Appendix: Selection of controls and Statistical methods

Selection of controls

ECAS dataset
For three studies, regional coordinators selected 6 live controls of the same gender and born at the same maternity ward 14 days subsequent to the index case. A delay period of 14 days was expected to assure that controls had an age similar to the SIDS baby when parents completed the questionnaire. Parents of the first four of these selected infants were invited to participate. If a family was unwilling to participate, another family among the two families of the remaining infants was invited.

For the other 12 studies included in this dataset it was intended that at least two live controls were obtained for each case. Almost all cases in these studies had 2 controls; all had one. These controls were selected from a list of births in the area and born within one week before or after the case. Controls were not matched for any other characteristic. Initially four controls were selected to be used as replacements if necessary.

GsSIDS dataset
For each case, 10 controls were selected that were matched for region, age, gender, and reference sleep. The control infants were recruited through the same or neighbouring local birth registration office where the case was registered. Control infants had been born 4–6 weeks after the case infant, so that by the time the interviews were done they had the same age as the index case (±2 weeks). From the control families who agreed, the three infants closest in age to the index case were selected. A total of 2702 controls were contacted; 58.7% agreed to participate.

Irish dataset
For every case notified to the SIDS register, four controls were selected randomly from the birth register and matched for date of birth and geographical location (same community care area as the index case). If an insufficient number of infants were born in the same community care area on a particular date, then a list of infants born on the two days either side of that date was also used. All families were invited by letter to participate in a standardized home interview. Where no response was obtained from controls families within one week, an additional four letters were sent, after which no further attempt at recruitment/replacement of controls was made. Information was collected on socio-demographics, pregnancy, the infant/child's medical history, the home environment, parenting practices and details of the last 48 hours, and last sleep period with a corresponding reference sleep period used for controls. An average of three controls per case were recruited; in the final dataset, the proportion of cases that had 4 completed control questionnaires was 33%, 3 control questionnaires = 22%, 2 controls = 20%, 1=11%, 6% had >4controls and 7% of cases had no corresponding control data.

New Zealand dataset
Controls were randomly selected from all births, except home births (less than 1%) in the participating regions. Controls had to be born and domiciled in the study region.

The following method was used to select controls:

(a) A date of interview (nominated date) was randomly selected.
(b) The control was then randomly allocated an age and date at interview.

(c) Births by day of the week vary considerably, probably because of induction of labour. The day of birth was adjusted to fit this distribution.
(d) An obstetric hospital was randomly chosen in proportion to the number of births over the previous year.
(e) In hospitals with more than one birth on the selected day random numbers were used to select a particular infant from among those born on the nominated day. For obstetric hospitals where there were no deliveries, a random direction indicator was used to indicate whether to go forwards or backwards in time to select the infant.

Thus, the controls were a representative sample of all live births in the study regions.

For questions on infant care practice that particularly related to the period of sleep prior to the death in the cases, parents of controls were given a nominated time of day which was randomly allocated to ensure that the distribution of this time among controls was similar to the estimated distribution of the time of death of the cases. If the infant was not asleep at the nominated time of day the direction indicator was used to select either the previous or subsequent sleep.

Scottish dataset

We identified babies born immediately before and after the index case in the same maternity unit to act as controls (2 controls for each index case). Controls were therefore matched for age, season, and maternity unit. If no contact could be made with the baby born immediately before the index case (or immediately after), then the baby born immediately before that first attempted control (or immediately after) was also attempted. If neither of the 2 babies born before or 2 babies born after could be contacted and a visit completed within 28 days of the index infant's death then no further attempts were made to contact other baby's parents to act as controls for the index case.

Statistical methods

Missing data

Preliminary analysis, together with the study context, showed that missing values were most plausibly missing at random dependent on study. Therefore, since we include study indicators as covariates, a complete records analysis will give unbiased if somewhat inefficient inference[A1]. To include the information from studies in which alcohol and drug use data were not observed, we used multiple imputation (under the missing at random assumption) to impute missing data. We used the REALCOM-IMPUTE software[A2] with a single level imputation model because alcohol and drug data were too sparse, among the studies in which they were recorded, to obtain convergence for a multilevel imputation model. Missing data were imputed for cases and controls separately. Ten imputed data sets were computed. Using STATA 12[A3] the substantive multilevel model was fitted to each in turn. Convergence was not achieved for one because the likelihood was flat in the region of the maximum; the results for the remaining 9 were combined for inference using Rubin's rules[A4].

