Methodological appendices for

Carpenter RG, Irgens LM, Blair P, England PD, Fleming P, Huber J, Jorch G, Schreuder P. Sudden unexplained infant death in 20 regions in Europe: case control study. Lancet 2004; 363: 185-191.

Appendix 1 Statistical Methods

Seasonality

Almost all the studies lasted complete calendar years, but the NESS studies were extended to include a few deaths outside the formal survey period, and the study in the 2nd CESDI region started 7 months after the studies in the other two CESDI regions. To examine the seasonality, deaths that occurred outside the formal survey period and in the first 5 months of the 2nd CESDI study were excluded. The remaining deaths were tabulated by month of occurrence. To model the seasonal variation the values of the cosine function was tabulated at $\pi/12$, $3\pi/12$, etc, corresponding to the mid point of each month. Then, ignoring the comparatively small variation in birth rates, a cosine curve was fitted to the counts using Poisson regression. The phase and amplitude of the seasonal variation were derived from the regression coefficients.

Logistic regression modelling

In the ECAS study, the controls were not matched, but were selected by sampling an appropriate register to represent the general population of infants in the district of similar age at the corresponding time of year. Therefore, as in other studies[1,2] unconditional logistic regression methods were used. In single factor analyses conditional and unconditional methods give almost identical estimates and confidence limits - see Tables in Appendix 2. Multivariately, conditional estimates have confidence limits about 40%

wider than unconditional estimates, and the method is much less efficient, because only differences between cases and controls in the same matched set are used.

Log(ORs) for the risk factors correspond to the logistic regression coefficients in a model. The constant terms in logistic regression largely depends on the ratio of cases to controls in the data. This may be seen by fitting a constant alone. Then the constant coefficient is the log of the ratio of the number of cases and controls. In this study, the number of controls per case varied from four to one in the different centres. As in multivariate regression, including a constant term for each centre in the model of the whole data ensures that the regression coefficients are the *pooled mean within centre* coefficients. The constants for centres take account of the differing ratios of cases and controls across the centres. The constants themselves are uninformative and the corresponding ORs for the centres are not shown.

Single factor analyses, Appendix 2, are based on records with entries in the relevant categories. The percentage of complete records is shown for each variable. When χ^2 test revealed more than random inter-centre variation in the odds ratios, the confidence limits for the pooled OR were adjusted using the Huber/White method[3].

Data on age and 32 variables were substantially complete (average 95.5%), but occasional missing readings were scattered through the records; e.g., there was often no data on partner's employment when the mother was single. If the multivariate analysis had been confined to cases and controls with complete data, 25% of records would have

been excluded. Missing data were therefore imputed for cases and controls separately to the values predicted by the other most closely related variables[4] These were determined, for continuous variables by multiple regression, for binary variables by multiple logistic regression, and for categorical variables by multiple multinomial (polytomous) logistic regression. The STATA routine **impute** was used to impute continuous variables, e.g., birth weight and gestational age. This routine was then modified to impute binary and categorical data using **logit** and **mlogit** routines[5]. Data on the variables included in models presented in table 2 and 4 were over 97.2% complete.)

The multivariate model was built up manually by a forward stepwise procedure where variable selection depended on the significance of the variable as a whole, and not on the significance of individual factors, with $p > 0.05$ as the critical level for both inclusion and exclusion. Interactions have not been generally explored except where suggested by other studies.

Age was included in all models described in the paper to account for any differences due to differences in age between cases and controls, but ORs presented in Appendix 2 have not been adjusted for age.

Two significant interactions with age were reported for the basic model, implying that the ORs for these factors change with the age at death. The ORs for these factors were reported at the median age of 10 weeks. These ORs were derived by centring age in days at 70 days, i.e., using (age - 70) in place of age. The confidence belts shown in Fig. 1

were derived by repeating the analysis with age centred at a succession of ages ranging from 2 to 26 weeks.

Adjusted ORs for incomplete variables were obtained by adjusting for all the other factors fitted for the model described in Table 2, the basic model. Technically, the ORs were obtained by logistic regression of the outcome variable, cases / control, on the *added variable residuals*[6] of each incomplete variable using as offset the predicted values derived from the basic model.

