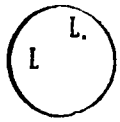
A STUDY OF THE INCUBATION PERIOD, OR AGE AT ONSET, OF THE TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES/PRION DISEASES.

A thesis submitted for the degree of
Doctor of Philosophy.

Marion Joan Anstee Wooldridge
BVetMed MSc(Epidem) MRCVS

London School of Hygiene and Tropical Medicine,
University of London.

April, 1995.



ABSTRACT

In order to model epidemics of infectious diseases, particularly to estimate probable numbers of cases with onset at any particular time, it is necessary to incorporate a term for the incubation period frequency distribution. Sartwell's hypothesis states that the incubation period frequency distribution for infectious disease is generally a log-normal distribution, based on his examination of disease with short incubation periods. However, it may not apply to diseases with long incubation periods. During the course of an epidemic of a disease with a long incubation period, left and right censoring makes direct observation of the frequency distribution highly unreliable; in addition, time of infection is often unknown. Therefore, for a previously undescribed disease, methods other than direct observation must be employed. One method is to extrapolate from information available for other diseases.

In evaluation of Sartwell's hypothesis as applied to diseases with long incubation periods, examination of transfusion-associated AIDS data was inconclusive. Examination of data for experimental transmissible spongiform encephalopathy (TSE)/prion disease in several species suggests that it may not apply. For natural TSE/prion disease, age at onset is used generally as a 'proxy' for incubation period since infection time is rarely known; the validity of this may vary with the disease type and species being examined. Using this measure, again Sartwell's hypothesis was not confirmed.

For both incubation period and age at onset, evidence presented suggests that observed frequency distribution coefficient of skewness is associated with modal age at onset (and thus indirectly with prior age at infection, where appropriate), an earlier modal age at onset resulting in a larger observed coefficient of skewness. The relationship of this association with Sartwell's findings is discussed; they are not incompatible. In addition, an association between observed coefficient of skewness and sample size is demonstrated and the implications discussed.

TABLE OF CONTENTS.

CONTENTS		3
ACKNOWLEDGEMENTS		12
ABBREVIATIONS USED IN THIS THESIS.		13
CHAPTERS		
CHAPTER 1	INCUBATION PERIOD: INTRODUCTION AND LITERATURE REVIEW	15
CHAPTER 2	AIDS: INTRODUCTION: THE INCUBATION PERIOD	25
CHAPTER 3	TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES (TSE'S)/PRION DISEASE: INTRODUCTION	30
CHAPTER 4	DATA USED AND ANALYSIS METHODS NOT DESCRIBED ELSEWHERE	41
CHAPTER 5	AIDS: DATA ANALYSIS AND DISCUSSION	44
CHAPTER 6	SHEEP: DATA ANALYSIS AND DISCUSSION: Natural and Experimental TSE/Prion disease.	51
CHAPTER 7	GOATS: DATA ANALYSIS AND DISCUSSION: Natural and Experimental TSE/Prion disease.	71
CHAPTER 8	PRIMATES: DATA ANALYSIS AND DISCUSSION: Experimental TSE/Prion disease.	86
CHAPTER 9	RODENTS: DATA ANALYSIS AND DISCUSSION: Experimental TSE/Prion disease.	98
CHAPTER 10	HUMANS: DATA ANALYSIS AND DISCUSSION: Natural and Iatrogenic TSE/Prion disease.	117
CHAPTER 11	CATTLE: DATA ANALYSIS AND DISCUSSION: BSE	135
CHAPTER 12	ASSOCIATION BETWEEN SAMPLE SIZE, STANDARD DEVIATION, AND COEFFICIENT OF SKEWNESS FOR A LOG-NORMAL FREQUENCY DISTRIBUTION.	148
CHAPTER 13	OVERVIEW OF MAJOR FINDINGS AND CONCLUSIONS.	154

APPENDICES

APPENDIX 1	EXAMPLES OF HISTOGRAMS OF INCUBATION PERIOD FREQUENCY DISTRIBUTIONS, PLUS OTHER DETAILS, FOR DISEASES WITH SHORT INCUBATION PERIODS AS DISCUSSED IN CHAPTER 1.	161
APPENDIX 2	DETAILS OF SARTWELL'S METHOD FOR THE EXAMINATION OF THE FREQUENCY DISTRIBUTION OF AN INCUBATION PERIOD.	165
APPENDIX 3	EXAMPLES FROM SARTWELL'S TABLE OF FREQUENCY DISTRIBUTION RESULTS FOR OUTBREAKS IN WHICH HE FOUND A 'LOGARITHMIC NORMAL' DISTRIBUTION.	166
APPENDIX 4	PRACTICAL DEMONSTRATION OF SARTWELL'S METHOD USING TWO SETS OF OUTBREAK DATA ALSO EXAMINED BY SARTWELL.	167
APPENDIX 5	REVISED CASE DEFINITION OF AIDS (CDC, 1987).	169
APPENDIX 6	A BRIEF OUTLINE OF THE BACK-CALCULATION METHOD USED (AFTER BROOKMEYER AND GAIL, 1988) TO OVERCOME THE PROBLEM OF LEFT AND RIGHT CENSORING IN AIDS MODELLING.	170
APPENDIX 7	DIAGRAMMATIC REPRESENTATION OF THE PRINCIPLES OF BACK-CALCULATION AND COMPARTMENTAL MODELS ILLUSTRATED FOR ONE INFECTION TIME POINT.	171
APPENDIX 8	EXAMPLES OF THE ESTIMATES OF THE INCUBATION PERIOD FOR AIDS.	172
APPENDIX 9	INFORMATION SUPPLIED WITH THE TA-AIDS DATASET DISCS.	173
APPENDIX 10	DETAILS OF DATA USED FOR TSE/PRION DISEASE EXAMINATION, INCLUDING SOURCE, VARIABLES, AND EXPERIMENTAL METHOD (WHERE APPLICABLE).	179
APPENDIX 11	FILIBEN'S METHOD FOR FREQUENCY DISTRIBUTION ANALYSIS FOR COMPATIBILITY WITH NORMALITY AND LOG-NORMALITY USING MINITAB STATISTICAL SOFTWARE.	196
APPENDIX 12	THE METHOD DESCRIBED BY SNEDECOR AND COCHRAN (1967) FOR ESTIMATION OF FREQUENCY DISTRIBUTION COEFFICIENT OF SKEWNESS USING MINITAB STATISTICAL SOFTWARE.	198

APPENDIX 13	SHEPPARD'S METHOD FOR THE CORRECTION OF THE ESTIMATE OF THE STANDARD DEVIATION FOR FREQUENCY DISTRIBUTIONS WITH GROUP INTERVALS.	200
APPENDIX 14	DISCRIMINATORY POWER OF THE TESTS USED TO ESTIMATE COMPATIBILITY WITH NORMALITY AND LOG-NORMALITY.	201
APPENDIX 15	TESTS USED IN THIS THESIS, APPLIED TO DISEASES SIMILAR TO THOSE EXAMINED BY SARTWELL, PLUS ONE WITH A 'PROXY' INCUBATION PERIOD AS USED BY ARMENIAN AND LILIENFELD (1974).	209
APPENDIX 16	CONSTRUCTED LOG-NORMAL DISTRIBUTION FOR SELECTED DATASETS FOR COMPARISON WITH ACTUAL DATA.	213
APPENDIX 17	SELECTION OF DATA FOR ESTIMATING THE INCUBATION PERIOD AND ITS FREQUENCY DISTRIBUTION FOR TA-AIDS.	216
APPENDIX 18	METHOD OF TRANSFUSION-WEIGHTING FOR ESTIMATING INCUBATION PERIOD IN TA-AIDS IN MULTIPLE-TRANSFUSION CASES.	218
APPENDIX 19	BACK-CALCULATION METHOD TO ESTIMATE THE MEAN INCUBATION PERIOD OF TA-AIDS.	220
APPENDIX 20	METHOD FOR ESTIMATING POTENTIAL SCRAPIE CASE NUMBERS ADJUSTED FOR CULLING OF SHEEP IN SCRAPIE INFECTED FLOCKS FROM DATA COLLECTED BY WOOLDRIDGE (1991).	222
APPENDIX 21	PRIMATE TAXONOMY RELEVANT TO DATA CLASSIFICATION USED IN THIS THESIS.	223
APPENDIX 22	PRIMATES: SELECTED NATURAL LIFE HISTORY MEASURES BY GENUS, AND COMPARISON WITH DATA USED IN ANALYSES.	224
APPENDIX 23	METHOD OF ESTIMATION OF DATA OVERLAP FOR HUMAN TSE/PRION DISEASE CASES WITH A GENETIC MUTATION, FROM THE TWO DATA SOURCES NIH AND GLS.	225
APPENDIX 24	METHOD OF ADJUSTING DATASET FOR THE EFFECT OF CULLING FOR BSE IN CATTLE AND AS APPLIED TO CATTLE GROUP 3.	226
APPENDIX 25	BACK-CALCULATION METHOD TO ESTIMATE THE MEAN INCUBATION PERIOD OF BSE.	228

APPENDIX 26	**CALCULATION OF THE THEORETICAL COEFFICIENT OF SKEWNESS OF A LOG-NORMAL FREQUENCY DISTRIBUTION USING EXCEL STATISTICAL SOFTWARE.**	**230**
APPENDIX 27	**METHOD FOR SIMULATION OF SAMPLES FOR ESTIMATION OF COEFFICIENT OF SKEWNESS.**	**233**
REFERENCES.		**235**

LIST OF TABLES

Table 3.1	TSE/prion disease spectrum.
Table 5.1.1	TA-AIDS; estimates of i/p for three datasets.
Table 5.1.2	TA-AIDS; the effect of left and right censoring on the i/p.
Table 5.1.3	TA-AIDS; estimates of i/p values and related data from back calculation.
Table 5.1.4	TA-AIDS; a/o f/d analyses by subgroup.
Table 6.1.1(i)	Natural sheep scrapie (Parry, 1962); a/o data.
Table 6.1.1(ii)	Natural sheep scrapie (Parry, 1962); f/d analysis.
Table 6.1.2	Natural sheep scrapie (Wooldridge, 1991); f/d analysis.
Table 6.1.3	Natural sheep scrapie (VIDA); f/d analysis.
Table 6.1.4.1	Natural sheep scrapie (GLS); f/d analysis.
Table 6.1.4.2(i)	Natural sheep scrapie (GLS); a/o data by exposure status of dam.
Table 6.1.4.2(ii)	Natural sheep scrapie (GLS); a/o f/d by exposure status of dam.
Table 6.1.4.3	Natural sheep scrapie (GLS); 'maximum i/p' f/d analysis.
Table 6.2.1.1	Experimental TSE in sheep; full dataset; i/p f/d analysis.
Table 6.2.1.2	Experimental TSE in sheep; full dataset; ANOVA results.
Table 6.2.2.1	Experimental scrapie in sheep; subgroup 1; i/p f/d analysis.
Table 6.2.2.2	Experimental scrapie in sheep; subgroup 1; i/p data by passage status.
Table 6.2.3.1	Experimental scrapie in sheep; subgroup 2 (2 mls of extract); i/p f/d analysis.
Table 6.2.3.2	Experimental scrapie in sheep; Subgroup 2; i/p data by volume of s/c inoculation.
Table 6.3.1.1(i)	Natural sheep scrapie; summary of f/d analyses.
Table 6.3.1.1(ii)	Natural sheep scrapie; descriptions of f/d.
Table 6.3.1.2(i)	Experimental TSE in sheep; summary of f/d analyses.
Table 6.3.1.1(ii)	Experimental TSE in sheep; descriptions of f/d.
Table 7.1.1.1	Natural goat scrapie (GLS); f/d analysis.
Table 7.1.2.2(i)	Natural goat scrapie (GLS); a/o data by contact status.
Table 7.1.1.2(ii)	Natural goat scrapie (GLS); a/o f/d by contact status.
Table 7.1.2	Natural goat scrapie (VIDA); a/o f/d analysis.
Table 7.2.1.1	Experimental TSE in goats; full dataset; i/p f/d analysis.
Table 7.2.1.2	Experimental TSE in goats; full dataset; ANOVA results.
Table 7.2.2.1	Experimental scrapie in goats; subgroup 1; i/p f/d analysis.
Table 7.2.2.2	Experimental scrapie in goats; subgroup 1; i/p by passage number.
Table 7.2.3	Experimental TSE in goats; subgroup 2; i/p f/d analysis.
Table 7.2.4	Experimental scrapie in goats; subgroup 3; i/p f/d analysis.
Table 7.3.1.1(i)	Natural goat scrapie; summary of f/d analyses.
Table 7.3.1.1(*n*)	Natural goat scrapie; descriptions of f/d.
Table 7.3.1.2(i)	Experimental TSE in goats; summary of f/d analyses.
Table 7.3.1.2(ii)	Experimental TSE in goats; descriptions of f/d.
Table 8.1.1(i)	Experimental TSE in primates; i/p data plus visual estimate of f/d skew, by primate group.
Table 8.1.1(ii)	Experimental TSE in primates; i/p f/d analysis by primate group.
Table 8.2.1(i)	Experimental CJD in primates; subgroup 6; sample size, mean i/p, and visual estimate of i/p f/d skew, by primate group.
Table 8.2.1(ii)	Experimental CJD in primates; subgroup 6; i/p f/d analysis by primate group.
Table 8.3 (i)	Experimental CJD in squirrel monkeys; subgroup 6: i/p data by inoculation route.
Table 8.3 (ii)	Experimental CJD in squirrel monkeys; subgroup 6. i/p f/d analysis by inoculation route.
Table 8.4	Experimental TSE in primates; full dataset; comparison of analysis results with life history measures, stratified by data representative of genus.
Table 8.5.1(i)	Experimental TSE in primates; full dataset; descriptions of f/d.

