.

Jacobson JL, Jacobson SW, Sokol RJ, Martier SS, Ager JW, Kaplan Estrin MG. (1993) Teratogenic effects of alcohol on infant development. *Alcohol: Clinical and Experimental Research* **17** (1) 174-83.

Jacobson J, Jacobson S. (1999) Drinking moderately and pregnancy: Effects on child development. *Alcohol Research and Health* **23**, 25-30.

Jastak JF, Jastak SR, Wilkinson GS.(1993) *The Wide-Range Achievement Test, Revised 2*. Wilmington, Guidance Associates of Delaware.

Jones KL, Smith DW. (1973) Recognition of the fetal alcohol Syndrome in early infancy. *Lancet* 2,999-1001.

Kalter H. (2003) Teratology in the 20th century: Environmental causes of congenital malformations in humans and how they were established. *Neurotoxicology and Teratology* **25**:131-282.

Kesmodel U, Wisborg K, Olsen SF, Henriksen TB, Secher NJ. (1987) Moderate alcohol intake in pregnancy and the risk of spontaneous abortion. *Alcohol and Alcoholism* **37** (1) 87-92.

Kesmodel U, Olsen S and Secher N. (2000) Does alcohol increase the risk of preterm delivery? *Epidemiology* **11**, 512-18.

Kesmodel U. (2001) Binge drinking in pregnancy-frequency and methodology. *Am J Epidemiol* 154(8),777-82.

Kesmodel U, Olsen SF. (2001) Self-reported alcohol intake in pregnancy: comparison between four methods. *J Epidemiol Community Health* 55,738-745.

Kesmodel U, Wisborg K, Olsen SF, Henriksen TB, Secher NJ. (2002) Moderate alcohol intake during pregnancy and the risk of stillbirth and death in the first year of life. *American Journal of Epidemiology* **155** (4):305-12.

Kesmodel U, Frydenberg M. (2004) Binge drinking during pregnancy--is it possible to obtain valid information on a weekly basis? *Am J Epidemiol* 159(8),803-8.

Knopik VS, Sparrow EP, Madden PAF, Bucholz KK et al. (2005) Contributions of parental alcoholism, prenatal substance exposure, and genetic transmission to child ADHD risk: a female twin study. *Psychological Medicine* 35,625-635.

Lazzaroni F, Bonassi S, Magnani M, Calvi A, et al. (1993) Moderate maternal drinking and outcome of pregnancy. *European Journal of Epidemiology* **9** (6) 599-606.

Lemmens PH, Tan ES, Knibbe RA. (1992) Measuring quantity and frequency of drinking in a general population survey: A comparison of five indices. *Journal of Studies on Alcohol* 53(5),476-486.

Lemon R, Dunnett SB. (2005) Surveying the literature from animal experiments. *BMJ* 330,977-978.

Little R, Weinberg C. (1993) Risk factors for antepartum and intrapartum stillbirth. *American Journal of Epidemiology* **137**, 1177-89.

Little JF, Hepper PG, Dornan JC. (2002) Maternal alcohol consumption during pregnancy and fetal startle behaviour. *Physiology and Behavior* 76, 691-694.

Long MG, Waterson EJ, MacRae KD, Murray Lyon IM. (1994) Alcohol consumption and the risk of first trimester miscarriage. *Journal of Obstetrics and Gynaecology* **14** (2) 69-70.

Lumley J, Correy J, Newman N, Curran J. (1985) Cigarette smoking, alcohol consumption and fetal outcome in Tasmania 1981-82. *Australia and New Zealand Journal of Obstetrics and Gynaecology* **25**, 33-40.

Lundsberg LS, Bracken MB, Saftlas AF. (1997) Low-to-moderate gestational alcohol use and intrauterine growth retardation, low birthweight, and preterm delivery. *Annals of Epidemiology* **7** (7) 498-5.

McCarthy P. (1972) *McCarthy scales of children's abilities*. New York: Psychological Corporation.

McDonald AD, Armstrong BG, Sloan M. (1992) Cigarette, alcohol, and coffee consumption and prematurity. *American Journal of Public Health* **82** (1) 87-90.

Macleod J, Davey Smith G, Heslop P, Metcalfe C et al. (2001) Are the effects of psychosocial exposures attributable to confounding? Evidence from a prospective observational study on psychological stress and mortality. *J Epidemiol Community Health* 55,878-884.

Makarechian N, Agro K, Devlin J, Trepanier E et al. (1998) Association between moderate alcohol consumption during pregnancy and spontaneous abortion, stillbirth and premature birth: A meta-analysis. *Canadian Journal of Clinical Pharmacology* 5(3), 169-176.

Marbury MC, Linn S, Monson R et al (1983) The association of alcohol consumption with outcome of pregnancy. *American Journal of Public Health* **73** (10) 1165-8.

May PA, Hymbaugh KJ, Aase JM, Samet JM. (2000) Epidemiology of fetal alcohol syndrome in a South African Community in the Western Cape Province. *American Journal of Public Health* 90(12),1905-1912.

May PA, Gossage JP. (2001) Estimating the prevalence of fetal alcohol syndrome. A summary. *Alcohol Research and Health* 25(3),159-167.

Mills JL, Graubard BI. (1987) Is moderate drinking during pregnancy associated with an increased risk for malformations? *Pediatrics* **80** (3) 309-14.

Mills J, Graubard B, Harley E, Rhoads G, Berendes H. (1984) Maternal alcohol consumption and birth weight. *JAMA* **252**, 1875-79.

Nulman I, Rovet J, Kennedy D, Wasson C, Gladstone J, Fried S, Koren G. (2004) Binge alcohol consumption by non-alcohol-dependent women during pregnancy affects child behaviour, but not general intellectual functioning; a prospective controlled study. *Archives of Women's Mental Health* **7** (3) 173-81.

O'Callaghan FV, O'Callaghan M, Najman JM, Williams GM, Bor W. (2003) Maternal alcohol consumption during pregnancy and physical outcomes up to 5 years of age: a longitudinal study. *Early Human Development* **71** (2) 137-48.