Inter-centre homogeneity of the ORs was tested as follows: first, the likelihood ratio χ^2 for the interaction of each variable with centre was computed and then totalled. The total χ^2 for the interactions was significant, $p < 0.003$, and a plot of standard Normal deviates, z, corresponding to the interaction χ^2s showed several were larger than expected. For these variables, computations of the deviations of the adjusted ORs for each centre from the pooled adjusted OR revealed the four outliers, noted in Table 3, for which the deviations were significant by a Bonferroni test. After fitting parameters for these outliers, the total interaction χ^2 was not significant, $p = 0.272$, and the z plot showed that none of the interactions with centres were significantly large. However, to exclude the data responsible for these outliers would have necessitated excluding all the data from the corresponding centres because complete records are required for the multivariate logistic regression. Therefore, parameters modelling the outliers were included in all appropriate logistic models.

Additional unreported parameters were also used to eliminate the effects on the OR for sex caused by the use of controls matched for sex in the NESS and Nordrein Westfallen studies.

When the searches for interactions and outliers were repeated for the variables listed in Tables 4 and 5, no further outliers were detected, but the confidence limits for three factors were adjusted for more than random inter-centre variation.

STATA's **lfit** command was used to test the goodness of fit of the multivariate logistic models. This command collapsed the data into 20 quantiles of probability and compared observed and expected numbers of cases and controls in each quantile and assessed goodness of fit using Hosmer-Lemeshow's[7] χ^2. The logistic link was also tested using STATA's **linktest** command. Neither test gave any indication of poor fit.

Calculation of attributable Fractions.

AFs were calculated from percentages and the corresponding ORs given in Tables 2, 3 & 4, see Methods. To determine the AFs for bed sharing and sleeping in another room, the model was re-run with '*room shared in last sleep (not bed sharing)*' included and omitting the age interaction term. The AF for bed sharing was calculated as

$$4.5 \times (1 - 1/2.06) + 15.4 \times (1 - 2/11.92) = 15.9$$

where the OR for smoking and bed sharing (11.92) is divided by 2, the average OR for mother smoking in the absence of bed sharing.

The OR for mother smoked in pregnancy was similarly derived. In this case, the OR for mother smoked & bed shared was divided by the OR for bed sharing in the absence of smoking. (1.206).

A full account of these analyses will be published elsewhere.

References

[1] Hoffman HJ, Denman DW, Damus K, Bell G van. Comparison of matched versus unmatched analyses in a case-control study of SIDS risk factors. In: Am. Statis. Ass. Pro. Social Statist Section. pp 318 - 323. 1987.

[2] L'Hoir MP, Engelberts AC, van Well GThJ, McClelland S, Westers P, et al. Risk and preventive factors for cot death in the Netherlands, a low-incidence country. European J. Paediatr. **157**, 681 - 688. 1998.

[3] Stata Corporation. *Stata User's Guide. Release 5.0.* College Station, TX 1997.

[4] The STATA impute routine was used for continuous variables and corresponding routines were written for binary and categorical variables.

[5] Flynn R. Missing data in the European Concerted Action on Sudden Infnat Death Syndrome, ECAS. MSc. Thesis, Medical Statistics Unit, London School of Hygiene & Tropical Medicine. 1999.

[6] Collett D. *Modelling Binary data.* p 133. Chapman & Hall.CRC. Boca Raton, Florida. 1999.

[7] Hosmer DW, Jr, Lemeshow S. Goodness-of-fit tests for the multiple regression model. Communications in Statistics A9: 1043-1069.

Appendix 2

Summary of Unifactor Analyses of Potential Risk Factors for SIDS
A: All Variables

Pooled Odds Ratios and 95% Confidence Limits[1]