LIST OF TABLES CONTINUED

Table 8.5.1(ii)	Experimental CJD in primates; subgroup 6; descriptions of f/d.
Table 8.5.1(iii)	Experimental TSE in primates; homogeneity, size and f/s skew comparisons.
Table 9.1.1.1	BSE in mice; full dataset; i/p f/d analysis.
Table 9.1.2.1(i)	BSE in mice: inoculation age and i/p data by BSE group.
Table 9.1.2.1(ii)	BSE in mice: i/p f/d analysis by BSE group.
Table 9.1.2.2	BSE in mice; inoculation age data and f/d c/s analysis; groups A to H
Table 9.2.2	Scrapie in mice; i/p data by scrapie agent strain.
Table 9.2.3(i)	Scrapie in mice; strain 22A: i/p data by 'sinc' genotype.
Table 9.2.3(ii)	Scrapie in mice; strain 22A; i/p f/d analysis by 'sinc' genotype.
Table 9.2.4(i)	Scrapie in mice; strain ME7; i/p data by 'sinc' genotype.
Table 9.2.4(ii)	Scrapie in mice; strain ME7; i/p f/d analysis by 'sinc' genotype .
Table 9.2.5	Scrapie in mice; analysis of strain & 'sinc' genotype by ANOVA.
Table 9.3.1	CJD in guinea-pigs; full dataset; f/d analysis.
Table 9.3.2(i)	CJD in guinea-pigs: i/p data by passage number.
Table 9.3.2(ii)	CJD in guinea-pigs; effect of passage number on i/p.
Table 9.3.3	CJD in guinea-pigs, f/d analyses for passage numbers 2-4.
Table 9.4.1.1(i)	BSE in mice; summary of f/d analyses.
Table 9.4.1.1(ii)	BSE in mice; descriptions of f/d.
Table 9.4.2.1(i)	Scrapie in mice; summary of f/d analyses.
Table 9.4.2.1(ii)	Scrapie in mice; descriptions of f/d.
Table 9.4.3.1	CJD in guinea-pigs; summary of f/d analyses.
Table 9.4.3.1(ii)	CJD in guinea-pigs; descriptions of f/d.
Table 10.1	Sporadic TSE in humans; a/o f/d analysis by data source.
Table 10.2.1(i)	Familial TSE in humans; a/o data by data source.
Table 10.2.1(ii)	Familial TSE in humans; a/o f/d analysis by data source.
Table 10.2.2.1	TSE in humans: a/o data by PrP gene mutation.
Table 10.2.2.2	TSE in humans; a/o f/d analysis by PrP gene mutation category.
Table 10.2.2.4	TSEs in humans; effect of met/val heterogeneity on mean a/o.
Table 10.3.1	Kuru in humans; a/o f/d analysis.
Table 10.3.2.1	Pituitary injection cases in humans; first set; i/p f/d analysis.
Table 10.3.2.2	Pituitary injection cases in humans; French set; i/p f/d analysis.
Table 10.3.3	Dura graft cases in humans; i/p and a/o data.
Table 10.4.1.1(i)	Sporadic and familial TSE in humans; summary of a/o f/d analyses.
Table 10.4.1.1(ii)	Sporadic and familial TSE's in humans; mean a/o, and f/d descriptions.
Table 10.4.1.2(i)	Recognised transmission route cases in humans; summary of a/o f/d analyses.
Table 10.4.1.2(ii)	Recognised transmission route cases in humans; a/o f/d descriptions.
Table 10.4.1.4(i)	TSE with known PrP gene mutation: summary of a/o f/d analyses.
Table 10.4.1.4(ii)	TSE's with known PrP gene mutation: modal and mean a/o, and f/d description.
Table 11.1.1(i)	BSE in cattle; a/o data by group.
Table 11.1.1(ii)	BSE in cattle; a/o f/d analysis by group.
Table 11.1.2(i)	BSE in cattle; group 3Y; a/o data by calving cohort.
Table 11.1.2(ii)	BSE in cattle; group 3Y; a/o f/d analysis by calving cohort.
Table 11.1.3	BSE in cattle; groups 1 and 3Y; comparison of minimum and mean a/o's per cohort.
Table 11.2.2(i)	BSE in cattle: group 3/87: a/o data by feedflag status.
Table 11.2.2(ii)	BSE in cattle: group 3 87: a/o f/d by feedflag status.

LIST OF TABLES CONTINUED

Table 11.4	BSE in cattle; group 3; estimates of a/o data from back calculation: unadjusted and (adjusted) for culling.
Table 11.5.1.1(i)	BSE in cattle; summary of a/o f/d analyses.
Table 11.5.1.1(ii)	BSE in cattle; a/o f/d descriptions.
Table 11.5.1.2	BSE in cattle; group 3Y; f/d skewness description by cohort.
Table 12.2(i)	Comparison of observed and theoretical c/s by sample size and SD; 10 observations per combination.
Table 12.2(ii)	Comparison of observed and theoretical c/s by sample size and SD; 5000 observations per combination.
Table A3	Examples from Sartwell's table of frequency distribution results.
Table A4	Two sets of outbreak data examined by Sartwell.
Table A8	Examples of estimates of the incubation period for AIDS.
Table A10.1	Data references; natural scrapie in sheep.
Table A10.2	Data references; experimental TSE in sheep.
Table A10.3	Data references; natural scrapie in goats.
Table A10.4	Data references; experimental TSE in goats.
Table A10.5	Data references and personal communications; experimental TSE in primates.
Table A10.6	Experimental TSE in primates; some major variables by primate group.
Table A10.7	Data references; TSE in humans; GLS.
Table A10.8	TSE in humans; codon substitutions recorded in data used, and resultant amino acid change.
Table A14.1	Correlation coefficient critical value (at P = 0.05) by sample size.
Table A17.1	Time to disease onset: Comparison of 'Cleaned' data for 1st, 2nd and 3rd transfusion with estimates where unknown month (coded 99) has been replaced by month 07 (i.e. mid-year), for transfusion, diagnosis, and report dates.
Table A20.1	Natural sheep scrapie; culling pattern estimate (Wooldridge, 1991).
Table A22.1	Primates; comparison of recorded maximum longevity with mean and maximum i/p (and as a percentage) for full dataset by genera.
Table A22.2	Primates; comparison of adult weight with certain i/p measures, for full dataset by genera.
Table A24.1	BSE in cattle: culling correction, and method as applied to group 3.

LIST OF FIGURES

Figure 5.1.2	TA-AIDS; mean i/p by year of transfusion.
Figure 5.1.4	TA-AIDS; subgroups A and B; i/p f/d.
Figure 6.1.1	Natural sheep scrapie; Parry (1962) data; a/o f/d by sex.
Figure 6.1.3	Natural sheep scrapie; VIDA data; a/o f/d.
Figure 6.1.4.1	Natural sheep scrapie; GLS data; a/o f/d.
Figure 6.1.4.3(a)	Natural sheep scrapie; GLS data; 'maximum i/p' f/d.
Figure 6.1.4.3(b)	Natural sheep scrapie; GLS data; 'maximum i/p' by age at 1st recorded exposure.
Figure 6.2.1.1	Experimental TSE in sheep; full dataset; i/p f/d.
Figure 6.2.2.1	Experimental scrapie in sheep; subgroup 1; i/p f/d.
Figure 6.2.2.2	Experimental scrapie in sheep; subgroup 1; i/p by age at infection.
Figure 6 2.3.1	Experimental scrapie in sheep; subgroup 2 (2 mls of extract); i/p f/d.
Figure 6 2.3.2	Experimental scrapie in sheep, subgroup 2; i/p by dose.
Figure 6 3.1.2	Experimental scrapie in sheep; CNS and s/c inoculation, superimposition of modal points.
Figure 6.3.3	Natural sheep scrapie: a/o f/d as percentage of cases by age, corrected for culling.
Figure 7.1.1.1	Natural goat scrapie; GLS data; a/o f/d.
Figure 7.1.1.2	Natural goat scrapie; GLS data; a/o f/d by contact status recorded.
Figure 7.1.1.3	Natural goat scrapie; GLS data; a/o by maximum age of exposure.

LIST OF FIGURES CONTINUED

Figure 7.2.1.1	Experimental TSE in goats; full dataset; i/p f/d.
Figure 7.2.2.2(a)	Experimental scrapie in goats; subgroup 1; i/p by age at infection.
Figure 7.2.2.2(b)	Experimental scrapie in goats; subgroup 1; i/p by dose.
Figure 7.3.1.3	Goat TSE; comparison of natural a/o and experimental (s/c, i/d and oral) i/p f/d; modes overlayed.
Figure 8.1.1	Experimental TSE in primates; i/p f/d by primate group.
Figure 8.1.2	Experimental TSE in primates; c/s by dataset size.
Figure 8.4	Experimental TSE in primates; comparison of certain i/p measures with mid-point of adult weight range, by primate group.
Figure 9.1.1.1	BSE in mice; full dataset; i/p f/d.
Figure 9.1.1.2	BSE in mice; full dataset; i/p by inoculation age.
Figure 9.1.2.2	BSE in mice; c/s by age at infection for groups A to H.
Figure 9.2.2	Scrapie in mice; full dataset; i/p f/d by agent strain.
Figure 9.2.3	Scrapie in mice; strain 22A; i/p f/d by 'sinc' genotype.
Figure 9.2.4	Scrapie in mice; strain ME7; i/p f/d by 'sinc' genotype.
Figure 9.2.5	Scrapie in mice; mean i/p by 'sinc' genotype and strain.
Figure 9.2.6	Scrapie in mice; association between f/d c/s and modal i/p.
Figure 9.3.1	CJD in guinea-pigs; full dataset; i/p f/d.
Figure 10.1	Sporadic CJD in humans; UK dataset; a/o f/d.
Figure 10.2.1	Familial TSE in humans; a/o f/d by data source.
Figure 10.2.2.1	Familial TSE in humans; mean a/o by PrP gene mutation and data source.
Figure 10.2.2.2	TSE in humans; a/o f/d by PrP gene mutation.
Figure 10.2.2.3	TSE in humans; f/d c/s by mean a/o and PrP gene mutation.
Figure 10.3.1	Kuru in humans; a/o f/d.
Figure 10.3.2.2	Pituitary injection cases in humans; French dataset; i/p f/d assuming infection in 1985, plus f/d of time from first treatment.
Figure 10.4.1.1	Sporadic and non-sporadic TSE in humans; NIH data; a/o f/d; stacked composite.
Figure 10.4.1.4	TSE in humans; f/d skewness and modal a/o; a comparison of three PrP gene mutations.
Figure 11.1.1	BSE in cattle; group 3; a/o f/d.
Figure 11.1.2	BSE in cattle; group 3Y: effect of censoring on a/o range, by calving cohort.
Figure 11.2.1	BSE in cattle; a/o by a/p.
Figure 11.2.2	BSE in cattle; group 3/87; a/o f/d by feedflag status.
Figure 11.3	BSE in cattle; group 3; effect of culling correction on a/o f/d.
Figure 12.2	Association between sample size and c/s for different SD values: log-normal distribution.
Figure 12.3	Coefficient of skewness frequency distribution for 5000 random-generations; SD of theoretical distribution = 1.5.
Figure A1.1	Frequency distribution: Serial interval for measles as an estimation of i/p (after Stillerman & Thalheimer, 1944).
Figure A1.2	Frequency distributions for incubation periods for 5 outbreaks of 'Typhoid Fever' reviewed by Miner (1922).
Figure A1.3	Details of i/p for 'Serum Jaundice'; Camp Polk series (after Parr, 1945)
Figure A1.4	Frequency distributions: Reported cases of GBS by period between vaccination and onset in 1-week periods and 2-day periods (after Schonberger et al, 1979).
Figure A1.5	Frequency distribution: Experimental gonorrhoea using human volunteers; i/p f/d (after Mahoney et al, 1946).
Figure A4.1	Sartwell's method plot for measles; data from Stillerman & Thalheimer, 1944.
Figure A4.2	Sartwell's method plot for 'typhoid fever'; data from Sawyer, 1914.
Figure A7	Diagrammatic representation of the principles of back-calculation and compartmental models illustrated for one infection time point.
Figure A11	Example of the normal probability plot produced by Minitab.
Figure A14.1	Composition of normal and log-normal distributions.

LIST OF FIGURES CONTINUED

Figure A14 2	The normal distribution Comparison by sample size.
Figure A14.3	The log-normal distribution: Comparison by sample size.
Figure A14.4	N=50; f/d; normality and non-normality.
Figure A14.5	N=50; f/d; log-normality and non-log-normality.
Figure A14.6	N=50; f/d; compatibility with both normality and log-normality.
Figure A14.7	N=5; f/d; symmetry, normality and log-normality.
Figure A14.8	N=5; f/d; asymmetry, normality and log-normality.
Figure A14.9	N=500; f/d; normality and non-normality.
Figure A14.10	N=500; f/d; log-normality (i).
Figure A14.11	N=500; f/d; log-normality (ii).
Figure A16.1	Log-normal construct for natural scrapie in sheep; GLS full dataset.
Figure A16.2	Log-normal construct for experimental TSE in sheep; full dataset.
Figure A16.3	Log-normal construct for scrapie in mice; strain 22A in 'sinc' s7p7 mice.
Figure A16.4	Log-normal construct for TSE in humans (i); sporadic CJD, UK dataset.
Figure A16.5	Log-normal construct for TSE in humans (ii); comparison of 3 datasets.
Figure A18.1	TA-AIDS; Single transfusion cases by proportion of transfusions per year.
Figure A21.1	Simplified taxonomy of the primates used in these experiments.
Figure A23.1	TSE in humans; codon 198 mutation; comparison of f/d for NIH and GLS data.