O'Connor MJ, Whaley SE (2003). Alcohol use in low income pregnant women.*J Studies on Alcohol* 64,773-783.

O'Connor TG, Heron J, Glover V; Alspac Study Team. (2002) Antenatal anxiety predicts child behavioral/emotional problems independently of postnatal depression. *J Am Acad Child Adolesc Psychiatry* 41(12),1470-1477.

Ogston S, Parry G. (1992) Results - Strategy of analysis and analysis of pregnancy outcome . *International Journal of Epidemiology* **21**, S45-S71.

Olsen J. (1994) Effects of moderate alcohol consumption during pregnancy on child development at 18 and 42 months. *Alcohol: Clinical and Experimental Research* **18** (5) 1109-13.

Olsen J, Tuntiseranee P. (1995) Is moderate alcohol intake in pregnancy associated with the craniofacial features related to the fetal alcohol syndrome? *Scandinavian Journal of Social Medicine* **23** (3):156-61.

Ørskou J, Henriksen TB, Kesmodel U, Secher NJ. (2003) Maternal characteristics and lifestyle factors and the risk of delivering high birth weight infants. *Obstetrics and Gynecology* **102** (1):115-20.

Parazzini F, Chatenoud L, Surace M, Tozzi L, Salerio B, Bettoni G, Benzi G. (2003) Moderate alcohol drinking and risk of preterm birth. *European Journal of Clinical Nutrition* **57** (10) 1345-9.

Parry G, Ogston S. (1992) Results - Child development at age 18 months. *International Journal of Epidemiology* **21**, S72-S85

Passaro KT, Little RE, Savitz DA, Noss J. (1996) The effect of maternal drinking before conception and in early pregnancy on infant birthweight. The ALSPAC Study Team. Avon Longitudinal Study of Pregnancy and Childhood. *Epidemiology* **7** (4) 377-83.

Peacock JL, Bland JM, Anderson HR. (1995) Preterm delivery: Effects of socioeconomic factors, psychological stress, smoking, alcohol, and caffeine. *BMJ* **311** (7004):531-36.

Peng Y, Yang P, Ng SSM, Wong OG et al. (2004) A critical role of Pax6 in alcohol-induced fetal microcephaly. *Neurobiology of Disease* 16(2),370-376.

Plant ML. (1984) Drinking amongst pregnant women: some initial results from a prospective study. *Alcohol and Alcoholism.* 19,153-157.

Plant ML. (1985) *Women, Drinking and Pregnancy*. New York: Tavistock Publications.

Plant M, Plant M. (1988) Maternal use of alcohol and other drugs during pregnancy and birth abnormalities: further results from a prospective study. *Alcohol and Alcoholism* **23**, 229-33.

Polygenis D, Wharton S, Malmberg C, Sherman N et al. (1998) Moderate alcohol consumption during pregnancy and the incidence of fetal malformations: a meta-analysis. *Neurotoxicology and Teratology* 21, 61-67.

Primatesta P, Del Corno G, Bonazzi MC, Waters WE. (1993) Alcohol and pregnancy: an international comparison. *Journal of Public Health Medicine* 15,69-76

Primatesta P, Del Corno G, Bonazzi MC (1994) Alcohol consumption as a fetal risk factor in a group of pregnant women of Northern Italy. *Journal of Preventitive Medical Hygiene* **35** (3-4):89-94.

Raymond E, Mills J. (1993) Placental abruption. Maternal risk factors and associated fetal conditions. *Acta Obstetrica Gynecolica Scandanavica* **72**, 633-39.

Rickards L, Fox K, Roberts C, Fletcher L, Goddard E (eds.) (2004) *Living in Britain: Results from the General Household Survey, no 31*. London: The Stationery Office.

Riley EP, Mattson SN, Lai T, Jacobson SW et al. (2003) Neurobehavioral consequences of prenatal alcohol exposure: an international perspective. *Alcoholism: Clinical and Experimental Research* 27(2),362-373.

Riley EP. (2004) Commentary on 'Paternal contribution to fetal alcohol syndrome' by EL Abel. *Addition Biology* 9,136-136.

Riley EP, McGee CL. (2005) Fetal alcohol spectrum disorders: an overview with emphasis on changes in brain and behaviour. *Experimental Biology and Medicine* 230(6),357-365.

Room R, Babor T, Rehm J. (2005) Alcohol and public health. *Lancet* 365,519-530.

Rutter M, Pickles A, Murray R, Eaves L. (2001) Testing hypotheses on specific environmental causal effects on behaviour. *Psychological Bulletin.* 127(3), 291-324.

Rutter M. (2005) Environmentally mediated risks for psychopathology: research strategies and findings. *Journal of the American Academy of Child and Adolescent Psychiatry* 44(1),3-27.

Sampson P, Bookstein F, Barr H, Streissguth A. (1994) Prenatal alcohol exposure, birthweight, and measures of child size from birth to age 14 years. *American Journal of Public Health* **84**, 1421-28.

Sampson PD, Streissguth AP, Bookstein FL, Little RE et al. (1997) Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. *Teratology* 56,317-326.

Schneider ML, Roughton EC, Lubach GR. (1997) Moderate alcohol consumption and psychological stress during pregnancy induce attention and neuromotor impairments in primate infants. *Child Development* 68(5), 747-759.

Shiono PH, Klebanoff MA. (1986) Ethnic differences in preterm and very preterm delivery. *American Journal of Public Health* **76** (11):1317-21.

Shu XO, Hatch MC, Mills J, Clemens J, Susser M. (1995) Maternal smoking, alcohol drinking, caffeine consumption, and fetal growth: results from a prospective study. *Epidemiology* **6** (2) 115-20.

Sokol RJ, Martier SS, Ager JW. (1989) The T-ACE questions: practical prenatal detection of risk-drinking. *American Journal of Obstetrics and Gynecology* 160,863-870.