			Percentage		Odds	Lower	Upper
		% Data[2]	Cases	Controls	Ratio	Limit	Limit
(a) Last days before death/interview							
Position usually left:	**side vs. back**	82.5	38.3	37.4	**1.72***	1.34	2.21
	front vs. back		30.4	12.0	**5.88***	4.02	8.60
Position last left:	**side vs. back**	97.8	35.6	37.5	**1.91***	1.49	2.46
	front vs. back		38.2	11.3	**9.10***	6.35	13.02
Position usually found:	**side vs. back**	83.5	24.3	14.7	**2.94**	2.25	3.70
	front vs. back		36.5	14.7	**5.00***	3.50	7.15
Position last found:	**side vs. back**	96.6	18.2	17.3	**2.95***	2.15	4.04
	front vs. back		56.5	14.8	**13.87***	10.12	19.00
Head covered when found:	**Yes vs. No**	84.3	25.0	3.4	**9.88**	7.27	13.42
Duvet used usually:	**Yes vs. No**	91.7	60.9	52.2	**1.58***	1.16	2.16
Duvet used on last occasion:	**Yes vs. No**	56.2	48.1	27.2	**2.53**	1.91	3.34
Hat worn to sleep usually:	**Yes vs. No**	92.0	4.3	2.5	**1.10**	0.61	2.00
Hat worn to sleep on last occasion:	**Yes vs. No**	56.3	8.6	4.5	**1.53**	0.91	2.57
Sweating in last 24 hours:	**Yes vs. No**	56.6	28.0	19.4	**1.55**	1.19	2.02
Evidence of sweating when found:	**Yes vs. No**	50.0	22.7	10.2	**2.51**	1.82	3.47
Dummy ever used:	**Yes vs. No**	90.7	62.5	66.9	**0.88**	0.72	1.06
Dummy used in last sleep:	**Yes vs. No**	49.0	36.2	55.1	**0.47***	0.34	0.64
Bed-sharing in last sleep:	**Yes vs. No**	99.7	19.4	8.1	**2.93**	2.3	3.72
Room shared (in own cot) usually:	**Yes vs. No**	56.2	46.8	58.8	**0.71***	0.53	0.95
Room shared (in own cot) in last sleep:	**Yes vs. No**	44.3	28.0	44.5	**0.47**	0.35	0.63
Cough, cold or earache in last week:	**Yes vs. No**	99.0	28.7	20.2	**1.59***	1.13	2.25
Immunised/vaccination in last 7 days:	**Yes vs. No**	62.7	10.9	10.4	**1.27**	0.89	1.81
(b) Development							
Initial feeding:	**breast & bottle only vs. breast only**	91.4	27.0	21.5	**1.83**	1.46	2.29
	bottle only vs. breast only		18.6	12.0	**2.83**	2.06	3.88
Admitted to hospital since birth:	**Yes vs. No**	98.4	23.5	13.6	**2.07**	1.67	2.57
History of ALTE:	**Yes vs. No**	98.8	11.2	3.0	**4.12**	2.94	5.77
(c) Birth							
Sex[3]:	**Male vs. Female**	100.0	61.2	49.1	**1.65**	1.32	2.06
Multiple Birth	**vs. Singleton**	99.6	6.4	1.5	**4.49**	2.86	7.05
Birthweight:	**> 3500g**	99.2	28.7	49.1	**1.00**		
	2500 to 3500		51.6	47.3	**1.88**	1.55	2.27
	2000 to 2500		11.4	2.5	**8.12**	5.6	11.78
	< 2000		8.3	1.2	**13.91**	8.6	22.51
Gestation:	**>= 40 wks**	98.4	52.4	72.4	**1.00**		
	38-39		22.2	19.5	**1.62**	1.28	2.05
	35-37		16.4	6.9	**3.42**	2.62	4.47
	< 35 wks		9.0	1.1	**12.60**	7.87	20.19
Birthweight Z-score:	**> 75th centile**	98.1	23.6	37.2	**1.00**		
	50th-75th		22.0	25.4	**1.37**	1.07	1.74
	25th-49th		22.3	21.7	**1.61**	1.26	2.05
	5th to 24th		24.0	12.9	**2.97**	2.31	3.82
	< 5th		8.1	2.9	**4.43**	2.99	6.56
APGAR at 5 mins < 9:	**Yes vs. No**	61.8	14.4	7.1	**2.63**	1.78	3.9
Breech delivery:	**Yes vs. No**	90.6	4.4	2.7	**1.70**	1.06	2.71
Caesarian section:	**Yes vs. No**	98.6	17.9	13.4	**1.44**	1.15	1.8
Admitted to SCBU*:	**Yes vs. No**	91.3	24.4	9.2	**3.39***	2.52	4.55
(d) Pregnancy							
Mother's Age:	**>=31**	99.3	22.5	36.2	**1.00**		
	26-30		30.3	36.5	**1.39**	1.11	1.75
	21-25		30.9	22.1	**2.52**	1.99	3.2
	19-20		10.2	3.8	**5.05**	3.51	7.24
	<=18		6.2	1.3	**9.81**	5.94	16.2
UTI in pregnancy:	**Yes vs. No**	97.5	11.2	6.4	**2.00**	1.49	2.68