ACKNOWLEDGEMENTS

I wish to thank Mr Simon Cousens for his supervision, encouragement, and statistical advice throughout the course of this project.

I am grateful to all the staff of the Epidemiology Department, Central Veterinary Laboratory, Weybridge, for their support, but especially to Mr John Wilesmith MRCVS who initiated this project and gave me both time and facilities to complete it; to Mr Robin Sayers for statistical help and advice, and to Ms Sue Kidd and Mr Andy Mitchell for data retrieval from VIDA and the BSE database.

Thanks are also due to the library staff at the CVL, all of whom have been very willing to help with my numerous requests for information.

In addition I wish to thank Professor Paul Fine, Mr Jamie Robinson, Dr Paddy Farringdon and Dr Rosalind Ridley for useful discussions in the early stages of this project.

Datasets for use in the project were kindly supplied by Dr Hugh Fraser and Dr Nora Hunter, both of the Neuropathology Unit, Edinburgh; Dr T Peterman, Centers for Disease Control, Atlanta, Georgia, USA; Dr R Will, UK CJD Surveillance Unit, Edinburgh, and Mr Michael Dawson MRCVS of the CVL.

In particular I thank Dr Paul Brown, National Institutes of Health, Bethesda, Maryland, USA, for the valuable time I was able to spend in his laboratory, for his generosity in giving me access to his extensive data collection, and for replying enthusiastically to all my faxes.

Finally, I would like to thank Dr M.Y.A. Gordon for many, many helpful discussions, and Dr John Morris for encouragement along the way.

ABBREVIATIONS USED IN THIS THESIS

adj.	adjusted
AIDS	acquired immune deficiency syndrome
ANOVA	analysis of variance
a/o	age at onset
a/p	age at purchase
APD	atypical prion disease
BSE	bovine spongiform encephalopathy
CC	correlation coefficient - specifically indicates result of f/d analysis
CDC	Centers for Disease Control, Atlanta, Georgia, USA
CI	confidence interval
CJD	Creutzfeldt-Jakob disease
Cn	codon
CNS	central nervous system
c/s	coefficient of skewness
C/v	critical value for CC (at P=0.05 level unless otherwise stated)
CVL	Central Veterinary Laboratory, Addlestone, Surrey
CWD	chronic wasting disease of mule deer and elk
DNA	deoxyribonucleic acid
exp	exponential
F	value calculated in a variance-ratio test for comparison with F distribution
f/d	frequency distribution
FFI	fatal familial insomnia
FSE	feline spongiform encephalopathy
GBS	Guillain-Barre syndrome
GLS	general literature search
GSS	Gerstmann-Straussler-Sheinker syndrome
HIV	human immunodeficiency virus
i/d	intra-dermal
i/p	incubation period
kg	kilogram
L	left skewed
Lit.	indicates data from GLS
MAFF	Ministry of Agriculture, Fisheries and Food
max	maximum
MBM	meat and bone meal
met	methionine

ml	millilitre
N	number (i.e. sample size)
NAG	National Algorithm Group Ltd
nc	not able to be calculated (due to small sample size)
NIH	National Institutes of Health, Bethesda, Maryland, USA.
NPU	Neuropathology Unit, Edinburgh.
NS	not significant (i.e. not statistically significant; $P > 0.05$)
OIE	Office Internationale des Epizooties
P	probability
p.c.	personal communication
Pg	passage
PrP	prion protein
PrP^{C}	normal cellular form of prion protein
PrP^{Sc}; PrP^{CJD} etc.	abnormal isoform of prion protein associated with scrapie, CJD etc.
r	correlation coefficient - other than those calculated in f/d analysis
R	right skewed
RNA	ribonucleic acid
SAF	scrapie-associated fibril
s/c	sub-cutaneous
SD	standard deviation
Sinc	'scrapie incubation period'; a gene in mice
Sip	'scrapie incubation period'; a gene in sheep
TA-AIDS	transfusion-associated (or transfusion-acquired) AIDS
TME	transmissible mink encephalopathy
TSE	transmissible spongiform encephalopathy
tx	transfusion or treatment
UK	United Kingdom
u/k	unknown
val	valine
VIDA	veterinary investigation diagnosis analysis (a database)

CHAPTER 1

INCUBATION PERIOD: INTRODUCTION AND LITERATURE REVIEW

1.1 The incubation period of a disease: Background and early work.

1.1.1 The definition of incubation period and related measures for an infectious disease.
1.1.2 Examination and uses of incubation periods for acute infectious diseases: Field outbreaks.
1.1.3 Examination of incubation periods in acute infectious disease: Experimental infection.
1.1.4 Patterns created by incubation period for acute infectious diseases.
1.1.5 Examination of incubation periods by Sartwell's method.

1.2 Examination of non-infectious, neoplastic and other diseases by use of 'proxy' incubation periods.

1.2.1 Definitions of 'proxy' incubation periods.
1.2.2 Sartwell's method applied to 'proxy' incubation periods.

1.3 Infectious diseases with long incubation periods.

1.3.1 Definition of a 'long' incubation period.
1.3.2 Problems inherent in the investigation of incubation period in diseases with long incubation periods.

1.4 Aims of this study.

CHAPTER 1

INCUBATION PERIOD: INTRODUCTION AND LITERATURE REVIEW

1.1 The incubation period of a disease: Background and early work.

1.1.1 The definition of incubation period and related measures for an infectious disease.

The incubation period (i/p) of a disease is generally defined as the period of time from infection with the disease-causing agent to onset of clinical signs (Benenson 1985; Thrusfield, 1986).

The latent time or latent period is the time from infection to infectiousness (Becker, 1989; Anderson & May, 1991); it will only equal i/p if infectiousness and clinical signs occur simultaneously. Latent period is, however, used by some authors to describe i/p (Meynell & Meynell, 1958; Kondo, 1977).

Serial interval, or generation time, is time from a particular stage of disease in the infecting case to the same stage in the infected case (Rodrigues, 1991); in measles, rash-to-rash is often used; for statistical purposes it is the best measure of i/p obtainable (Stocks, 1931). It is equivalent to latent period plus time from the start of infectiousness to transmission; duration is therefore partly dependent upon contact rate with susceptible individuals. Anderson & May (1991) define the generation time as the sum of average latent and infectious periods.

1.1.2 Examination and uses of incubation periods for acute infectious diseases: Field outbreaks.

With acute diseases with short i/p's, for example 'food poisoning', or many viral diseases of childhood, the link between infection source and disease onset was originally noted precisely because i/p is short; for example virtually all those affected remembered attending the same function. Investigations of such outbreaks or epidemics, either to find infection source, or to learn more on behaviour of epidemics, allowed description of i/p ranges for specific clinical presentations, though an infectious agent may not have been identified.

Examples include an investigation into an outbreak described as typhoid fever (Sawyer, 1914), and measles cases occurring in the same household (Stocks and Karn, 1928) using serial interval (rash-to-rash).

Various explanations were advanced when the same disease in different host groups produced

differing results. Stillerman and Thalheimer (1944), also looking at successive cases of measles in the same household, report a serial interval range of 8 to 19 days. They note a tendency for younger children to experience longer intervals, particularly those aged from 12 to 23 months, and suggest this may be connected with 'velocity of infection' (figure A1.1); but for young children contact and infection may be delayed. Benenson (1985) gives measles i/p as 8-13 days, but occasionally longer or shorter.

With microbiological development, it became apparent that sometimes clinical presentation of disease caused by different organisms was similar, resulting in refinement of i/p frequency distribution (f/d) estimates for some infections. The separate causal organisms of typhoid and paratyphoid were first described in 1896 (Achard and Bensaude), but were still treated by Miner (1922) as one disease, in his review of five outbreaks of 'typhoid fever' in which infection time was known (figure A1.2). He reported that i/p f/d for 'typhoid fever' appeared to vary with outbreak, concluding that it is highly variable from one outbreak to another, possibly depending on "virulence of infection". However, for outbreaks 1 and 2, foodborne as opposed to water-borne, the i/p resembles that of paratyphoid fever (given as 1-10 days by Benenson, 1985), more usually associated with foodborne outbreaks than with contaminated water (Huckstep, 1962).

Once the stability of many i/p's became documented, they were routinely used to trace outbreak sources. Since i/p was short, from hours (salmonellosis; Benenson, 1985) to months (serum hepatitis; average 2-3 months, Benenson, 1985), recall was comparatively accurate; this often provided further refined i/p data.

Parr (1945) reported an incident involving contaminated yellow fever vaccine used for US Army personnel, giving i/p details for serum jaundice (probably viral hepatitis type B). A standard vaccine dose was given to many young healthy adults, mostly male, experiencing similar lifestyle. Vaccination date and place was known, follow-up was good, and large vaccinee groups were sent to different post-vaccination destinations, for example 'Camp Polk' (1004 cases) (figure A1.3). Parr concluded that not enough importance had previously been attached to host and environment in determining attack rate, i/p length, and disease severity, and that the common practice of smoothing time-based data results in information loss.

Taylor et al (1974) report an outbreak of 132 cases of typhoid fever in Trinidad in 1971. Taking typhoid fever i/p as approximately two weeks, they identified the source, contaminated ice-cream, then calculated the actual i/p; mean, 19.25 days; median, 19 days. Most cases were children. Huckstep (1962) stated that i/p in children is generally shorter than the 10-14 days of adults. However, the authors cite Hornick et al (1966) as showing that i/p varies inversely with number of ingested typhoid organisms, conclude that this outbreak was probably caused by low level contamination, and estimate contamination in organisms per gram of ice-cream.

Increasing reports of Guillain-Barre syndrome (GBS) after influenza vaccination was noticed during a United States nationwide vaccination programme (Schonberger et al; 1979). Surveillance revealed 1098 cases, a rapid rise in recent vaccinees strongly suggesting an association. Based on attack rates, the relative risk of GBS in those vaccinated compared with those unvaccinated was greater than 1 ($P < .05$) for ten weeks post-vaccination, peaking at over 12 in weeks 2-3; the authors conclude that many cases were vaccination-related, and comment on the non-randomness of the i/p f/d, with the mode, 10%, on the 16-17th days post-vaccination (figure A1.4).

Differences in i/p ranges for a given disease were specifically examined in a review of children hospitalised for measles in Copenhagen from 1915 to 1925 (Aaby, 1991). Using rash-to-rash to estimate i/p, fatal secondary cases had a shorter mean serial interval (9.3 days; SD 2.0) than survivors (11.9 days; SD 2.6) ($P <0.05$). Since the prodromal period is also prolonged in severe cases, the author warns that using the rash-to-rash interval, in severe cases, will lead to an overestimation of i/p length.

1.1.3 Examination of incubation periods in acute infectious diseases: Experimental infection.

Different infective doses, ages at infection, and other variables were suggested by several workers as likely to affect i/p. In estimating i/p in natural disease, in humans and animals, often the time of infection is unknown. Experimental infection can overcome some of these problems. In an early example, Topley (1919) describes a series of experimental infections of rodents with an organism called *Bacillus danysz*, undertaken to investigate the phenomenon of recurrent epidemics, separated by considerable periods of time.

Unfortunately, although i/p information in experimental studies is usually obtainable, it is often reported incompletely, if at all. This omission occurs in a study in human volunteers, of *Vibrio cholerae* (Music et al, 1970). Although it is clear from experimental design that i/p data could have been obtained, it is only reported for 2 of 27 cases; conversely, attack rates and disease severity are reported in detail. In contrast, a study describing a canine model of *Vibrio cholerae* (Sack & Carpenter, 1968) details i/p for 21 dogs; and for all 90 dogs in the study, some i/p information was given, being less than 18 hours in 96% of the dogs.

Mahoney et al (1946) report a study in human volunteers, using experimental infection with *Neisseria gonorrhoeae*, which causes gonorrhoea, giving detailed results from 30 separate transmission experiments, most using several men. Although i/p was generally given, it was

not specifically considered further by the authors, although strains of organism, techniques of infection and attack rate were examined in detail. The range recorded was from less than 24 hours to 31 days; mode 3 days (21 of 76 cases) (figure A1.5). Benenson (1985) gives 2-7 days for natural disease but some infection techniques used were unlikely to have simulated natural transmission.

Experimental work was used to investigate biological changes during the apparently 'silent' i/p, serial sacrifice and post-mortem examination being undertaken before onset of clinical signs. Extensive investigation into pathological mechanisms in bacterial infections using this method with mice was undertaken by Orskov and collaborators (e.g. Orskov, 1932; reported in detail by Masden, 1937). Sterile autopsies allowed organ and tissue culture to follow bacterial routes through the body during i/p. Similarly for viral pathology; using infectious ectromelia of mice as a model for smallpox, both the epidemiology, including i/p f/d details under differing circumstances, and the pathogenesis during i/p was described (Fenner, 1948a; 1948b). Clearly, biological events were occurring during i/p, for both bacteriological and viral diseases.