Sokol RJ, Delaney-Black V, Nordstrom B. (2003) Fetal alcohol spectrum disorder. *JAMA.* 290(22),2996-2999.

Sood B, Delaney-Black V, Covington C, Nordstrom-Klee B et al. (2001) Prenatal alcohol exposure and childhood behavior at age 6 to 7 years: I. Dose-response effect. *Pediatrics* 108(2),1-9.

Sproston K, Primatesta P (eds.) (2005) *Health Survey for England 2003*. London: The Stationery Office.

Stockwell T, Donath S, Cooper-Stanbury M, Chikritzhs T, Catalano P, Mateo C. (2004) Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall. *Addiction* 99, 1024-1033.

Stoler JM, Ryan LM and Holmes LB. (2002;) Alcohol dehydrogenase 2 genotypes, maternal alcohol use, and infant outcome. *J Pediatr* **141** (6):780-5.

Stratton KR, Howe CJ, Battaglia FC. (1996) *Fetal alcohol syndrome: diagnosis, epidemiology, prevention and treatment.* Washington DC: National Academy Press.

Streissguth AP, Martin DC, Martin JC, Barr HM. (1981) The Seattle longitudinal prospective study on alcohol and pregnancy. *Neurobehav Toxicol Teratol* 3,223-233.

Streissguth A, Barr H, Martin D. (1983) Maternal alcohol use and neonatal habituation assessed with the Brazelton Scale. *Child Development* **54**, 1109-18.

Streissguth AP, Barr HM, Sampson PD, Bookstein FL, Darby BL. (1989) Neurobehavioral effects of prenatal alcohol: Part I. Research strategy. *Neurotoxicology and Teratology* **11** (5) 461-76.

Streissguth AP, Barr HM, Sampson PD. (1990) Moderate prenatal alcohol exposure: effects on child IQ and learning problems at age 7 1/2 years. *Alcohol: Clinical and Experimental Research* **14** (5) 662-9.

Streissguth AP, Aase JM, Clarren SK, Randels SP et al. (1991) Fetal alcohol syndrome in adolescents and adults. *JAMA* 265(15),1961-1967.

Streissguth AP, Bookstein FL, Sampson PD, Barr HM. (1993) *The enduring effects of prenatal alcohol exposure on child development: Birth through seven years a partial least squares solution.* Michigan: University of Michigan Press.

Streissguth AP, Dehaene P. (1993) Fetal alcohol syndrome in twins of alcoholic mothers: concordance of diagnosis and IQ. *American Journal of Medical Genetics* 47,857-861.

Sulaiman ND, Florey CD, Taylor DJ, Ogston SA. (1988) Alcohol consumption in Dundee primigravidas and its effects on outcome of pregnancy. *BMJ* 296,1500-1503.

Sulik KK. (2005), Johnston MC, Webb MA. (1981) Fetal alcohol syndrome: embryogenesis in a mouse model. *Science* 214,936-938.

Sulik KK Genesis of alcohol-induced craniofacial dysmorphism. *Experimental biology and Medicine* 230(6),366-375.

Surgeon General (2005) Surgeon General's Advisory on Alcohol Use in Pregnancy. Available at http://www.hhs.gov/surgeongeneral/pressreleases/sg02222005.html.

Testa M, Quigley B, Das Eiden R. (2003) The effects of prenatal alcohol exposure on infant mental development: a meta-analytical review. *Alcohol and Alcoholism* **38**, 295-304.

Thorndike RL, Hagen EP, Sattler JM. (1986) *Stanford-Binet Intelligence Scales. 4th ed*. Itasca, IL: Riverside Publishing Company.

Tolo KA, Little RE. (1993) Occasional binges by moderate drinkers: implications for birth outcomes. *Epidemiology* **4** (5) 415-20.

Verkerk PH. (1992) The impact of alcohol misclassification on the relationship between alcohol and pregnancy outcome. *International Journal of Epidemiology* 21(4 Suppl.1),S33-37.

Verkerk PH, Van Noord Zaadstra BM, Du Florey VC, De Jonge GA, Verloove Vanhorick SP. (1993) The effect of moderate maternal alcohol consumption on birth weight and gestational age in a low risk population. *Early Human Development* **32** (2-3) 121-9.

Verkerk PH, Buitendijk SE, Verloove Vanhorick SP. (1994) Differential misclassification of alcohol and cigarette consumption by pregnancy outcome. *Int J Epidemiology* **23** (6) 1218-25.

Virji SK, Talbott EO. (1990) The relationship between occupational classification and low birth weight in a national sample of white married mothers. *International Archives of Occupational and Environmental Health* **62** (5):351-6.

Virji SK. (1991) The relationship between alcohol consumption during pregnancy and infant birthweight. An epidemiologic study. *Acta Obstetrica Gynecologica Scandinavia* **70** (4-5) 303-8.

Warren KR, Ting-Kai Li. (2005) Genetic polymorphisms: impact on the risk of fetal alcohol spectrum disorders. *Birth Defects Research (Part A)* 73,195-203.

Waterson EJ, Murray-Lyon IM. (1989) Drinking and smoking patterns amongst women attending an antenatal clinic: II Drinking in pregnancy. *Alcohol and Alcoholism* 24,163-173.

Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. (2005) The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm [accessed Jan 2006].

Wechsler D. (1974) *The Wechsler Intelligence Scale for Children - Revised*. New York, NY: The Psychological Corporation.

Whitehead N, Lipscomb L. (2003) Patterns of alcohol use before and during pregnancy and the risk of small-for-gestational-age birth. *Am J Epidemiol* **158** (7) 654-62.

Wilczynski N, Haynes R, Hedges Team. (2003) Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. *AMIA Annual Symposium Proceedings* 713-23.

Windham GC, Fenster L, Swan SH. (1992) Moderate maternal and paternal alcohol consumption and the risk of spontaneous abortion. *Epidemiology* **3** (4) 364-70.