Analysis showed that the between imputation variation across the 9 imputed data sets was small relative to the within imputation variance, so 9 imputations were sufficient.

Reliability of results based on observed and imputed data

First, define the *key sub group* as babies < 3 months who were breast fed whose parents did not smoke and whose mother took less than 2 units of alcohol in the last 24 hours and was not a drug user, who either bed shared or room shared. We have data on both maternal and paternal alcohol consumption in the last 24 hours and drug use after birth for two datasets, and for the key subgroup of cases and controls, we have extracted the paternal data from the original records. The unadjusted OR for bed sharing in this group is 5.6 (1.6– 20.3), p = 0.009. And for this group, in both the bed sharing and room sharing groups the control partners had taken slightly *more* alcohol in the last 24 hours than the cases' partners. Consequently, after adjusting for partner's alcohol consumption in the last 24 hours, the OR is 7.7 (1.8 – 32.3), although the OR for partner's alcohol is not significant; OR = 0.73 (0.41 – 1.27), p =0.265.

For cases, belonging to the key subgroup in the three studies for which maternal alcohol use in the last 24 hours was not available , we have checked the original records, most of which include pertinent questions about alcohol use, and ensured that alcohol and drugs were not contributory factors in any.

Second, the prevalence of alcohol and drug use among mothers varies considerably across the studies where the information was collected. For controls, the prevalence of mother having more than 2 units of alcohol in the last twenty four hours (henceforth 'mother using alcohol') ranged from 0 to 9%, and the prevalence of mother using any illegal drug (henceforth 'mother using drugs') ranged from 0 to 0.6%. For cases the corresponding percentages range from 0 to 39% and 0 to 3% respectively. Consequently the ORs for mother using alcohol vary significantly across the studies. However, there is no evidence that the three-way interaction of mother using alcohol, bed sharing and study is significant, p = 0.429. Therefore, the relationship between bed sharing and study does not vary by mother using alcohol. In consequence, the OR for bed sharing is not affected by varying prevalence of mother using alcohol across the studies. For mother using drugs the data are too sparse for the analogous three-way interaction to be tested. However, it seems unlikely it would be significant. In consequence, the OR for bed sharing is not affected by varying prevalence of mother using drugs across the studies.

Third, because the alcohol and drug data are plausibly missing at random, MAR, dependent on study, which is included as an indicator variable in both the substantive model and the imputation model, theory suggests that the point estimates in the complete records analysis should be unbiased,[A5] and within sampling variation of those obtained after multiple imputation. The advantage of multiple imputation here is thus the recovery of information, primarily through the inclusion of the partially observed data from the three studies in which alcohol and drug use were not collected, c.f., Carpenter and Kenwood, p 220.[A5] The results are in line with this, as shown in Table 1, columns 8-11. Also as reported above the OR for the key subgroup is 5.6 (1.6– 20.3). The number of observations in this subgroup are too small to attempt adjustment for other factors like maternal age parity and birth weight. Compare this subgroup OR with the fully adjusted AOR of 5.1 ((2.3 – 11.4) for breast fed babies < 3 month, whose parents do not smoke and whose mother did not take two units alcohol or more in the last 24 hours alcohol. or use drugs. This AOR is also adjusted for all the other factors in the model, see Table 3. The narrower CI results from the recovery of the partially observed data.

Calculation of univariate and multivariate odds ratios

Odds ratios were calculated by logit regression. Univariate analyses were adjusted for age and study because controls were on average 3 weeks older than cases, and the number of controls varied between studies. For multivariate AORs, multilevel logit regression model was fitted with 'bed sharing' random across studies; this was done to take account of a significant interaction of bed sharing with studies. Some other AORs showed significant interaction with studies; however, it was found that these were due to significant deviations in one or at most two studies. When parameters were added to the overall model, to account for these interactions, they had little effect on the main parameters, and only slightly *increased* the estimate of risk associated with bed sharing. The additional parameters were therefore dropped in the final model and these interactions ignored.