Hornick and Woodward (1966) investigated dose effect, both on i/p and disease severity, in an investigation of typhoid vaccine efficacy in experimentally infected humans. Individual i/p's are not given but the authors state that they were inversely related to size of infecting dose, range being 3-26 days. Further evidence for a dose effect on i/p is reported (Wear et al, 1968); intracerebral inoculation of measles virus into mice showed association between high dose and short i/p. Similarly, vaccination of humans with attenuated measles virus resulted in i/p's from 7 to 14 days; duration varied inversely with numbers of virus particles injected (Aaby, 1985).

1.1.4 Patterns created by incubation periods for acute infectious diseases.

The earliest recorded use of modelling in the context of dynamics and mathematical properties of infectious disease concerns smallpox (Bernoulli, 1760). The mathematical theory of epidemics, including the phenomenon of epidemic periodicity was specifically considered by, for example, Hamer (1906) and Ross (1915, 1916, 1917).

Studying patterns created both by disease outbreaks, and 'waves' of disease within more extensive epidemics resulted in the description of serial interval, and differentiation between primary and secondary cases, for example using measles (Stocks & Karn, 1928). Hope-Simpson's (1952) derivation of "susceptible exposure attack rate" distinguishes between several secondary cases all with onset within a single i/p, and the same number of cases arranged with onset in 'waves' (i.e. as a series of secondary, tertiary, etc. cases), within successive i/p's. He then assigns to the

two situations high infectiousness and lower infectiousness respectively, on the basis of these two patterns.

Many mathematical models of epidemics, for example that of Soper (1929), and further derivations incorporating the concept of chance, as in the Reed-Frost model (Abbey, 1952), rely on a constant known i/p to describe the behaviour of the epidemic and to predict expected numbers in successive epidemic waves. From around 1970, the publishing of papers on dynamics and mathematical properties of infectious diseases increased rapidly (Bailey, 1975), generally based on observations of diseases with short i/p's. However, modelling of aspects of tuberculosis (ReVelle, Lynn & Feldmann, 1967) and leprosy (Lechat, 1971), two long i/p diseases, also began, paving the way for work on acquired immunodeficiency syndrome (AIDS).

1.1.5 Examination of incubation periods by Sartwell's method.

Sartwell (1950, 1952, 1966) examined i/p f/d for outbreaks of relatively acute infectious diseases, mainly from published data; the stated purpose was to demonstrate a consistent f/d pattern. Having observed that i/p f/d's usually had some resemblance to a normal curve, he plotted cumulative distributions on normal probability paper, examining results for linearity. The resultant plot was usually more nearly linear if the logarithm of the i/p was plotted; i.e. the data usually more closely resembled a 'log-normal' curve than a normal curve. Using those 'log-i/p' plots where the result was nearly linear, he fitted a straight line by inspection, and calculated the "estimated median" and "dispersion factor" (detailed method; appendix 2).

He found a "logarithmic normal" distribution for 18 datasets treated as above, including bacterial and viral diseases, amoebic dysentery and malaria (examples; appendix 3). Although noting that a minority of datasets were otherwise, irregular, or with greater or lesser skewness, he concluded that f/d for most infectious diseases resembles a "logarithmic normal curve". From examination of "dispersion factors", he also concluded that the "degree of dispersion of incubation periods is proportionate to the usual length of incubation". This method (reproduced; appendix 4), subsequently became known as 'Sartwell's method'.

When data did not fit the 'log-normal' pattern, it was often experimental data with a longer right tail to the distribution (Sartwell, 1966), for example experimental gonorrhoea (Mahoney et al, 1946).

Although the longest i/p Sartwell examined was that for serum hepatitis (median i/p; 105 days), he concluded that most diseases fitted the 'log-normal' pattern whether their i/p was 'short' or 'long'. Although he only examined graphed data by eye, applying no statistical test for log-normality, and noted that this was not the best description for all datasets, 'log-normality'

gradually passed into acceptance as THE i/p f/d for infectious diseases.

1.2 Examination of non-infectious, neoplastic and other diseases by use of 'proxy' incubation periods.

1.2.1 Definitions of 'proxy' incubation periods.

Many workers adopted Sartwell's methods to examine diseases of non-infectious or unknown origin, including neoplasia and drug-induced disease, re-defining i/p as required; these time periods are here called 'proxy' i/p's.

Cobb et al (1959) define malignant disease i/p as "the period of multiplication of an invading strain of abnormal host cells", stating that "when exposure (to the causal agent) is prolonged or continuous, the incubation period is almost always less than the appearance time after first exposure"

Generally, however, for diseases of non-infectious aetiology, 'i/p' is used simply to mean time from exposure (to the assumed causal agent) to clinical disease onset (Armenian & Lilienfeld, 1974). For multiple or continuous exposure, ' i/p' generally denotes total duration of exposure (from first exposure) plus any interval after exposure ceases, until disease onset (Kennaway, 1957, Armenian & Lilienfeld, 1974). These time intervals are also called latent period (Armitage & Doll, 1961; Stuart & Kneale, 1970), induction time (Druckey, 1967) and 'time-to-tumour' (Guess & Hoel, 1977). Studying blood dyscrasia attributed to chloramphenicol, Polak et al (1972) use 'latent period' for the period from last dose to first appearance of signs or symptoms. 'Latent period' is therefore often equated with i/p of an infectious disease.

In experimental work on neoplasia, latent period has also been described as the time from initial exposure (to causal agent) to disease detection (Jones & Grendon, 1975). Since detection, for example by X-ray, may precede clinical signs, use of this definition is significantly different.

In diseases where interval from first clinical signs to death is generally rapid and early signs are non-specific, i/p has been equated with time from exposure (to causal agent) to death, as in lung cancer (Armenian & Lilienfeld (1974)).

For diseases of possible genetic origin, 'i/p' has been equated to age at onset (a/o) of clinical signs; with inherited human prion disease, for example Gerstmann-Straussler-Sheinker syndrome (GSS), Baker et al (1991) have implied this definition. For congenital conditions, 'i/p' has been

equated with the mothers age at the birth under investigation, as in an investigation into causes of twinning (Phillipe, 1990).

1.2.2 Sartwell's method applied to "proxy' incubation periods.

Armenian and Lilienfeld (1974) examine neoplasia using Sartwell's method, to "demonstrate the possible usefulness, and range of application of the 'log-normal' model". Datasets include bronchogenic carcinoma in asbestos workers (Selikoff & Hammond, 1968); thyroid cancer following childhood radiation (Beach & Dolphin, 1962); bladder tumours following occupational exposure (Goldblatt, 1949); several types of leukaemias (e.g. Court Brown & Doll, 1965; Tomonaga, 1962) and Burkitt's lymphoma (e.g. Latunde Odeku & Osuntokun, 1968). They also examine data on pancytopenia after chloramphenicol (Polak et al, 1968).

Their conclusions are similar to Sartwell's; they also note that some datasets do not fit particularly well, for example for Wilm's tumour after intra-uterine x-ray exposure (Stewart & Kneale, 1970). They note that for a dataset collected from several different sources and populations, the finding of log-normality should be viewed with caution, and discuss problems of continuous exposure, and lack of precise exposure time. Nevertheless, they conclude that Sartwell's model may have potential for application to chronic diseases.

Some investigators appear convinced of the probable universal validity of Sartwell's conclusions, applying the hypothesis to totally different conditions. Phillipe (1990) examines the mother's age (the 'i/p') at the birth of twins, to investigate whether the likelihood of having twins is congenital, or an effect of approaching menopause, conclusions being based on how well the resultant f/d's of the mother's age, and of the remaining time to menopause, fitted log-normality. They conclude that unlike-sex (di-zygotic) twins have their "causal origin in the maternal pre-natal period, as the age of onset is definitely log-normal", whereas the situation for like-sex twins (a mixture of mono- and di-zygotic) is less clear-cut.

Other diseases where a/o is considered as i/p, and investigated by Sartwell's method include various hereditary and metabolic diseases (Armenian & Khoury, 1981), giving a log-normal distribution for a/o, and familial and sporadic Alzheimer's disease (Horner, 1987), with similar results.

1.3 Infectious diseases with long incubation periods.

1.3.1 Definition of a 'long' incubation period.

Diseases considered to have a 'long' i/p in this project generally have an i/p measured in

years. However, there are exceptions, particularly for laboratory animal models of disease. Therefore the definition used here is that these diseases usually take one or more years, but occasionally less, to develop in their natural host range, and may often take less than a year in artificially or abnormally acquired infections, or unusual hosts. Inherent biological variation means that definition of a precise minimum period would be impractical.

1.3.2 Problems inherent in the investigation of incubation period in diseases with long incubation periods.

Certain problems are inherent in investigating natural diseases with long i/p's. Infection is generally long before disease onset, so often no connection is made. A disease may go unrecognised as infectious for many years, with enduring aetiological uncertainty for some diseases; for example motor neurone disease (Swash, 1991); multiple sclerosis (Fischman, 1981); grass sickness, a disease of equidae (Pavord & Fisher 1987); some types of neoplasia, for example African Burkitt's lymphoma, nasopharyngeal cancer, and Hodgkin's disease, all of which are associated with higher antibody titres for Epstein-Barr virus (Evans, 1982).

Even when a disease is recognised as infectious, the distant time of infection may be unknown or forgotten, making it impossible to trace infection source, or calculate i/p with accuracy. Sometimes it is possible to demarcate an infection 'time-window'; often a/o is the only information available. Mays & Ricketts (1975) describe a case of melioidosis with "a 26-year interval between probable exposure and confirmed diagnosis"; Doyle (1953) report a leprosy patient in Great Britain with a presumed i/p of between 8 and 13 years, exposed whilst on service abroad.

Many diseases with long i/p's are insidious in development, with non-specific signs such as weight loss, fatigue, a mild cough, or memory impairment; it may be impossible to accurately record the time of disease onset. In some diseases, particularly tuberculosis, routine or post-contact screening for subclinical infection, followed by treatment for humans (e.g. after contact with infected elk; Fanning & Edwards, 1991), or culling for livestock (e.g. dairy cattle infected from a farmer; Englert & Milbradt, 1977) eliminates clinical disease onset, necessitating amendment of i/p definition.

With a high disease prevalence, contact with potential infection may occur regularly, making it impossible to evaluate which contact was responsible. Screening, as in the British BCG vaccination programme (Sutherland & Springett, 1987), may indicate infection status at a particular time, and thus indicate minimum or maximum i/p if disease subsequently occurs. Nevertheless, accurate i/p data collection remains difficult; Irgens (1985) describes some of the problems in a study of leprosy.

1.4 Aims of this study.

The emergence of acquired immune deficiency syndrome (AIDS) and bovine spongiform encephalopathy (BSE), both epidemic diseases with long i/p's, resulted in a need to predict future epidemic behaviour; for AIDS, to plan health care facilities; for BSE, to ensure disposal facilities; for both to plan financial resources and control spread.

In order to model epidemics of infectious diseases, particularly to estimate probable numbers of cases with onset at any particular time, it is necessary to incorporate a term for the i/p f/d (Lui et al, 1986; Gail & Brookmeyer, 1988). Sartwell's hypothesis may not apply to diseases with long i/p's. During the course of an epidemic of a long i/p disease, left and right censoring makes direct observation of this distribution highly unreliable (Peterman, 1987); in addition, time of infection is often unknown. Therefore, for a previously undescribed disease, methods other than direct observation must be employed. One method is to extrapolate from information available for other diseases.

The main aim of this study is therefore to evaluate Sartwell's hypothesis in relation to the group of diseases known as transmissible spongiform encephalopathies (TSE's) or prion disease, both natural and experimental, utilising information and methods from AIDS if appropriate; complications include aetiological uncertainties for many forms of prion disease, and unknown date of infection/disease initiation for natural disease, often necessitating the use of a/o as 'proxy' i/p.

Underlying this aim is the attempt to validate, or otherwise, the incorporation of a log-normal i/p (or a/o) f/d term into BSE epidemic modelling.

A subsidiary aim is, where appropriate, to assess the effect of other variables on i/p (or a/o) and its f/d.

CHAPTER 2

AIDS: INTRODUCTION: THE INCUBATION PERIOD

2.1 Introduction to AIDS

2 1.1 The disease of AIDS - a summary.
2 1.2 The epidemiology of AIDS - a summary.

2.2 The incubation period in AIDS.

2 2.1 Predicting numbers of AIDS cases: Methods used, with special reference to incubation period.
2 2.2 Published estimates of the incubation period of AIDS.

CHAPTER 2

AIDS: INTRODUCTION: THE INCUBATION PERIOD

2.1 Introduction to AIDS

2.1.1 The disease of AIDS - a summary.

AIDS epidemiology had suggested an infectious aetiology (Peterman et al, 1985); most investigators now accept this is human immunodeficiency virus (HIV). Gazzard (1990) considers the evidence overwhelming; HIV is cultured from both donor and recipient lymphocytes in transfusion acquired AIDS (TA-AIDS). The T-lymphotropic retrovirus was isolated in 1983 (Barre-Sinoussi et al, 1983; Gallo et al, 1984); evidence of infection was found in virtually all AIDS patients (Peterman et al, 1985).

At sero-conversion, possibly several months post-infection (Hartsburgh et al, 1989; Longini & Clarke, 1989), over 50% of patients experience a mononucleosis-like syndrome (Lifson et al, 1988; Tindall et al, 1988;); incubation period (i/p) from infection to the development of AIDS is long and variable.

The original definition of AIDS (CDC, 1982a) was modified in 1987 (CDC, 1987) (appendix 5), encompassing wider recognition of possible manifestations (Peterman et al, 1985; Gazzard, 1990), and differences in world region being studied, depending on accessibility to laboratory testing and prevalent local infections (Matondo, 1992).