Windham GC, Fenster L, Hopkins B, Swan SH. (1995) The association of moderate maternal and paternal alcohol consumption with birthweight and gestational age. *Epidemiology* **6** (6) 591-7.

Windham GC, Von Behren J, Fenster L, Schaefer C, Swan SH. (1997) Moderate maternal alcohol consumption and risk of spontaneous abortion. *Epidemiology* **8** (5) 509-14.

Wisborg K, Henriksen TB, Hedegaard M, Secher NJ. (1996) Smoking during pregnancy and preterm birth. *British Journal of Obstetrics and Gynaecology* **103** (8):800-5.

Wright JT, Waterson EJ, Barrison IG, Toplis PJ et al (1983) Alcohol consumption, pregnancy and low birth weight. *Lancet* 1,663-665.

Yang Q, Witkiewicz BB, Olney RS, Liu Y et al. (2001) A case-control study of maternal alcohol consumption and intrauterine growth retardation. *Annals of Epidemiology* **11** (7) 497-503**.**

Part c

Appendix 1: Members of advisory group on fetal effects of alcohol

Dr Margaret Barrow
Consultant Clinical Geneticist
Clinical Genetics Dept
Leicester Royal Infirmary NHS Trust
Leicester, LE1 5WW

Professor Nigel Brown
Director
Centre for Developmental and Endocrine Signaling
Basic Medical Sciences
St George's University of London
Cranmer Terrace
London SW17 0RE
United Kingdom

Dr Frances M Cowan
Senior Lecturer in Neonatal Neurology
Department of Paediatrics and Neonatal Medicine
Imperial College School of Medicine
Hammersmith Hospital
London W12 ONN
United Kingdom

Professor Colin Drummond
Professor of Addiction Psychiatry
Dept of Psychiatry of Addictive Behaviour
Level 6, Hunter Wing
St George's Hospital Medical School
Cranmer Terrace
London SW17 0RE
United Kingdom

Susan Fleisher
The National Organisation for Fetal Alcohol Syndrome - UK
PO Box 33397
London NW11 6WQ
United Kingdom

Mr Robert B Fraser
Reader in Reproductive and Developmental Medicine
Level 4,The Jessop Wing
Tree Root Walk
Sheffield S10 2SF
United Kingdom

Dr Ulrik Kesmodel
University of Aarhus
Institute of Public Health
Department of Epidemiology
Vennelyst Boulevard 6, 8000 Aarhus C, Denmark

Dr Patricia McElhatton
Director
National Teratology Information Service
c/o the Regional Drug and Therapeutic Centre
Claremont Place
Newcastle upon Tyne NE2 4HH
United Kingdom

Ms Laura Oakley
Research Networker
The National Childbirth Trust
Alexandra House
Oldham Terrace
Acton
London W3 6NH
United Kingdom

Professor Moira Plant
Director of the Alcohol and Health Research Trust
University of the West of England
Glenside Campus
Blackberry Hill
Stapleton
Bristol BS16 1DD
United Kingdom

Dr Mark Prunty
Senior Medical Officer
Substance Misuse Team
Department of Health
Room 580D, Skipton House
80 London Road
London SE1 6LH
United Kingdom

Dr Judith Rankin
Principal Research Associate
School of Population and Health Sciences,
University of Newcastle upon Tyne
William Leech Building
Newcastle upon Tyne NE2 4HH
United Kingdom

Professor Sir Michael Rutter
Professor of Developmental Psychopathology
Institute of Psychiatry
Box P080
De Crespigny Park
London SE5 8AF
United Kingdom

Professor Marianne Thoresen
Professor of Neonatal Neurology
Clinical Science South Bristol
St Michaels Hospital
Southwell Street
Bristol BS2 8EG
United Kingdom

Professor Dieter Wolke
Scientific Director
Jacobs Foundation
Seefeldquai 17
PO Box
CH-8034 Zurich
Switzerland

NPEU members

Ron Gray
Consultant Clinical Epidemiologist

Jane Henderson
Health Services Researcher

Peter Brocklehurst
Consultant Clinical Epidemiologist/Director of NPEU

Maggie Redshaw
Social Scientist

Jenny Kurinczuk
Consultant Clinical Epidemiologist

Rona McCandlish
Trials Development Research Fellow

Tricia Boyd
Senior Clinical Research Fellow

Kirstie McKenzie-McHarg
Clinical Research Psychologist

Maria Quigley
Statistical Epidemiologist

Jessica Gibson
Specialist Registrar in Psychiatry

Appendix 2: Report of meeting of advisory group – 8th Dec 2005

Summary of meeting

The meeting of the expert advisory group was to present the preliminary results of the systematic review and to set research priorities for the field of prenatal alcohol exposure.

Short presentations

Short presentations covered the following areas:

Measuring alcohol consumption by maternal self report - Dr Ulrik Kesmodel

Ulrik Kesmodel described the various ways in which women could be asked about their alcohol consumption. Biological measures are not useful for measuring low levels of consumption and indirect questions, whilst potentially helpful in a screening context, do not quantify consumption.

To accurately quantify consumption one needs to know the frequency and quantity consumed, the variability in this, and container size. Asking about timing of drinking and location may also help. Reported frequency and quantity can be ascertained from interviews, questionnaires and diaries. If one assumes that the higher the intake admitted to, the better the method, then interviews and diaries appear to be better. Questionnaires sometimes have high item non-response whereas, in the experience of Ulrik Kesmodel, diaries have high overall response rates. Anonymity improves response rates in questionnaires but precludes making associations with outcome. More detailed questioning about type of beverage consumed appears useful as does asking about average consumption in pregnancy rather than asking about a specific time period such as last week or last fortnight.

Regarding binge drinking, in educated populations this seems to be well remembered both in pregnancy and prior to pregnancy recognition.

Generally, drinking behaviour varies according to location but if too many locations are specified, estimates may overlap leading to over-estimates of consumption. Container size is an important and neglected factor. Glass sizes are very variable and, when this is taken into account, estimates of consumption can more than double.