Continued

	% Data	Percentage Cases	Percentage Controls	Odds Ratio	Lower Limit	Upper Limit
(e) Obstetric History						
Time since previous live birth < 12 months: Yes vs. No	42.4	13.4	4.0	**3.86**	2.45	6.07
Interval in months: > 48 months	35.4	21.7	30.8	**1.00**		
37-48		14.5	15.5	**1.39**	0.87	2.22
25-36		24.6	24.1	**1.60**	1.07	2.42
13-24		29.0	26.4	**1.83**	1.23	2.71
<=12		10.1	3.2	**5.57**	3.01	10.29
Previous Live Births: None	99.6	26.9	41.5	**1.00**		
1		34.6	35.8	**1.57**	1.27	1.93
2		21.7	15.5	**2.42**	1.9	3.1
3		9.9	5.1	**3.45**	2.47	4.83
4+		6.8	2.1	**5.50**	3.56	8.51
Previous fetal losses: Yes vs. No	98.3	31.4	26.3	**3.86**	2.45	6.07
Number of losses: None	98.3	68.6	73.7	**1.00**		
1		20.2	19.1	**1.13**	0.91	1.4
2		8.3	5.1	**1.81**	1.29	2.53
3		1.8	1.6	**1.29**	0.67	2.48
4+		1.1	0.6	**2.21**	0.9	5.44
Previous infant death: Yes vs. No	92.2	3.9	1.7	**2.22**	1.33	3.72
(f) Smoking, Alcohol, Drugs and Caffeine						
Mother smoked before pregnancy: Yes vs. No	77.1	67.0	36.0	**3.87***	3.06	4.90
Amount smoked: None	76.8	33.1	64.1	**1.00**		
1 to 9		13.9	10.9	**2.42**	1.76	3.32
10 to 19		31.3	16.3	**4.17**	3.24	5.37
20 to 29		17.8	7.9	**4.64**	3.41	6.3
30 +		4.7	0.8	**12.18**	6.16	24.08
Mother smoked during pregnancy: Yes vs. No	99.7	61.8	26.5	**4.90***	3.93	6.10
Amount smoked: None	98.2	39.1	73.6	**1.00**		
1 to 9		23.9	15.1	**3.17**	2.52	3.98
10 to 19		24.7	7.9	**6.48**	5.06	8.3
20 to 29		9.0	2.7	**7.47**	5.1	10.94
30 +		3.3	0.7	**10.37**	5.4	19.92
Mother smoked after pregnancy: Yes vs. No	91.4	60.8	27.9	**4.38**	3.62	5.30
Amount smoked*: None	89.8	39.7	73.5	**1.00**		
1 to 9		21.7	11.4	**3.79**	2.82	5.11
10 to 19		22.3	10.1	**4.45**	3.11	6.38
20 to 29		12.5	4.6	**5.64**	4.19	7.58
30 +		3.8	0.4	**24.29**	16.31	36.16
Others smoked before pregnancy: Yes vs. No	45.9	55.4	38.1	**2.12**	1.65	2.73
Amount smoked: None	44.7	45.7	63.6	**1.00**		
1 to 9		9.0	8.9	**1.33**	0.84	2.09
10 to 19		18.8	15.1	**1.96**	1.38	2.78
20 to 29		20.2	10.3	**2.99**	2.08	4.28
30 +		6.4	2.0	**4.42**	2.29	8.53
Others smoked during pregnancy: Yes vs. No	67.5	57.0	35.3	**2.58***	1.84	3.60
Amount smoked: None	59.1	44.0	65.5	**1.00**		
1 to 9		8.4	9.8	**1.21**	0.8	1.83
10 to 19		19.4	13.4	**2.38**	1.74	3.26
20 to 29		19.0	9.2	**3.42**	2.46	4.75
30 +		9.1	2.0	**8.63**	5.04	14.8
Others smoked after pregnancy: Yes vs. No	98.6	60.2	34.5	**3.13**	2.42	4.04
Amount smoked: None	90.0	41.8	66.0	**1.00**		
1 to 9		10.3	9.7	**1.76**	1.28	2.41
10 to 19		21.2	13.4	**2.76**	2.14	3.56
20 to 29		17.8	9.1	**3.34**	2.54	4.4
30 +		8.9	1.8	**8.77**	5.64	13.66
Mother used drugs before pregnancy: Yes vs. No	79.4	6.3	2.6	**2.95**	1.87	4.66
Mother used drugs during pregnancy: Yes vs. No	79.4	4.2	0.7	**6.89**	3.49	13.59
Mother used drugs after birth: Yes vs. No	79.5	3.7	1.1	**4.12**	2.21	7.71
Tea/Coffee consumption during pregnancy: None	48.6	39.5	44.9	**1.00**		
1-2 cups per day vs. None		24.4	30.6	**0.88**	0.65	1.19
3+ cups per day vs. None		36.1	24.5	**1.85**	1.39	2.45
Mother's alcohol consumption in last 24 hours: None	77.9	84.0	87.1	**1.00**		
1-2 drinks per day vs. None		9.0	10.2	**0.91**	0.65	1.27
3+ drinks per day vs. None		7.0	2.8	**2.92**	1.9	4.51
Partner's alcohol consumption in last 24 hours: None	74.6	70.0	74.2	**1.00**		
1-2 drinks per day vs. None		16.6	16.1	**0.88**	0.66	1.18
3+ drinks per day vs. None		13.5	9.7	**1.47**	1.08	2.01