The trend in the ln(OR) for bed sharing with age was best represented by a linear downward trend on the logit scale, for the first six months followed by a constant term thereafter. In all four models were used for the analysis:

Model 1. A multilevel logit model of the whole data, including the interaction of age and bed sharing, modelled by the linear trend,

Model 2. To obtain rates applicable to all ages, the same model, excluding the age×bedsharing interaction was fitted, thereby obtaining average AOR for the year.

Models 3 & 4. To obtain average AORs for the first three months and later, a logistic forms of the rates model was fitted to records of infants under 3 months and 3 month or more. Logistic models were used because of convergence problems with multilevel models.

Goodness of fit of the models to the data

Goodness of fit tests are not available for multilevel logit models nor are they available after using Rubin's combination rules for the analysis of multiple imputed data sets. Therefore single level (i.e., standard) logistic models, using the same parameters as the overall model plus fixed effect parameters for study, were fitted to each of the 10 data sets completed with imputed data; both the log link and goodness of fit tests were applied to each. The link tests confirmed that all the models were correctly specified: p(for regression on hat^2) averaged 0.44 and all were> 0.15, and p(for the constant) averaged 0.75 and all were >0.56. The average Hosmer-Lemeshow goodness of fit $\chi^2(48) = 40.3$ was less than expectation. and none had a p value < 0.13. It was, therefore, concluded that the model fit was excellent. Checks on the model, without the age trend, fitted to infants aged <3 months showed equally good fit.

To check the fit of the overall model to the data relating to the breast fed cases, age <3 months, whose parents did not smoke and whose mothers did not consume alcohol or use drugs but who were bed sharing, their deviance residuals were computed. The AOR for this groups is represented by the lower line in Fig 2. As above, the deviance residuals could only be computed after fitting a logistic model to each of the 10 completed data sets. Again, the The results were pooled using Rubin's rules[A4]. It was found that the mean deviance for this group = - 0.098, s.e. 0.1004. Also there was not evidence of any systematic deviation from the fitted line in that there was no evidence of a trend in the residual deviances with age; b = -0.0015, s.e. 0.005.

Similarly residual deviances were computed for this group after fitting model 3. The pooled average residual deviance was -0.147 with s.e. 0.096; p = 0.122. The trend in the residuals was 0.00012 with s.e. 0.005. Thus, there is no suggestion that the model parameters do not represent these crucial data.

The Attributable Fraction

The attributable fraction (of deaths, computed as described by Brussi et al.[29]), was similarly computed for each of the 10 logistic models fitted to the imputed data sets. The results were combined using Rubin's combination rules.[A4]

Mortality rates

Rates were derived from the parameters of Model 2. Rates are given for all infants, computed by a weighted combination of the rates for boys and girls. The base rate for girls was the SIDS rate when none of the model risk factors were present. Then, logit(base rate) = model constant scaled by the addition of the logit of the population SIDS rate and the subtraction of the log(ratio of the number of cases to controls in the model). Combinations of AORs gave other rates from the base rate.

Estimating AORs and Rates for other groups

The AORs computed for other groups, as described on page 7 are approximate because the AORs for the factors which do not interact with age or bed sharing vary, but not significantly, across the 4 models used for the analyses. The AORs shown in the penultimate column of Table 1 are those given by model 2. These differ a little from the comparable AORs given by the Model 1, which includes the age×bed sharing interaction. Thus for the example on page 7, the AOR predicted by model 1 is 4,402 (1,758–11,022) compared with 4514 shown.