Diagnosis of fetal alcohol syndrome – Dr Margaret Barrow

Fetal alcohol syndrome (FAS) is difficult to diagnose unless all of the clinical features are present. There are questions around who should make the diagnosis depending on how and when problems present. Characteristics of FAS are well documented including abnormalities of the face, pre- and postnatal growth restriction and neurodevelopmental delay. Scoring systems can be used but they are lengthy, time consuming and costly. The four key categories are growth abnormalities, facial dysmorphism, structural and functional abnormalities of the brain and confirmed alcohol use in pregnancy. There are some objective measurements related to, e.g.growth and palpebral fissure length, which can aid diagnosis. This is sometimes clear cut but often difficult and controversial. No specific neurodevelopmental phenotype reflecting alcohol exposure has yet been definitely identified. Affected children often end up in care so it can be difficult to get a history of maternal alcohol consumption. There are up to five diagnostic categories within fetal alcohol spectrum disorders.

In North America diagnosis is time consuming and complex and varies according to research and service availability. In the UK these resources are not available so diagnosis of fetal alcohol spectrum disorders (FASD) is patchy. The differential diagnosis includes genetic disorders such as Williams syndrome, where children may look similar to those with FAS.

Neuroimaging studies – Dr Frances Cowan

The clinical problems of fetal alcohol syndrome are, in some ways, comparable to those of prematurity. Autopsies of children with FAS are not representative because the babies who have died may be most severely affected . It is thought that alcohol and prematurity both affect the brain by damaging the myelination of the nerves. Brain effects include decreased cerebral volume, altered perisylvian white and grey matter, corpus callosum thinning and white matter atrophy; however, the hippocampus is spared in FAS for unknown reasons. Functional Magnetic Resonance Imaging shows different patterns between children and adults born preterm which may be due to neuroplasticity and consequent adult compensation. Even if it is not possible to see any effect on the nervous system, it is possible to measure deficits in function. The fetal brain can be imaged from 23 weeks onwards and techniques are improving. It may be possible to develop a 'normal standard template' and then assess how the brains of babies affected by FAS are different from this. The way forward could be to have early matching of cases and normal controls with a view to analysing structural and functional differences where possible by using repeated scans over time. It may be possible to identify a brain phenotype for FAS in the early neonatal period.

Study designs for prenatal alcohol exposure – Professor Michael Rutter

Professor Rutter outlined six key research challenges in this area. They included the wide range of overlapping risks, variations in genetic susceptibility, postnatal environment, smoking and gender. He then suggested some potential research designs which could pull apart the different risks. These might include comparing outcomes in children whose mothers and fathers had different levels of alcohol consumption, adoption studies, siblings from pregnancies which were discordant for prenatal alcohol exposure, and comparing outcomes in relation to indicators of alcohol exposure such as dysmorphic features or head circumference.

Treatment of alcohol problems in women – Professor Moira Plant

Professor Plant suggested that the UK pattern of drinking is a major risk factor. The term binge drinking as a new way of drinking in the UK is misleading. This is a typical Northern European pattern of drinking. What is changing is the increased amount drunk on each occasion, particularly amongst young women. Barriers to treatment of problem alcohol use can be both internal, such as shame, and external, such as a lack of awareness amongst health professionals – women are not generally perceived as 'drinkers'. Problem drinking women have an increased likelihood of having experienced childhood trauma and increased chance of co-morbid psychiatric illness. Where women do have a history of childhood trauma, this must be treated at the same time as treating the substance abuse problem. Moira Plant stressed the importance of acceptance of the women and the need for special training of staff. Other factors of importance in treatment of alcohol problems in women include early intervention, case-management services, continuum of care, and long term support.

Discussion and small group work

Preliminary results of the systematic review were presented showing no consitent effect of low-to-moderate alcohol consumption on postnatal growth and a possible effect of binge drinking on neurodevelopmental outcomes.

During subsequent small group work and discussion, the advisory group concluded that there is considerable scope for further research in this area. The difficulties in accurately estimating maternal alcohol consumption, of diagnosing less florid forms of FAS, and treating both alcoholic mothers and their affected children all present challenges for research and practice. In particular, the difficulty of disentangling the separate effects of low-to-moderate alcohol consumption from the 'risk environment' needs to be addressed.

Research questions prioritised by the advisory group included the following:

- What are the effects of low-to-moderate prenatal exposure on IQ, socio-emotional development and behaviour?
- What is the prevalence of alcohol consumption in UK pregnant women?
- Are the risks of fetal alcohol exposure at levels below dysmorphology contingent upon other prenatal risks and/or postnatal risk environment?
- Are the behavioural and cognitive sequelae of overt FAS modifiable? Are treatment implications different from non-FAS?
- What are the reasons for the large differences between the UK and USA in rates of FAS and FASD?
- Are preventive and treatment interventions effective?
- What is the contribution of prenatal alcohol exposure to neurodevelopmental disorders and neurobehavioural functions?
- What is the prevalence of FAS in the UK?

Appendix 3:
UK and USA standard measurement and quantity of alcohol

	Quantity of alcohol in one measure		
Measure	*Grams of alcohol*	*ml of alcohol*	*fl oz of alcohol*
UK one unit	8	10	0.3
USA one drink	12	15	0.5

When converting data from ‘drinks’ or ‘units’ to grams, the following general rule was used in this report:

None or abstainers = 0g
1-2 drinks = 1-24g
3-4 drinks = 25-48g
etc

Appendix 4: Neurodevelopmental outcomes

Neonatal Neurobehavioural functioning

Three sets of behaviours:-

- Reflexes (primitive) - sucking, head turning
- Motor tone
- Orienting behaviour - response to auditory and visual stimuli

Measured either singly or in combination e.g. Brazelton (NBAS)

- Psychophysiological measures e.g. EEG
- Neuroimaging

Infant

- Mental development
- Motor development
- Behaviour

- Measured in combination e.g. Bayley Scales of Infant Mental and Motor development
- Psychophysiological measures e.g. EEG
- Neuroimaging

Child

Will depend on age at assessment.