Continued

		% Data	Percentage Cases	Percentage Controls	Odds Ratio	Lower Limit	Upper Limit
(g) Social Circumstances							
Mother's Race:	**Caucasian - No vs. Yes**	99.4	5.1	2.9	**1.93**	1.26	2.94
Marital Status:	**Married**	99.6	53.0	68.5	**1.00**		
	Cohabiting vs. Married		29.8	24.8	**1.87***	1.05	3.32
	"Single" vs. Married		17.2	6.7	**3.90***	2.67	5.71
Mother's Education:	**FE College and Higher**	97.8	13.0	24.0	**1.00**		
	> Secondary		29.3	32.3	**1.47***	1.04	2.09
	Secondary only		39.2	35.3	**2.99***	1.91	4.68
	Less than Secondary		18.6	8.4	**6.64***	4.40	10.03
Father's Education:	**FE College and Higher**	80.3	16.2	26.7	**1.00**		
	> Secondary		32.4	33.8	**1.53**	1.15	2.05
	Secondary only		32.6	29.6	**2.53**	1.88	3.40
	Less than Secondary		18.8	9.9	**4.63**	3.25	6.63
Mother Unemployed vs. Employed		98.8	74.3	63.0	**2.25**	1.81	2.79
Partner Unemployed vs. Employed		95.2	29.8	11.7	**3.79***	3.04	4.71
Rooms per person in household	**> 2**	55.8	2.3	3.9	**1.00**		
	1.5 to 2		9.3	17.5	**1.08**	0.48	2.45
	1 to 1.5		19.4	31.4	**1.21**	0.55	2.65
	0.5 to 1		51.4	42.0	**2.56**	1.2	5.5
	< 0.5		17.9	5.1	**13.36**	5.51	32.39
Moved house since birth:	**Yes vs. No**	78.8	9.5	5.0	**2.03**	1.43	2.89

B: Variables Considered in Combination

Pooled Odds Ratios and 95% Confidence Limits[1]

		% Data[2]	Percentage Cases	Percentage Controls	Odds Ratio	Lower Limit	Upper Limit
Position usually left and usually found:		81.2					
Usually Left	**Usually found**						
Supine	**Supine or side**		29.2	47.2	**1.00**		
Side	**Supine or side**		35.0	36.1	**1.61***	1.24	2.09
Prone	**Any position**[4]		30.6	12.1	**5.90***	3.93	8.85
Supine	**Prone**		2.0	3.1	**0.98**	0.50	1.92
Prone	**Prone**		3.2	1.5	**3.80**	2.05	7.03
Position last left and last found:		97.1					
Last left	**Last found**						
Supine	**Supine or side**		18.6	48.4	**1.00**		
Side	**Supine or side**		23.0	35.7	**1.63***	1.29	2.07
Prone	**Any position**[5]		38.7	11.3	**13.54***	8.95	20.50
Supine	**Prone**		7.8	2.8	**7.55***	4.53	12.57
Side	**Prone**		11.9	1.8	**20.77***	16.60	25.98
Mother smoked during pregnancy, and bed shared on last occasion:		98.5					
Mother smoked	**Shared Bed**						
No	**No**		34.6	67.8	**1.0**		
No	**Yes**		4.5	5.8	**1.70***	1.06	2.74
Yes: <10 per day	**No**		18.5	13.7	**2.86***	2.22	3.69
Yes: 10+ per day	**No**		27.0	10.3	**5.72**	4.50	7.28
Yes (<10 & 10+)	**Yes**		15.4	2.3	**16.02***	9.07	28.28

Continued

Notes:

[1] The odds ratios and confidence limits reported are unconditional estimates adjusted for Centre.

[2] '% Data' is the percentage of data which are not missing, coded "Don't know" or "Not applicable", and therefore available for analysis.

[3] The odds ratios and confidence limits reported for "Sex" were estimated after excluding data from the Nordic and German studies,which were matched additionally for sex.

[4] 82% of cases & 70% of controls were usually found prone. The remainder usually found in other positions.

[5] 95% of cases and 90% of controls were last found prone. The remainder were last found in other positions.

* .Denotes that statistically significant intercentre heterogeneity was detected and that the confiidence limits for these estimates have been adjusted accordingly.