Other classification systems include grouping by infection stage (CDC, 1987), and level of T-helper cells (OKT4 count) (Gazzard, 1990); some patients do not fit any category exactly.

2.1.2 The epidemiology of AIDS - a summary.

AIDS was recognised in 1981; from October 1980 to May 1981 five cases of *Pneumocystis carinii* pneumonia in young homosexual men in Los Angeles were recorded (CDC, 1981a) whilst in New York and California, reported diagnoses of Kaposi's sarcoma in young homosexual men increased markedly (CDC, 1981b). By May 1985, there were 10,000 recorded United States cases of AIDS (Peterman et al, 1985).

Cases were soon recognised in other groups; bisexual males, intravenous drug users (Selik et al, 1984), recipients of blood transfusions (Ammann et al, 1983) haemophiliacs (CDC, 1982b), and

Haitian immigrants to the USA (CDC, 1982c). For TA-AIDS, traceable blood donors were often in other recognised risk groups (Curran et al, 1984), strengthening evidence for infection, which also included direct links, for example sexual partners, or intravenous needle-sharers (Auerbach et al, 1984; Curran et al, 1985a; Freidland & Klein, 1987)

Lack of denominator information in many risk groups made incidence rate estimation difficult (Peterman et al, 1985).

Subsequently, disease was reported from Europe (CDC, 1985a), Haiti (Pape et al, 1983), and Africa (Van der Perre et al, 1984; Piot et al, 1984). In Europe, major risk groups were homosexual males and intravenous drug users (Peterman et al, 1985), but in both Haitians and Africans, heterosexual males and females (including many prostitutes) comprised the majority of cases (Clumeck et al, 1984; Van der Perre et al, 1984; Pitchenic et al; 1983).

Soon, heterosexual women and men with no identified risk factors, both in USA and Europe, were diagnosed with AIDS (CDC, 1983; Harris et al, 1983); many had high risk group sexual partners (Harris et al, 1983), or were born in Africa or Haiti (Curran et al, 1985a), although sometimes no risk factor was identified (CDC, 1984). Most areas of the world subsequently experienced AIDS cases (Curran 1985; Quinn et al, 1986; Blattner, 1991).

Once tests were available for HIV antibody, it was realised that the number infected was far greater than those with disease. Early evidence, and the probability of a long i/p for AIDS came from the 'San Francisco cohort study' (Jaffe et al, 1985), originally set up to monitor hepatitis B infection in male homosexuals, supported by studies on transfusion recipients (e.g. Eyster et al, 1987; Lifson et al, 1988; Giesecke et al, 1988); thus the scale of the epidemic was realised.

Additional infection routes include needle-stick injuries (CDC, 1985b; Curran et al, 1985a; Peterman et al, 1985), transplants (Erice et al, 1991), maternal transmission (Rogers et al, 1987) and transmission via breast-feeding (Zeigler et al, 1985; Thiry et al; 1985; WHO/UNICEF, 1992). Up to 65% of babies born to infected mothers develop AIDS (Scott et al, 1985; Rogers et al, 1986; Thomas et al, 1987); evidence suggests i/p is shorter in the very young (Curran et al, 1985b; Medley et al, 1987; Medley et al, 1988a; Lui et al, 1988a).

2.2 The incubation period in AIDS.

2.2.1 Predicting numbers of AIDS cases: Methods used, with special reference to incubation period

Gail and Brookmeyer (1988) critically review three major methods; extrapolation, 'back calculation', and compartmental (transmission) models; Anderson's (1988) review highlights transmission models.

Earliest predictions extrapolated reported case numbers, usually with allowance for reporting lag (e.g. Curran et al, 1985a; McEvoy & Tillet, 1985; Downs et al, 1987). However, ascertainment bias alters with improved disease recognition; reporting lag changes with time; the mathematical function assigned to the epidemic curve may be one of many which fit the data up to that point; and no changes in transmission rate are allowed for. The method uses no i/p, infection rate, or transmission method knowledge, and leads to wide variation in predictions, but was the obvious choice early in the epidemic (Gail & Brookmeyer, 1988).

Additional information allowed for use of more complex methods. For data grouped by infection cohort, left and right censoring reduces the apparent mean i/p over time, (Peterman, 1987); conversely, grouping by diagnosis cohort causes the apparent mean i/p to increase (Rogers et al, 1987). 'Back calculation' is a parametric method (see appendices 6 and 7) to overcome the censoring effect (e.g. Lui et al, 1986; Lui et al, 1988a; Kalbfleish & Lawless, 1989a; Day et al, 1989; Hendriks et al, 1992; Newton et al, 1993) but evaluating best i/p frequency distribution (f/d) function (necessary for use in this method) from observed cases is complicated by unknown infection time, multiple exposures, insidious onset, variable disease definitions, recording errors (Brookmeyer & Gail, 1988; Geisecke et al, 1988; Kalbfleish & Lawless, 1989b; Peterman & Ward, 1989).

These problems are exacerbated by the interaction of changing infection rates on the apparent i/p; if stable, there is no problem. However if it increases, as in the early stages of any epidemic, or decreases with behavioural alterations (Brookmeyer & Gail, 1988; Bacchetti & Moss, 1989), the effect on apparent i/p depends on the rate of that change, since different numbers of cases infected at different times will differentially weight the 'censoring effect' (Lui et al, 1986; Kalbfleish & Lawless, 1988). Anderson (1988) concludes that the complete i/p f/d may take decades to evaluate; hence the use of parametric curves which fit early data.

However, early in the epidemic, many functions may fit i/p data equally well; different functions give different right 'tails', with widely differing predictions (Medley et al, 1987; Brookmeyer & Gail, 1988; DoH, 1988; Kalbfleish & Lawless, 1989a; Day et al, 1989). The real i/p may change

with time, due to drug therapy or other co-factors (Brookmeyer & Gail, 1988; Schechter et al, 1989; Brookmeyer, 1991); or with risk group (Brookmeyer & Gail, 1988; Day et al, 1989). Much i/p information comes from TA-AIDS cases, who may be less healthy at infection, or receive a larger dose (Brookmeyer & Gail, 1988).

Parametric methods are applicable for short or (possibly) medium term estimates of case numbers, which are not highly sensitive to i/p f/d function; however, the calculated mean i/p, and long-term estimates are extremely sensitive to the function used, and therefore differ widely (Gail & Brookmeyer, 1988; Anderson, 1988).

With compartmental models based on transmission pathways (see appendix 7), (e.g. Anderson et al, 1986; Blythe & Anderson, 1988; Longini & Clark, 1989; Velasco-Hernandez & Hsieh, 1994), classes are identified (e.g. susceptible, infected, diseased, dead), and transfer rates between classes allocated, based on available information. The complexity of these models depends on categories included, but still utilises unverified assumptions, including transfer rate from infected to diseased dependant upon the i/p f/d, with the uncertainties previously discussed. Incorporation of transmission rates allows simulation of different hypotheses, sub-sets, treatments etc. for comparisons; they may eventually be useful for quantitative predictions (Gail & Brookmeyer, 1988; Taylor, 1989a).

Other methods used in analysing the AIDS epidemic include survival analysis of a particular infection cohort, which can give information on the i/p f/d and risk of developing AIDS (e.g. Eyster et al, 1987; Giesecke et al, 1988), and other non-parametric methods (e.g. Bacchetti & Moss, 1989; Kalbfleish & Lawless, 1989a). Rees (1987) investigated transfusion data utilising a normal i/p f/d, concluding that mean i/p was far longer than generally accepted at that time. For this, and his methodology, he was heavily criticised (Barton, 1987; Costagliola & Downs, 1987; Lui et al, 1987).

2.2.2 Published estimates of the incubation period of AIDS.

The estimated i/p of AIDS varies with method used and group studied, and tends to lengthen over time (Taylor et al, 1991). In addition, published i/p information is given in different ways (see appendix 8).

CHAPTER 3

TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES (TSE'S)/ PRION DISEASE: INTRODUCTION

3.1. The TSE/prion disease spectrum.

3.2. The diseases - summaries.

3.2.1 Scrapie: the prototype.

3.2.1.1 Natural disease.

3.2.1.2 Experimental scrapie; early work.

3.2.2 Bovine spongiform encephalopathy: the epidemic.

3.2.3 Other veterinary TSE's/prion diseases.

3.2.4 Human TSE/prion diseases.

3.3 The aetiological debate.

3.3.1 The first debate: the biochemical nature of the transmissible agent.

3.3.2 The second debate: the aetiological mechanism in natural disease.

3.4 Relevance of animal TSE's to human public health.

CHAPTER 3

TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES (TSE'S)/ PRION DISEASE: INTRODUCTION

3.1 The TSE/prion disease spectrum.

Characteristic of the group of diseases traditionally called transmissible spongiform encephalopathies (TSE's) is neurodegeneration with vacuolation and amyloid fibrils (scrapie-associated fibrils, SAF's), a variety of neuropathologically-associated clinical signs, fatal outcome, experimental and iatrogenic transmissibility of infection, and a 'long' incubation period (Kimberlin, 1990); additionally, there is no easily detectable immune reaction to the presence of infectious agent (Bennett et al, 1992; Berg, 1994), precluding any simple method for identifying infection. Hadlow (1959) drew attention to the similarity between scrapie and kuru, and Gajdusek (1977) postulated similar aetiological mechanisms for scrapie, CJD and kuru.

It has since been recognised that all TSE's involve the presence of an aberrant isoform (PrP^{Sc}) of normal cellular prion protein (PrP^{C}); it forms the main constituent of the amyloid deposits. Presence of PrP^{Sc} is regarded by Roberts and Collinge (1990) as diagnostic of transmissible dementias; the name 'prion disease' is recommended (Collinge & Prusiner, 1992).

There is, as yet, no consensus on terminology; nevertheless there is broad consensus on disease range (table 3.1).

Table 3.1 TSE/prion disease spectrum (after Prusiner and DeArmond, 1994).

Human diseases	
Sporadic CJD*	[* CJD = Creutzfeldt-Jakob disease]
Familial CJD	
Gerstmann-Straussler-Scheinker syndrome (GSS)	
Fatal familial insomnia (FFI)	
Iatrogenic CJD	
Kuru	
Other species diseases	**Species**
Scrapie	Sheep, goats
Bovine spongiform encephalopathy (BSE)	Cattle
Feline spongiform encephalopathy (FSE)	Cats (domestic and large)
Transmissible mink encephalopathy (TME)	Mink
Chronic wasting disease (CWD)	Elk, mule deer
TSE of exotic ungulates	Several, e g. nyala, eland.

Recent overviews include 'Prion Diseases of Humans and Animals' (1992) edited by Prusiner et al; 'Molecular Biology of Prion Diseases' (1994), edited by Collinge & Weissmann; 'Transmissible spongiform encephalopathies - Impact on animal and human health' (1993), edited by F Brown; TSE's of animals are reviewed generally (OIE, 1992).

3.2. The diseases - summaries.

3.2.1 Scrapie: the prototype.

3.2.1.1 Natural disease.

Scrapie was first documented in sheep in England in about 1730 (Comber, 1772); Besnoit and Morel (1898) describe associated neuropathology. Onset, usually between 2-5 years of age, is generally insidious; signs vary, and include intense pruritis, hyperexcitability, incoordination, ataxia and loss of condition; outcome is fatal (Mitchell & Stamp, 1983). Sudden death is possible (Clark, 1991).

Classic pathology is confined to the CNS and includes non-inflammatory neurodegeneration with vacuolation and spongiform change (Parry, 1983). Amyloid deposits are sometimes found (Beck et al, 1964). Lesion severity varies, not necessarily correlated with clinical signs (Hadlow et al, 1982). Since there is no easily detectable immunological reaction to infection, diagnosis relies on clinical signs and histopathological confirmation (Detwiler, 1992).

Scrapie has been reported in most world regions. Epidemiology in the United States (Hourrigan et al, 1979; Hadlow, 1990) and Iceland (Sigurdsson,1954; Palsson, 1979) is described. Parry (1983) extensively reviews historical and world epidemiology; Detwiler (1992) gives recent data.

Although an infectious origin was considered likely as far back as 1794 (Stockman, 1913, citing Young, 1799), in 1848, Roche-Lubin suggested scrapie was caused by sexual excess, or thunderstorms. Proponents of the infectious theory include M'Fadyean (1918) and Kimberlin (1990). That the disease is transmissible, from sheep to sheep under experimental conditions, was demonstrated by Cuille & Chelle (1936); Gordon et al (1939) reported iatrogenic scrapie by contamination of louping-ill vaccine.

Conversely, Yorkshire sheep farmers considered scrapie to be hereditary (Stockman, 1913, citing a report to the Board of Agriculture, 1812), as did French farmers (Chelle, 1942). Bosanquet et al (1956) described the disease as a primary myopathy, probably of genetic origin, and Parry (1962, 1983) presented much data on which he based his 'recessive gene hypothesis'.

As early as 1913, Stockman reported that 'some people' considered the disease to be both infectious and hereditary; Prusiner and DeArmond (1994) also consider this possible. Pattison (1974) concluded that two cases which he observed in genetically susceptible sheep had arisen 'spontaneously'. The aetiological debate surrounding natural scrapie is still unresolved.

General reviews of scrapie in sheep, documenting accumulation of information and changing ideas, include those by Stockman (1913), Greig (1940a), Palmer (1959), Parry (1983) and Detwiler (1992).