General domains:-

- Behavioural
- Sensory
- Motor
- Language
- Cognitive
  - Intelligence
  - Learning
  - Memory
  - Executive function
  - Communication
- Learning disabilities
- Mental disorders

Measures may vary considerably but intelligence measures such as WISC and general behaviour measures such as CBCL may be expected.

Also:-

- Detailed neuropsychological testing
- Psychophysiological measures e.g. EEG
- Neuroimaging

Summary descriptions of standard psychometric tests

Achenbach Child Behaviour Checklist (CBCL)

An instrument by which parents/carers rate a child's problem behaviours and competencies. The test focuses on behavioural and emotional problems in the previous 6 months. It measures aggression, hyperactivity, bullying, conduct problems, defiance and violence. Teacher Report Forms, Youth Self-Report Forms and Direct Observation Forms are also available.

Brazelton Neonatal Behavioural Assessment Scale (NBAS)

A scale devised to assess babies from birth to 2 months. The scale assesses infants across 4 different developmental areas: autonomic system, motor system, "state" regulation, social interaction.

Bayley Scales of Development

A standardised test of infant development for children in the age range 1 to 42 months. Development is measured in 3 domains: cognitive, motor and behavioural. Age standardised scores range from 50-150 with a mean=100 and s.d.=15. Significant delay is indicated in scores with two standard deviations below the mean e.g.<70.

Connors Rating Scale for ADD/ADHD

Scale consisting of 2 separate subscales to measure children's behaviour. The Connors Teacher's Rating Scale measures hyperactivity, conduct problems, emotional over-indulgence, anxious passivity, asocial behaviour and day-dream attention problems. The Connors Parent Rating Scale assesses conduct and learning problems, psychosomatic, impulse hyperactivity and anxiety. Both scales ml the behaviours of a child and compare them to levels of appropriate normal groups.

Griffiths Child Development Scale

Scale to assess child development in 5 areas: locomotor, personal-social, hearing & speech, eye & hand coordination, performance. There are 2 scales for children aged 0-2 years and 2-8 years. The score generates a development quotient (DQ) by summing all subscales with a mean=100 and s.d.=15. Global delay is a DQ score <70.

McCarthy Scales of Children's Ability (MSCA)

A measurement used to assess the abilities of pre-school children aged 2.5 to 8.5 years. It produces 6 scale scores: verbal, perceptual-performance, quantitative, composite (general cognitive), memory, motor.

Peabody Individual Achievement Test (PIAT)

A measure of academic achievement comprising 6 subtests: general information, reading recognition, reading comprehension, written expression, mathematics, spelling. It is designed for children 5 years and over. The standard scores have mean=100 and s.d.=15.

Stanford-Binet Intelligence Test

A test measuring intellectual and cognitive ability in children and adults aged 2 - 23 years. It covers 4 areas: verbal reasoning, quantitative reasoning, abstract/visual reasoning, short-term memory. Test mean=100 s.d.=16.

Wechsler Intelligence Scale Children (WISC)

A measure of general intellectual function for children aged 6-16 years. 12 subtests assess 2 areas of intelligence: verbal IQ (VIQ) and performance IQ (PIQ). Summated scores provide a full scale IQ (FSIQ) with mean=100 and s.d.=15.

Wide Range Achievement Test (WRAT)

An instrument to measure basic school skills and code comprehension for children aged 5 and over. Comprises 3 subtests: reading, spelling, arithmetic. The test is not designed to measure reasoning or judgement. It is used to assess level of learning ability/disability.

Appendix 5:
Medline search strategy

#22	(#19 not #20) and (LA:MEDS = ENGLISH) and (PY:MEDS >= 1970) and (TG:MEDS = HUMANS)	598
#21	#19 not #20	639
#20	(risk.mp or explode cohort studies / all subheadings or between groups.tw.) and (((low or light or social or moderate or dose or bing*) and ((explode “Alcohol-Related-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Alcohol-Drinking” / all SUBHEADINGS in MIME,MJME) or (explode “Ethanol-” / all SUBHEADINGS in MIME,MJME) or alcohol* or drinking)) and (#15 or #14 or #11 or #10 or #9 or #8 or #7 or #6 or #5 or #4)) and ((PT:MEDS = CASE-REPORTS) or (PT:MEDS = EDITORIAL) or (PT:MEDS = LETTER) or (PT:MEDS = REVIEW))	35
#19	(risk.mp or explode cohort studies / all subheadings or between groups.tw.) and (((low or light or social or moderate or dose or bing*) and ((explode “Alcohol-Related-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Alcohol-Drinking” / all SUBHEADINGS in MIME,MJME) or (explode “Ethanol-” / all SUBHEADINGS in MIME,MJME) or alcohol* or drinking)) and (#15 or #14 or #11 or #10 or #9 or #8 or #7 or #6 or #5 or #4))	674
#18	risk.mp or explode cohort studies / all subheadings or between groups.tw.	529208
#17	((low or light or social or moderate or dose or bing*) and ((explode “Alcohol-Related-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Alcohol-Drinking” / all SUBHEADINGS in MIME,MJME) or (explode “Ethanol-” / all SUBHEADINGS in MIME,MJME) or alcohol* or drinking)) and (#15 or #14 or #11 or #10 or #9 or #8 or #7 or #6 or #5 or #4)	4851
#16	#15 or #14 or #11 or #10 or #9 or #8 or #7 or #6 or #5 or #4	1000764
#15	(explode “Child-Development-Disorders-Pervasive” / all SUBHEADINGS in MIME,MJME) or (explode “Child-Language” / WITHOUT SUBHEADINGS in MIME,MJME) or (((mental retard*) or (learning disabil*) or neuro?development* or wisc* or cbcl) or (explode “Mental-Disorders-Diagnosed-in-Childhood” / all SUBHEADINGS in MIME,MJME) or (explode “Child-Development” / all SUBHEADINGS in MIME,MJME))	180847
#14	#12 and #13	293325