When computing SIDS rates for other groups from those give in Table 4, the procedure is similar. However, the observed rate must first be divided by 7.43 to reduce the rate baseline – the rates reported in Table 4 relate the second infant with birth weight 2500 – 3499g of a cohabiting white women age 26 to 30. The appropriate baseline rate, i.e., for various smoking groups may then be scaled up according to the other risk factors present. However, if the computed rate is r > 0.003 per 1000, it should be reduced by $-r^2$, because the scaling is based on AORs and rates are probabilities. Conversely if the starting rate is >0.003 it has first to be scaled to an AOR by adding its square.

For example the estimated SIDS rate for a bed sharing 18 year old cohabiting white mother, with her 1st baby, birth weight 2240g. bottle fed when both parents smoke and mother often has 2+units of alcohol is estimated to be

$$r = \{(0.0275 + 0.0275^2)/7.43\}\times 4.2\times 9.1 = 145.4$$

where:

- 0.0275 = rate from Table 4 when both smoke, mother uses alcohol and baby is bottle fed
- 0.0275^2 is added to obtain the corresponding AOR because the starting rate is >0.003
- /7.43 to obtain the corresponding baseline AOR
- ×4.2 from Table 1 for babies 2000-2499
- ×9.1 from Table 1 for mothers aged 18

Thus, r > 0.003. Hence

$$\text{Predicted rate per 1000} = 1000*(r-r^2) = 125 \text{ per } 1000,$$

which is exact because the AORs in Table 1 are derived from Model 2.

Supplementary tables show predicted SIDS rates for two groups of women other than those in Table 4.

Rates may also be scaled up or down in direct relation to the population SIDS rate. Thus if the population SIDS rate is 0.4 per 1000 instead of 0.5 the the estimated rates will be reduced by 4/5 =0.8.

Supplementary tables of predicted rates for two other groups of women.

a) Cohabiting white women age 30+ with 1st baby birth weight >3500g

Group No.	Risk factors present			Room sharing	Bed sharing	Ratio of rates	
	Feeding	smoking	Alcohol	Rate/1000	Rate/1000	Ratio	95% CI
Baseline	Br	no	no	**0.011**	**0.031**	**2.7**	1.4–5.3
1	**Bot**	no	no	**0.017**	**0.047**	**2.7**	1.4–5.3
2	Br	**P**	no	**0.013**	**0.070**	**5.6**	2.9–10.8
3	Br	**M**	no	**0.018**	**0.171**	**9.7**	4.4–21.7
4	Br	**B**	no	**0.033**	**0.254**	**7.7**	4.3–13.8
5	**Bot**	**B**	**Y**	**0.235**	**3.74**	**16.0**	5.8–44.2

OK 9/9/12

b) Cohabiting white women age 18 - 19 with 1st baby with birth weight 2000 - 2499g

Group No.	Risk factors present			Room sharing	Bed sharing	Ratio of rates	
	Feeding	smoking	Alcohol	Rate/1000	Rate/1000	Ratio	95% CI
Baseline	Br	no	no	**0.4**	**1.2**	**2.7**	1.4–5.3
1	**Bot**	no	no	**0.6**	**1.8**	**2.7**	1.4–5.3
2	Br	**P**	no	**0.5**	**2.7**	**5.6**	2.9–10.8
3	Br	**M**	no	**0.7**	**6.5**	**9.7**	4.4–21.7
4	Br	**B**	no	**1.2**	**9.5**	**7.6**	4.3–13.6
5	**Bot**	**B**	**Y**	**8.8**	**124.6**	**14.1**	5.7–39.0

References

A1 Carpenter JR, Goldstein H, Kenwood MG. (2012) Statistical modelling of partially observed data using multiple imputation: principles and practice. p20. In: Modern Methods for Epidemiology. Ed. Greenwood DC & Tu Y. Springer, London.2012.

A2 Carpenter JR, Goldstein H, Kenward MG. (2012). REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response Types. J Statistical Software. **45** :5: 1-14.

A3 StataCorp 2011. Stata Statistical Software: Release 12.1. College Station, TX: Stata Corporation.

A4 Rubin D. (1987) Multiple Imputation for Non-response in Surveys. Wiley. Chichester.

A5 Carpenter JR, and Kenward MG. Multiple Imputation and its Application. Chichester: Wiley, pp28 and 220.