Natural scrapie in a goat was first reported by Chelle (1942). There is a paucity of information; Wooldridge and Wood (1991) reviewed the epidemiological information available; pathology (Wood & Done, 1992) and clinical history (Wood et al, 1992) is described for 20 cases.

Genetic studies have now identified mutations of the sheep PrP gene which have been associated with increased susceptibility to natural scrapie. For example, at codon 136 in Shetland, Scottish Halfbred and Bleu du Maine breeds, valine is associated with scrapie whilst alanine is associated with healthy animals (Hunter et al, 1994a), and in Suffolk sheep, homozygous codon 171 mutations are associated with disease (Westaway et al, 1994). No PrP gene mutations have been reported as affecting susceptibility in goats.

3.2.1.2 Experimental scrapie; early work.

After experimental transmission of scrapie from sheep to sheep (Cuille & Chelle, 1936) further landmarks include the following:-

1957: Experimental transmission, sheep to goats (Pattison, 1957).
1961: Experimental transmission to mice (Chandler, 1961).
1961: Oral transmission to sheep and goats (Pattison & Millson, 1961a).
1968: Sip (*scrapie incubation period)* gene found in sheep (Dickinson et al, 1968).
1969: Sinc (*scrapie incubation period*) gene found in mice (Dickinson et al, 1969).
1969: Experimental transmission to monkeys (Gibbs & Gajdusek, 1969).
1980: Oral transmission to monkeys (Gibbs et al, 1980).

Experimental transmission changed the focus of investigation from natural to experimental disease; in particular, to disease in small laboratory rodents, since this reduced incubation periods by several years (reviewed by Wooldridge, 1991). Transmission could be used for diagnosis, and also allowed extensive pathology and pathogenesis studies, for example in goats (Hadlow, 1961; Eklund & Hadlow, 1969) and mice (Gibbs et al, 1965; Eklund et al, 1967; Outram, 1976).

Although the most efficient experimental transmission method was to use brain material from the donor (Pattison & Millson, 1962) inoculated into the CNS of the recipient, the involvement of the lymphoreticular system was also identified; it's role in pathogenesis is reviewed by Fraser et al (1992). Different 'strains' of scrapie agent are described, as is the effect of host genotype on incubation period, which varies with 'strain' (Dickinson & Mieckle, 1971; Bruce & Dickinson, 1979; Kimberlin & Walker, 1978).

Periodic reviews of accumulating information on many aspects of both experimental and natural scrapie include those in 'Slow, Latent and Temperate Virus Infections' (1965), edited by Gajdusek et al, 'Slow Virus Diseases of Animals and Man' (1976), edited by Kimberlin, and 'Slow Transmissible Diseases of the Nervous System' (1979), edited by Prusiner and Hadlow.

Pattison (1988; 1992) reviews early scrapie research in the UK, emphasising that natural disease is not the same as experimental disease.

3.2.2 Bovine spongiform encephalopathy: the epidemic.

Bovine spongiform encephalopathy (BSE) was first reported in 1987 in Great Britain (Wells et al, 1987); neuropathology with degenerative changes and vacuolation reminiscent of scrapie is described. Subsequent pathological investigation indicates that, unlike scrapie, only one 'strain' of agent is involved (Wells et al, 1992; Bruce, 1994), and no genetic association with differences in incubation period has been found (Hunter et al, 1994b).

Epidemiological evidence strongly suggests foodborne infection with a scrapie-like agent derived from ruminant tissues, from approximately 1981, with the first case in 1985 (Wilesmith et al, 1988). Changes in rendering methods for producing meat and bone meal (MBM) are implicated, particularly cessation of hydrocarbon solvent extraction (Wilesmith et al, 1991). Rapid epidemic growth (Wilesmith et al, 1992) resulted in a statutory ban on feeding of ruminant protein to ruminants (Order, 1988). Incidence rapidly declined in cattle born subsequent to the ban (Hoinville, 1994).

BSE in Northern Ireland is epidemiologically similar to Great Britain (Denny et al, 1992). Occasional cases from Switzerland (Anon, 1990a) France (Anon, 1991), Eire, the Falkland Islands (Bradley & Lowson, 1992), Oman (Carolan et al, 1990) and Portugal (Telo, 1994) are reported.

Onset is generally from 3-5 years of age (Wilesmith et al, 1988). Clinical signs vary; most frequently recorded are apprehension, hyperaesthesia and ataxia; others include kicking, tremors, loss of weight and reduced milk yield (Wilesmith et al, 1992; Austin, 1993). As with scrapie, there is no test for infection; diagnosis relies on clinical signs and histopathology (Kimberlin, 1992).

Using brain-derived material, BSE has been transmitted to mice, both by intracerebral inoculation (Fraser at al, 1988), and orally (Barlow & Middleton, 1990); to pigs (Dawson et al, 1990), to sheep and goats (Foster et al, 1993), and to marmosets (Baker et al, 1993). Infectivity has also been reported in the ileum (Wells et al, 1994). The disease is not simply inherited (Wijeratne & Curnow, 1990), and thus far, no PrP gene mutations associated with susceptibility have been found (Wilesmith, personal communication, 1995).

Recent reviews include Wilesmith (1994) on epidemiology, and a general review by Kimberlin (1992), plus the Office International des Epizooties report (OIE, 1994).

3.2.3 Other veterinary TSE's/prion diseases.

Transmissible mink encephalopathy (TME) was first reported (Hartsough and Burger, 1965) as several outbreaks, in Wisconsin, of a disease of unknown aetiology with similarities to scrapie. A foodborne infection was hypothesised, possibly of bovine or ovine origin. Transmissibility was demonstrated (Marsh et al, 1969). Subsequently, an outbreak where sheep had never been fed suggested the possibility of unrecognised scrapie-like disease in cattle (Marsh and Hartsough, 1985). Marsh and Hadlow (1992) review TME.

Chronic wasting disease (CWD) has been reported in mule deer and elk, both species of the family cervidae (Williams & Young, 1980; 1982). Williams and Young (1992) review the disease; it is scrapie-like, occurring in both captive and free-living animals. It is not associated with foodborne infection; lateral and probably maternal transmission occur.

Feline spongiform encephalopathy (FSE) was first reported in 1990 (Wyatt et al) in a domestic cat; pathological findings were consistent with a TSE. Other cases followed (Legget et al, 1990; Wyatt et al, 1991), including a captive puma (Willoughby et al, 1992), and several cheetah (Kirkwood & Cunningham, 1994). Pearson et al (1993) review findings for 24 cases. Transmission to mice is reported (Fraser et al, 1994). Kimberlin (1992) suggests contaminated MBM in catfoods as the most likely source of infection.

TSE of exotic ungulates was first observed in nyala in 1986, and subsequently in gemsbok (Jeffrey & Wells, 1988), Arabian oryx and greater kudu (Kirkwood et al, 1990) and eland (Fleetwood & Furley, 1990), all species of the family bovidae. Kimberlin (1992) suggests that the majority of these animals were fed contaminated MBM in their diets, however, for kudu, there is evidence that maternal and lateral infection subsequently occurred (Kirkwood et al, 1992; Kirkwood et al, 1993; Kirkwood et al, 1994).

Wells (1993) reviews pathology of animal TSE's, whilst Kirkwood and Cunningham (1994) review epidemiology in captive wild animals.

3.2.4 Human TSE/prion diseases.

Creutzfeldt-Jakob disease (CJD) was first described in the 1920's (Creutzfeldt, 1920; Jakob, 1929), characterised mainly by a rapidly progressive dementia. It is generally sporadic, but around 15% are familial cases (Masters et al, 1979). Gerstmann-Straussler- Scheinker syndrome (GSS) was first described in 1936 (Gerstmann et al), and is familial, with an autosomal dominant pattern of inheritance (Brown, 1993); dementia is slowly progressive and ataxia often present. Families exist which exhibit both clinical syndromes (Ridley and Baker, 1993).

In 1957, Gajdusek & Zigas reported a localised epidemic of kuru in the Fore people of Papua New Guinea, associated with ritual cannibalism (Klitzman et al, 1994). Typical progression is locomotor ataxia, tremors, severe speech impairment and death (Zigas and Gajdusek, 1959). Hadlow (1959) pointed out similarities to scrapie, suggesting transmission experiments which proved successful both by inoculation (Gajdusek et al, 1966), and orally (Gajdusek, 1979).

CJD was also transmitted experimentally to chimpanzees and monkeys (Gibbs and Gajdusek, 1969). Iatrogenic CJD has resulted from neurosurgery (Will & Matthews, 1982), dura-mater and corneal grafts (Baker, 1990), and injections of pituitary extract (Gibbs et al, 1985), confirming its transmissibility to other humans.

Fatal familial insomnia (FFI) was recognised as a prion disease in 1992 (Medori et al); sleep disturbance is a particularly prominent sign; 'atypical' prion dementia (APD) is described by Ridley & Baker (1993), identified by accumulation of PrP^{Sc} or PrP gene mutation.

Brown (1993b) reviews the clinical picture, emphasising continuity of the spectrum across all these diseases, and Kretzschmar (1993) reviews neuropathology. Masters et al (1979) review

worldwide epidemiology of CJD; Will et al (1992) review CJD epidemiology in the UK with particular emphasis on occupation.

Many genetic mutations have been identified in the PrP gene in families with familial disease (e.g. Hsiao et al, 1989, 1991a, 1992; Brown et al, 1991; Collinge et al 1992, 1993). Slovakian cases are associated with mutation at codon 200 (Meyer et al, 1977; Goldfarb et al, 1990), as are cases of Libyan Jewish ancestry, where dietary factors (in particular sheep's eyes) were originally suspect (Kahana et al, 1974; Hsiao et al, 1992b).

Polymorphism of the PrP gene at codon 129 is found in the general population, and in TSE patients. The two allelic forms code for methionine (met) or valine (val), giving met/met, met/val, or val/val (Collinge & Palmer, 1992). Codon 129 homogeneity is associated with susceptibility to development of CJD in pituitary hormone injection cases (Collinge et al, 1991; Deslys et al, 1994); homozygosity may also predispose to sporadic CJD (Palmer et al, 1991). Polymorphism at codon 129 may also affect clinical manifestation or age at onset in the presence of other PrP gene mutations (Crow et al, 1992; Ridley & Baker, 1993). In the presence of codon 178 mutation, codon 129 polymorphism may produce different clinical syndromes, CJD and FFI (Monari et al, 1994). There may be other genetic susceptibility factors (either on the PrP gene or elsewhere), as yet undiscovered (Ridley & Baker, 1993).

Annual incidence of all forms of human spongiform encephalopathies is usually given as no more than one case per million population (Baker, 1990).

3.3 The aetiological debate

Controversy on the aetiology of these diseases continues. There are two separate, but associated, debates. That experimental transmission is possible is generally accepted.

3.3.1 The first debate: the biochemical nature of the transmissible agent.

The hypothesis that the agent is comprised of protein alone has developed largely from the 'prion' hypothesis. Prusiner (1982) proposed a small proteinacious infectious particle, the prion, as the transmissible scrapie agent, suggesting it might comprise either a tightly packed protein coat surrounding a protected small nucleic acid, or an infectious protein alone, and that host genes may code for both a protein-only scrapie prion and a related, necessary protein. This hypothesis was based on the agent's physico-chemical properties; it was unusually stable (e.g. Hunter & Millson, 1964, Pattison, 1965c); smaller than viruses (Alper et al, 1966) and not isolated by

conventional virus-isolation techniques (Hunter and Millson, 1967). Millson et al (1976) review the agent's physico-chemical properties. Hunter (1992) summarises this early work.

Following publication of the 'prion' hypothesis, McKinley et al (1983) report that the concentration of a protease-resistant protein designated PrP, from scrapie-infected hamster brain, is proportional to infectivity, suggesting it as the major component of infectious prions. This PrP assembles into filaments, which accumulates to form amyloid plaques (DeArmond et al, 1985). Oesch et al (1985) report that a cellular gene encodes for this protein, now called PrP^{27-30}, and Basler et al (1986) show that PrP^{27-30} and normal cellular PrP isoforms are encoded by the same gene, suggesting differences are due to post-translational events. Safar et al (1990) report diminished infectivity of scrapie precursor protein (now called PrP^{Sc}) preparations after addition of polyclonal antibodies raised against PrP^{27-30}; biochemical and physicochemical analyses indicate that PrP^{Sc} may be the only component of the infectious unit. PrP^{27-30} is present in cattle brains after scrapie inoculation (Gibbs et al, 1990).

Brown et al (1990) describe resistance of scrapie agent to autoclaving and formaldehyde treatment, results compatible with replication initiation by nucleation and auto-patterning (e.g. by an inorganic crystal) of a configurational change in scrapie precursor protein, followed by crystalline growth, thus mimicking replication of a microbiological agent. A similar conclusion was reached by Come et al (1993). Jarrett and Lansbury (1993) review possible mechanisms.

Pan et al (1993) demonstrate conformational differences in PrP^{C} and PrP^{Sc}, whilst Kocisko et al (1994) report 'seeded' conversion of PrP^{C} to PrP^{Sc} in a cell-free medium. Work with transgenic models indicates that mice producing no PrP^{C} are resistant to scrapie (Bueler et al, 1993); and that efficiency of interaction between PrP^{C} and introduced PrP^{Sc} depends on the genetic source of the two PrP's, with similar sources leading to increased efficiency of disease production, reduced incubation times and reduction of the 'species barrier' (Scott et al, 1989; Prusiner et al, 1990; Weissmann, 1991a; Scott et al, 1993); all add support to the possibility of conformational change in native PrP^{C} initiated by incoming, $PrP^{Sc/CJD/etc.}$.