#13	(((explode “Intelligence-Tests” / all SUBHEADINGS in MIME,MJME) or (explode “Intelligence-” / all SUBHEADINGS in MIME,MJME) or (brain imag*) or (explode “Diagnostic-Imaging” / all SUBHEADINGS in MIME,MJME) or (neuro?behav*) or (explode “Neurobehavioral-Manifestations” / all SUBHEADINGS in MIME,MJME) or (explode “Psychophysiology-” / all SUBHEADINGS in MIME,MJME) or (explode “Psychological-Tests” / all SUBHEADINGS in MIME,MJME) or ((explode “Motor-Activity” / all SUBHEADINGS in MIME,MJME) or (explode “Hyperkinesis-” / all SUBHEADINGS in MIME,MJME) or (explode “Psychomotor-Performance” / all SUBHEADINGS in MIME,MJME))) or ((explode “Motor-Skills” / all SUBHEADINGS in MIME,MJME) or (explode “Motor-Skills-Disorders” / all SUBHEADINGS in MIME,MJME)) or ((explode “Language-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Language-Development” / all SUBHEADINGS in MIME,MJME) or (explode “Language-Development-Disorders” / all SUBHEADINGS in MIME,MJME)) or (executive function*) or ((explode “Memory-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Memory-” / all SUBHEADINGS in MIME,MJME)) or ((explode “Learning-” / all SUBHEADINGS in MIME,MJME) or (explode “Learning-Disorders” / all SUBHEADINGS in MIME,MJME)) or ((explode “Attention-Deficit-and-Disruptive-Behavior-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Attention-Deficit-Disorder-with-Hyperactivity” / all SUBHEADINGS in MIME,MJME) or (explode “Attention-” / all SUBHEADINGS in MIME,MJME))) or (cognit*) or ((explode “Cognition-” / all SUBHEADINGS in MIME,MJME) or (explode “Cognition-Disorders” / all SUBHEADINGS in MIME,MJME))	1763936
#12	neonat* or prenat* or infant* or child*	1802707
#11	(explode “Fetal-Alcohol-Syndrome” / all SUBHEADINGS in MIME,MJME) or (f?etal alcohol) or (alcohol embryopathy)	2883
#10	(explode “Abnormalities-” / all SUBHEADINGS in MIME,MJME) or (congenital anomal*) or malformation* or (birth defect*) or microcephaly or (head circumference)	318497
#9	(explode “Birth-Weight” / all SUBHEADINGS in MIME,MJME) or ((birth?weight) or (birth weight) or ((explode “Fetal-Growth-Retardation” / all SUBHEADINGS in MIME,MJME) or (explode “Growth-Disorders” / all SUBHEADINGS in MIME,MJME) or (growth restrict*) or (growth retard*) or (small for gestational age) or (low birth weight) or (antepartum h?emorrhage) or sga or lbw or elbw or vlbw or iugr))	74747
#8	(gestation*) or (explode “Gestational-Age” / WITHOUT SUBHEADINGS in MIME,MJME) or ((explode “Labor-Premature” / all SUBHEADINGS in MIME,MJME) or (explode “Infant-Premature” / all SUBHEADINGS in MIME,MJME) or (explode “Fetal-Membranes-Premature-Rupture” / all SUBHEADINGS in MIME,MJME) or (explode “Premature-Birth” / all SUBHEADINGS in MIME,MJME) or (explode “Infant-Premature-Diseases” / all SUBHEADINGS in MIME,MJME) or prematur* or preterm*)	206714

#7	neonatal death*	2939
#6	(explode “Fetal-Death” / all SUBHEADINGS in MIME,MJME) or (fetal loss*) or stillbirth*	23871
#5	(explode “Pregnancy-Complications” / all SUBHEADINGS in MIME,MJME) or (explode “Pregnancy-Outcome” / all SUBHEADINGS in MIME,MJME)	231043
#4	(explode “Abortion-Spontaneous” / all SUBHEADINGS in MIME,MJME) or miscarriage* or (spontaneous abortion*)	26183
#3	(low or light or social or moderate or dose or bing*) and ((explode “Alcohol-Related-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Alcohol-Drinking” / all SUBHEADINGS in MIME,MJME) or (explode “Ethanol-” / all SUBHEADINGS in MIME,MJME) or alcohol* or drinking)	66108
#2	low or light or social or moderate or dose or bing*	2080332
#1	(explode “Alcohol-Related-Disorders” / all SUBHEADINGS in MIME,MJME) or (explode “Alcohol-Drinking” / all SUBHEADINGS in MIME,MJME) or (explode “Ethanol-” / all SUBHEADINGS in MIME,MJME) or alcohol* or drinking	271952

Appendix 6:
Newcastle - Ottawa quality assessment scale

Cohort and cross-sectional studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort

a) truly representative of the average mother/child in the maternity hospital / community ❋
b) somewhat representative of the average mother/child in the maternity hospital / community ❋
c) selected group of users e.g.nurses, volunteers
d) no description of the derivation of the cohort

2) Selection of the non exposed cohort

a) drawn from the same community as the exposed cohort ❋
b) drawn from a different source
c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure

a) secure record (e.g.surgical records) ❋
b) structured interview ❋
c) written self report
d) no description

4) Demonstration that outcome of interest was not present at start of study

a) yes ❋
b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis

a) study controls for _____ (select the most important factor*) e.g. smoking, other drugs, postnatal environment*❋
b) study controls for any additional factors - please list ❋

Outcome

1) Assessment of outcome

a) independent blind assessment ❋
b) record linkage ❋
c) self report
d) no description

2) Was follow-up long enough for outcomes to occur

a) yes (select an adequate follow up period for outcome of interest) ❋
b) no

3) Adequacy of follow up of cohorts

a) complete follow up - all subjects accounted for ❋
b) subjects lost to follow up unlikely to introduce bias - small number lost - > 80 % follow up, or description provided of those lost) ❋
c) follow up rate < 80% and no description of those lost
d) no statement

CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?

a) yes, with independent validation ❋
b) yes, e.g.record linkage or based on self reports
c) no description

2) Representativeness of the cases

a) consecutive or obviously representative series of cases ❋
b) potential for selection biases or not stated

3) Selection of Controls

a) community controls ❋
b) hospital or clinic controls
c) no description

4) Definition of Controls

a) no history of disease (endpoint) ❋
b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis

a) study controls for _____ (Select the most important factor.) *) e.g. smoking, other drugs, postnatal environment* ❋
b) study controls for any additional factors - please list❋

Exposure

1) Ascertainment of exposure

a) secure record (e.g.surgical records) ❋
b) structured interview where blind to case/control status ❋
c) interview not blinded to case/control status
d) written self report or medical record only
e) no description

2) Same method of ascertainment for cases and controls

a) yes ❋
b) no

3) Non-Response rate

a) same rate for both groups ❋
b) non respondents described
c) rate different and no designation

Appendix 7:
Data extraction form

Study Id *(leave blank)* _____

Study. *(First author surname, year of publication)* _____

Relevant Outcomes *(Please refer to protocol for list of outcomes)*

1. _____ ***5.*** _____

2. _____ ***6.*** _____

3. _____ ***7.*** _____

4. _____ ***8.*** _____

Instructions: -

1. **Please ring multi-choice answers**
2. **Please print free text answers neatly**
3. **Please use terms 'no information', 'no description' or 'unclear'**
4. **Please use 'Notes' section to continue free text answers**
5. **If results are unclear or do not simply give adjusted and/or unadjusted ORs or RRs please don't waste time on this - leave blank and refer to RG. He and MQ will try to extract these later.**

Date extracted _____/05 Extracted by (initials) _____ Journal (initials) _____

List any relevant connected studies from same cohort or study *(First author surname, year of publication)*

_____

_____

_____

Methods

Design of study: 1.Cohort 2.Case control 3.Cross-sectional

Period of recruitment _____

Participants

Country _____

Number in study _____

Setting _____

Participation rate _____

Age (1. Mean, 2. Median 3. Range) of children at assessment _____

Alcohol exposure

Data collection: 1. Structured interview 2. Written self report 3. Other (specify) _____

Was data collected 1. Prospectively 2. Retrospectively

Did investigators measure exposure using:-

(a) Average amount consumed Y N

(b) Bingeing Y N

Did investigators report on timing of exposure during pregnancy Y (Specify) N

_____

Did investigators report on duration of exposure during pregnancy Y (Specify) N

_____

Was amount of alcohol consumed quantified in:- *(Please ring all that apply)*

1. 'drinks' 2. 'units' 3. grams, 4. ounces 5. ml

Describe how levels of average exposure (including any 'control' or 'abstainer' group) were named by the investigators, defined by them and quantified - **only include those within the range of interest (see protocol)**

(Where possible: give authors' definitions verbatim, followed by quantification of range in drinks or units and then equivalent in grams, ml or ounces as stated by the authors)

If there is a bingeing measure - describe how this is defined, used and quantified by the investigators. In particular is there a measure of the number of binge episodes and their timing during pregnancy?

Outcomes

How were outcomes defined? *(e.g. case definitions, events, test score, test cut-off)*

How were outcomes measured? *(e.g. maternal report, health record, clinical or psychological test)*

Study quality

Please score the appropriate Newcastle-Ottawa form, referring to the manual, and enter the number of stars for each section in the box below:-

	SELECTION	COMPARABILITY	EXPOSURE	OUTCOME
CASE CONTROL				
COHORT				

Study findings

Crude (and stratified) effect sizes with CIs/SEs *(use a table if possible and continue in Notes if you need more space)*

List factors adjusted for in the design and/or analysis and state method used for adjustment e.g. stratification, statistical modelling.

(If more than one adjustment – report the effect which is adjusted for the most confounders)

Notes:-

Appendix 8:
Summary of outcomes by first author

	Misc-arriage	Stillbirth/ APH	IUGR	Preterm birth	Low birth weight	Growth	Birth defect/ FAS	Neuro-develop-mental	Binge
Albertsen K				*					
Armstrong BG	*								
Bailey BN								*	*
Bell R				*	*		*		*
Berkowitz GS				*					
Brooke OG					*				
Davis PJ							*		
Day NL						*			
Day NL					*	*			
Day NL						*			
Ernhart CB							*		
Faden VB		*			*				
Forrest F								*	
Harlap S	*								
Henriksen TB	*								
Jacobson JL						*			
Jacobson JL								*	
Jacobson SW								*	
Kesmodel U				*					
Kesmodel U	*								
Kesmodel U		*							
Lazzaroni F				*					
Little RE		*							
Long MG	*								
Lumley J					*		*		
Lundsberg LS			*	*	*				

Marbury MC		*		*	*		*		
McDonald AD			*	*	*				
Mills JL							*		
Mills JL					*				
Nulman I								*	*
O'Callaghan FB					*	*			*
Ogston SA									
Olsen J								*	*
Olsen J							*		*
Ørskou J					*				
Parazzini F				*					
Parry GJ								*	
Passaro KT				*	*				*
Peacock JL					*				
Peacock JL				*					
Plant ML							*		*
Primatesta P					*	*			
Raymond EG		*							
Sampson PD					*	*			*
Shiono PH				*					
Shu XO					*				
Sood B									
Stoler JM							*		
Streissguth AP								*	*
Streissguth AP								*	*
Streissguth AP								*	
Streissguth AP								*	*
Sulaiman ND				*	*	*			
Tolo KA			*	*	*	*			*

Verkerk PH			*	*					
Verkerk PH				*	*				
Virji SK					*				
Virji SK					*				
Whitehead N			*						*
Windham GC			*	*	*				
Windham GC	*								
Windham GC	*								
Wisborg K				*					
Yang Q			*						