The main problem with acceptance of the protein-only hypothesis is the existence of distinct strains (Weissmann, 1991a); however, both protein-only and 'unified' theory of prion propagation (in which nucleic acid is involved, but not necessary for infection) (Weissmann, 1989; 1991b) can account for strain differences.

Aspects of both natural and experimental transmission in TSE's/prion diseases are reviewed (Prusiner, 1987), as well as the chemistry and neurochemistry of prions and prion diseases (Prusiner, 1989; DeArmond and Prusiner, 1993). Prusiner (1995) argues that the weight of

evidence now accumulated indicates that the prion is a new class of pathogen, with prion disease resulting from aberrations of protein conformation.

However, the agent was originally assumed to contain nucleic acid, its small size suggesting a virus (Hunter 1965), and the search for nucleic acid continues. Sklaviadis et al (1989, 1990), describe studies of the physical properties of the CJD agent, and conclude that infectivity is associated with nucleic acid-protein complexes, the agent being of virus-like size; a nucleic acid binding protein is described (Sklaviadis et al, 1993). Narang (1990, 1992, 1993) describes a single-stranded DNA as part of a SAF 'coat', Akowitz et al (1990) and Murdoch et al (1990) report RNA, possibly of retroviral origin, as co-purifying with CJD infectivity, whilst Aiken et al (1989, 1990) and Narang et al (1991) implicate mitochondrial DNA as involved in scrapie infection. Aiken and Marsh (1990) review evidence for scrapie agent nucleic acid; they conclude that evidence points to an agent with host-encoded protein protecting an infectious nucleic acid; the virino hypothesis.

Schreuder (1993) reviews evidence for virus, virino, prion, and unified hypotheses, including evidence from transgenic experiments.

Prusiner was not the first to postulate an infectious agent devoid of nucleic acid. Results of radiobiological experiments led Alper et al (1967) to suggest the agent might replicate without nucleic acid; Gibbons and Hunter (1967) put forward the 'membrane hypothesis' that plasma membrane is essential for infectivity. Alper (1993) argues the case for membrane fragment involvement.

3.2.2 The second debate: the aetiological mechanism in natural disease.

The historical debate on the aetiology of scrapie has been outlined; that debate continues, and extends to all TSE/prion disease. One viewpoint holds that all cases of TSE involve a transmissible agent, linked, in certain cases to genetic susceptibility which may be hereditary or the result of somatic mutation (Diringer et al, 1993).

Alternative theory allows for hereditary, infectious and sporadic disease initiation mechanisms (DeArmond & Prusiner, 1993). Hereditary genetic mutations in the PrP gene may result in instability of the PrP^{C} (the initial gene product), allowing spontaneous post-translational change to, for example, PrP^{CJD}, PrP^{GSS}, PrP^{Sc} or PrP^{FFI} dependant on the disease in question. Once produced, this altered PrP is infectious if introduced, under appropriate conditions, into a suitable host, as in experimental transmission, iatrogenic CJD, kuru, BSE, and probably most scrapie

cases; precise transmission pathway may be unknown. Under this scenario, sporadic CJD may be either the result of transmission, possibly coupled with susceptibility (perhaps due to a somatic mutation - an 'age effect') or spontaneous, again perhaps due to a somatic mutation.

That transgenic mice containing a mutant prion protein gene from a GSS patient spontaneously developed neurodegenerative disease (Hsiao et al, 1990) lent early support to this theory; reviews of subsequent supporting research and its interpretation include those of Hsiao and Prusiner (1990), Prusiner and Westaway (1991) Prusiner (1994) and Prusiner and DeArmond (1994). Nevertheless, there is still dispute regarding whether PrP gene mutation is sufficient aetiological cause for disease, or merely a factor increasing susceptibility to infection or disease. Hunter et al (1992) review the significance of PrP gene polymorphisms in natural scrapie.

Alper (1993) proposes a variant on the second theory, whereby genetic-mutation-derived abnormal PrP results in failure of plasma membrane to develop correctly, thus eventually causing disease. The infectious agent is not altered PrP alone, but PrP^{Sc}-loaded plasma membrane vesicles, which incorporate into host membrane.

3.4 Relevance of animal TSE's to human public health

Epidemiological studies have not linked occurrence of CJD to exposure to scrapie (review: Taylor, 1989b). However, the discovery, and subsequent epidemic of BSE in the UK led to renewed speculation that animal TSE's could pass to humans, since the agent had apparently crossed the 'species barrier' to cattle and cats; potential hazards are examined (Blakemore, 1989; Taylor, 1989b; Collee, 1990; Kimberlin, 1990; Brown, 1993a). Committees (Southwood, 1989; Tyrrell, 1990) were set up to investigate the epidemic, evaluate necessary research, and advise the government; methods, including legal sanctions, used to contain the epidemic and minimise risk to other species are summarised (Taylor, 1991; Kimberlin, 1992), and have recently been extended (Anon, 1994).

Surveillance of human TSE cases in the UK has been established (Anon, 1990b), and monitoring of the BSE epidemic by MAFF continues at the CVL; on 23/12/94 over 141 thousand confirmed cases had been recorded. No BSE transmission to humans has yet been proven; it may be impossible to prove that it never occurs (Matthews, 1990).

But there are other ways in which BSE can affect human health; at least one case of 'delusional BSE' has occurred in a human (Lovestone, 1990).

CHAPTER 4

DATA USED AND ANALYSIS METHODS NOT DESCRIBED ELSEWHERE

4.1 Data used.

4.1.1 The AIDS dataset used.

4.1.2 Data used for TSE/prion disease examination.

4.2 Outline of data examination.

4.3 Methods not described elsewhere.

4.3.1 Data storage.

4.3.2 Analyses; methods not described elsewhere.

CHAPTER 4

DATA USED AND ANALYSIS METHODS NOT DESCRIBED ELSEWHERE

4.1 Data used.

4.1.1 The AIDS dataset used.

Data comprises the set of 4010 TA-AIDS cases recorded in the USA up to and inclusive of December 1990 and was kindly supplied on disc by T Peterman, CDC, Atlanta, USA. Appendix 9 is a copy of the accompanying explanatory leaflet. Records supplied were as received by CDC; they were not 'cleaned' (i.e. not checked for completeness or errors).

4.1.2 Data used for TSE/prion disease examination.

Data was from 3 sources; a general literature search (GLS) for each disease, data supplied directly by other investigators (either written or on disc), and data extracted directly from case records. Appendix 10 gives details of data-source by species, and describes and explains the selection of datasets and dataset subgroups used in the analyses in chapters 6 to 11.

4.2 Outline of data examination.

Chapters 5 to 11 have a similar construction. Each starts with a brief summary, followed by a set of quantitative analysis results pertaining to TSE/prion disease in a particular type of animal (or for chapter 5, AIDS in humans), and concludes with a qualitative discussion of those results. Chapters 12 and 13 examine aspects of the overall findings.

For chapters 6 to 11, after selection of datasets and subgroups within these datasets (for details, see appendix 10), each dataset is examined for incubation period (i/p) or age at onset (a/o) parameters. The i/p and/or a/o frequency distribution (f/d) is then generally illustrated with a histogram, examined statistically for compatibility with normality and log-normality, and the coefficient of skewness (c/s) calculated. This initial pattern of analyses follows a broadly similar format for each dataset, thus avoiding the necessity for lengthy, repetitive descriptions within the analysis sections of each chapter. Additional analyses either on complete datasets or subgroups,

for example to examine the effect of censoring, or of variables, or of culling, are undertaken where appropriate and individually described or referenced.

4.3 Methods not described elsewhere.

4.3.1 Data storage.

All data was entered into specifically written databases in the epidemiological software programme Epi Info, version 5.01a (1991). Data on disc was transferred directly, other data was entered manually.

4.3.2 Analyses; methods not described elsewhere.

Unless otherwise stated, analyses were performed using Epi Info as above, and methods given in the reference manual (version 5). For non-parametric testing, Epi Info uses the Kruskal-Wallis test (Siegel, 1956). The following additional methods were used frequently; other methods are described or referenced when used.

Compatibility of the f/d for normality or log-normality was tested by the method of Filiben (1975), using Minitab statistical software, release 7.2 (1989) (see appendix 11). Evidence of skewness in the untransformed and log-transformed data was tested for by the method of Snedecor and Cochran (1967), again using Minitab software (see appendix 12). Regression analysis and analysis of variance were also performed using Minitab statistical software, plus methods given in the reference manual, release 7 (Minitab, 1989). For f/d's with cases grouped by more than 1 time-unit, standard deviation was corrected by the method of Sheppard (1897) (see appendix 13).

The meaning of the following abbreviations are explained in detail in method appendices 11 and 12 and are regularly used in the tables of results without further comment:-

Correlation coefficient, CC	Sample size, N
Critical value, C/v	Probability, P
Coefficient of skewness, C/s	

Appendix 14 gives an indication of the discriminatory power of the tests used to estimate compatibility with normality and log-normality. Appendix 15 shows the application of these tests to diseases with short i/p's, and appendix 16 illustrates constructs of log-normal distribution for selected datasets for comparison with the observed f/d.

CHAPTER 5

AIDS: DATA ANALYSIS AND DISCUSSION

5. Summary

5.1 Incubation Period Analyses

5.1.1 Estimate of mean incubation period for three groups.
5.1.2 The effect of left and right censoring.
5.1.3 Back-calculation to overcome censoring.
5.1.4 Examination of incubation period frequency distribution.

5.2 Discussion

5.2.1 Mean incubation period estimates.
5.2.2 Incubation period frequency distribution.
5.2.3 Mean incubation period; comparison of back-calculation with direct estimates.

CHAPTER 5

AIDS: DATA ANALYSIS AND DISCUSSION

Summary.

The effect of censoring in an epidemic with a long incubation period (i/p) renders the observed mean i/p estimation and i/p frequency distribution (f/d) description liable to substantial errors until well into the epidemic's course.

Back-calculation methods may be applied to other epidemics in a similar fashion to that described here, in an effort to overcome the problem of censoring. However, since they rely on the assumption of an i/p f/d they do not help to elucidate the true i/p f/d. They may give a closer approximation to the true mean i/p if censoring is great, but validation is only possible once the true i/p f/d can be found from other sources. Estimates of the mean and maximum i/p may vary greatly with i/p f/d assumed, rendering long-term predictions unsafe.

It is impossible to accurately estimate the mean i/p or i/p f/d of a disease with a long i/p, by using data from within that epidemic, until several successive stable i/p range infection cohorts have passed, showing the maximum possible i/p; and there has been enough further time for such cohorts to have been recorded, analysed, and identified; monitoring is necessary for this stage to be recognised as soon as possible.

Sartwell's hypothesis can not be verified, or refuted, from this data, since the stage of stable i/p range by infection cohort has not been reached.

5.1 Incubation period analyses.

5.1.1 Estimate of mean incubation period for three groups.

Estimates are given, in table 5.1.1, for cases with a single transfusion (tx) only, for those with a traced suspect donor and for the most complete dataset possible, including both of these plus cases with multiple transfusions (see appendix 17 for selection of data used, appendix 18 for method of estimation of multiple transfusion infection date):-

Table 5.1.1 TA-AIDS; estimates of i/p (in months) for three datasets (see text).

Dataset	N	Mean i/p	SD	Range
Single tx	911	52.5	27.0	3-208
Suspect donor	53	29.3	15.1	6-68
Complete	2570	55.1	28.0	1-415

5.1.2 The effect of left and right censoring.

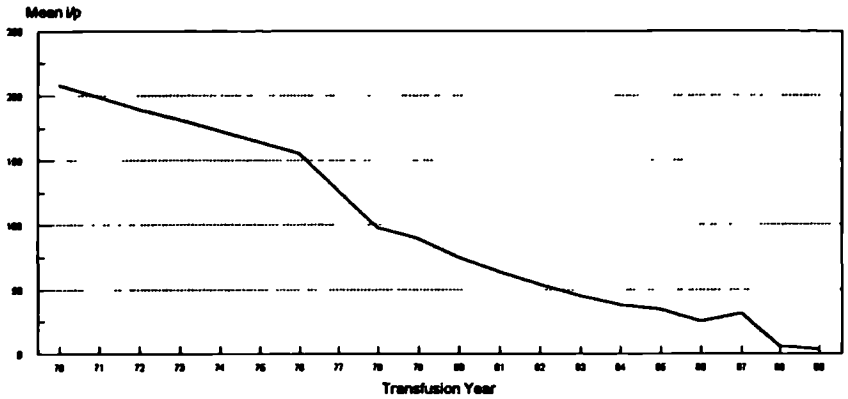
In demonstrating censoring, to reduce infection date uncertainty, only single (or suspect donor) transfusion cases are used. Table 5.1.2 gives i/p data by transfusion year.

Table 5.1.2 TA-AIDS; the effect of left and right censoring on the i/p.

Results:-	Number N	Mean i/p (months)	Range (months)	SD (months)
Complete set:	911	52.5	3-208	27.0
Tx year:				
1970	1	208.0	-	-
1971	0	-	-	-
1972	1	189.0	-	-
1973	1	181.0	-	-
1974	0	-	-	-
1975	0	-	-	-
1976	1	155.0	-	-
1977	5	126.4	93-142	20.4
1978	13	97.8	44-139	27.3
1979	40	89.1	47-125	19.8
1980	68	74.8	29-122	24.1
1981	124	63.7	9-105	24.0
1982	175	53.6	9-96	20.2
1983	206	45.0	3-86	19.3
1984	199	37.9	4-75	15.9
1985	61	34.6	8-62	13.6
1986	11	24.8	4-41	13.2
1987	3	31.7	28-34	3.2
1988	1	6.0	-	-
1989	1	3.0	-	-

The mean i/p decreases from 208 months (1970) to three months (1989) (figure 5.1.2).

Figure 5.1.2 TA-AIDS; mean i/p (in months) by year of transfusion.



Regression analysis gave the annual mean i/p decrease as 9.7 months (95% CI 9.0-10.4) ($r = -0.67$; $F=743.95$; $P<0.001$).

5.1.3 Back-calculation to overcome censoring.

The single transfusion dataset was again used. Back-calculation (kindly undertaken by R Sayers; see appendix 19 for method, and assumptions made) was performed for three different i/p f/d assumptions; gamma, Weibull, and log-normal. Several estimations were performed (data not shown) each using different additional assumptions; all gave broadly comparable results. Chi-square goodness-of-fit tests, and plausibility of assumptions made, were combined in selecting the specific set of results shown (table 5.1.3). There were 724 cases in the final dataset used.

Table 5.1.3 TA-AIDS; estimates of i/p values (in months) and related data from back calculation.

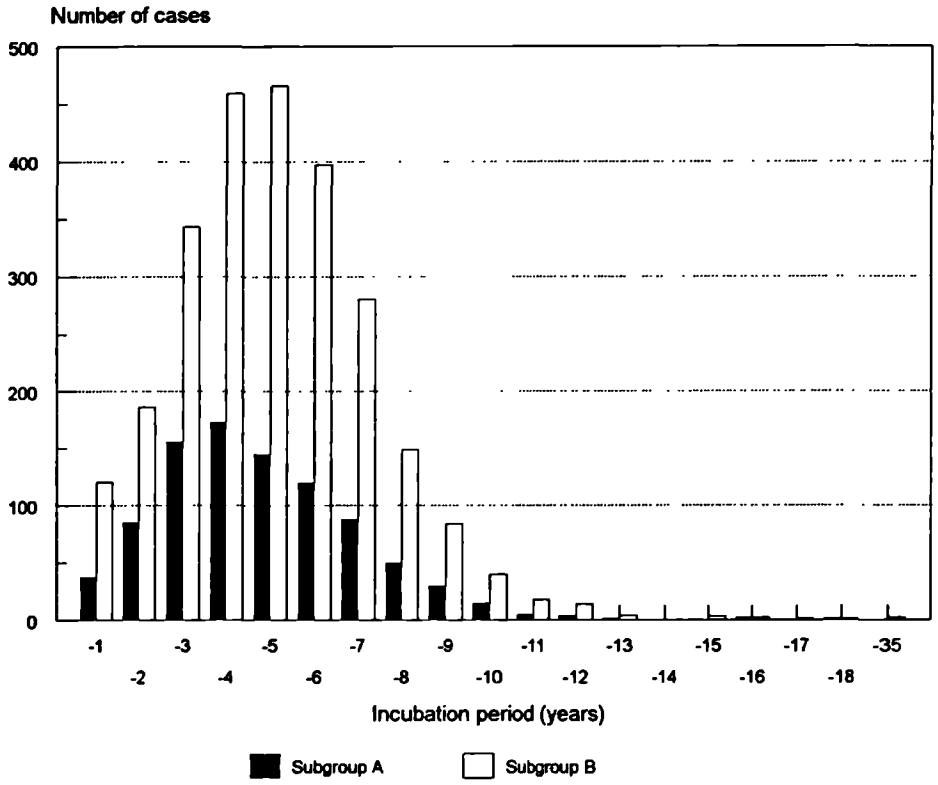
Distribution	Percentiles			
	1%	50%	95%	99%
Gamma	6	66	189	265
Weibull	4	66	176	234
Log-normal	9	93	487	968
	Mean	Mode	Max log likelihood	
Gamma	79.3	39	-1999.98	
Weibull	75.9	40	-2002.17	
Log-normal	154.4	34	-1998.56*	

As expected, for all f/d's the mean i/p has increased markedly from the 52.5 months for the simple mean, to 6-7 years for gamma and Weibull distributions, and to almost 13 years for the log-normal distribution, for which the maximum log-likelihood is actually the largest (although all are close). Even more striking is the 99% value, at almost 20 years (234 months) for the Weibull distribution, and over 80 years (968 months) for the log-normal distribution.

5.1.4 Examination of the incubation period frequency distributions.

Two subgroups were examined for i/p frequency distribution (f/d); the single transfusion dataset (subgroup A; 911 cases), and that obtained after weighting for multiple transfusion cases (subgroup B; 2570 cases) (figure 5.1.4).

Figure 5.1.4 TA-AIDS; subgroups A and B; i/p f/d.



[Note: I/p scale: -1 indicates cases with onset within one year of transfusion etc.]

There is a difference in the mean of these subsets (P=0.004, non-parametric test), but both are right skewed. Table 5.1.4 gives results of testing for normality, log-normality and skewness.

Table 5.1.4. TA-AIDS; a/o f/d analyses by subgroup.

Sub-group	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
A	911	Normality	0.977	0.998	0.971	<0.001
B	2570	"	0.964	0.999	1.586	<0.001
A	911	Log-normality	0.972	0.998	-1.056	<0.001
B	2570	"	0.939	0.999	-1.738	<0.001

Neither subgroup is compatible with normality or log-normality; for both, right skew is too extensive for normality but not extensive enough for log-normality, log-transformation resulting in a left skew.

Examination of comparable f/d analyses of data for transfusion cohorts, for individual years 1979-1985 for subgroup A (complete data not shown) gives, for each year, an f/d compatible with normality, and for only 1979 an f/d also compatible with log-normality. P values for significance of skew of the log-transformed data are: 0.089 (1979); 0.013 (1980); < 0.001 (1981 - 1985).

5.2 Discussion.

5.2.1 Mean Incubation period estimates.

Mean i/p for single and multiple transfusion cases are similar. However, with a traced suspect donor, the mean is much reduced. A donor is more likely to be traced for a recent transfusion, and a short i/p case will have a more recent transfusion at the time of onset, resulting in bias. The effect of left and right censoring on mean i/p is clearly demonstrated.

5.2.2 Incubation Period Frequency Distribution.

Sartwell (1950, 1966), and others using his method, have examined either outbreaks of acute infectious disease, where case ascertainment was as complete as possible, or collections of data on sporadic disease, gathered into one dataset for analysis. In either case, with large enough numbers, it is assumed that the observed range of i/p's would be relatively complete.

During a major epidemic of a newly observed disease with long i/p, relatively complete case ascertainment cannot occur until after the end of the epidemic, probably many years into the future. Neither can there be confidence in complete case ascertainment for earlier subgroups, for example those infected in a particular year, since it is unknown how long it will be before the last case derived from that cohort will occur; this is particularly the situation early in the epidemic.

Once the true maximum possible i/p from the earliest cohort has passed, the maximum i/p observed should tend to stabilise for subsequent infection cohorts (if numbers per cohort are large), which will give an indication that this stage has been reached. Unfortunately, the AIDS dataset used here covers a period prior to demonstrable i/p f/d stability by infection cohort.

Results indicate that as the year of transfusion recedes, the resultant f/d approximates more closely to log-normality; a wider range of i/p's has been observed for the early transfusion years;

the effect of left censoring appears to be less severe than that of right censoring; and i/p range stability in any pair of sequential years is not yet demonstrable. It is therefore impossible to say whether the theoretical maximum possible i/p has been observed.

Therefore, due to censoring, the dataset is virtually certain to be incomplete, and highly unlikely to be representative of the full i/p f/d. Although this dataset gives an i/p f/d which is less right skewed than is compatible with log-normality, it is not possible to verify, or refute, the hypothesis that the complete i/p f/d is a log-normal curve, or any other particular type of f/d. The range may well be incomplete.

5.2.3 Mean Incubation Period; comparison of back-calculation with direct estimates.

Taking into account the effect of censoring dramatically increases the estimated mean and maximum i/p, particularly if assuming a log-normal f/d. The main difference in the three f/d's is in the frequency and extent of cases in the right 'tail'. For the model illustrated here, if Sartwell's hypothesis is correct, infection with HIV means a lifetime at risk of AIDS.

CHAPTER 6

SHEEP: DATA ANALYSIS AND DISCUSSION: Natural and experimental TSE/prion disease.

6. Summary.

6.1 Natural scrapie in sheep.

6.1.1 Data from Parry (1962).
6.1.2 Data from Wooldridge (1991).
6.1.3 Data from VIDA.
6.1.4 Data from general literature search.
    6.1.4.1 Analysis of age at onset for full dataset.
    6.1.4.2 Analysis of age at onset by exposure status of dam.
    6.1.4.3 Analysis of incubation period.

6.2 Experimental TSE in sheep.

6 2.1. Full experimental dataset.
    6.2.1.1 Analysis of incubation period.
    6.2.1.2 Variables within the full experimental data set; analysis of variance (ANOVA).
6 2.2 Subgroup 1: Experimental CNS inoculation of CNS tissue.
    6.2.2.1 Analysis of incubation period.
    6.2.2.2 Variables within subgroup 1; further analysis.
6 2.3 Subgroup 2: Infection by sub-cutaneous inoculation.
    6.2.3.1 Analysis of incubation period.
    6.2.3.2 Variables within subgroup 2; analysis of dose.

6.3 Discussion.

6 3.1 Age at onset and incubation period frequency distribution analysis findings.
    6.3.1.1 Natural cases.
    6.3.1.2 Experimental cases.
6.3.2 Effect of variables
    6.3.2.1 Natural cases
    6.3.2.1 Experimental cases
6.3.3 Culling.
6.3.4 Comparison with Sartwell's conclusions.

CHAPTER 6

SHEEP: DATA ANALYSIS AND DISCUSSION

Natural and experimental TSE/prion disease.

Summary.

In the absence of heavy culling or data truncation, and in the presence of good flock observation and recording, natural scrapie age at onset (a/o) frequency distribution (f/d) may have a very elongated right skew, more extreme than that of log-normality, but this is rarely observed in practice. Without a test for infection, it is impossible to be certain that an exposed sheep will not succumb to disease, whatever its age. Culling is likely to seriously affect only the right extremity of the f/d, but in doing so may reduce the observed right skewness.

From both the a/o and incubation period (i/p) datasets, there is some indication that sample size is associated with the observed f/d; smaller datasets may tend to be less right skewed than larger sets.

Examination of variables in experimental TSE indicates that passage status, preceding donor species, infection route, and infecting substance affect mean i/p. The i/p f/d of the full experimental dataset is very unlike any other f/d, for either i/p or a/o; however, if the modal values for the subcutaneous (s/c) and CNS inoculated cases of scrapie are aligned, there is an approximation to a/o f/d of the natural disease. For CNS inoculated cases, i/p f/d has a right skew too extreme for log-normality.

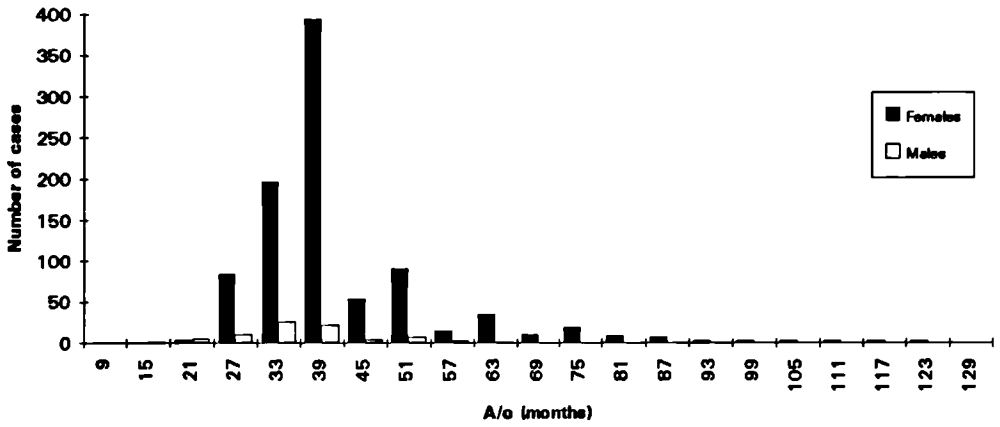
Whilst findings broadly agree with those of Sartwell, interpretation and conclusions differ. For natural or experimental scrapie, log-normal i/p f/d cannot be assumed; the underlying i/p f/d may be considerably more right skewed than that, although not necessarily the observed f/d.

6.1 Natural Scrapie in Sheep.

6.1.1 Data from Parry (1962).

The histogram of the a/o f/d (in months) is shown by sex (figure 6.1.1). (See appendix 10 for all sheep data source and categorisation details, unless otherwise indicated).

Figure 6.1.1 Natural sheep scrapie; Parry (1962) data; a/o f/d by sex.



Both male and female data appears right skewed. For analysis results see tables 6.1.1(i) and 6.1.1(ii).

Table 6.1.1(i) Natural sheep scrapie (Parry, 1962); a/o data (months).

Subgroup	N	Mean a/o	SD*	Range	Mode	Median
Female	930	42.0	13.7	21-123	39	39
Male	78	36.2	8.9	15-63	33	33

[* Sheppard's correction applied]

Females have a mean a/o at 3-4 years but a maximum of over ten years. For males, mean a/o at just over 3 years is less than that for females ($P < 0.001$, non-parametric test) and maximum recorded age is just over 5 years. Sample sizes for males and females are very different.

Table 6.1.1(ii) Natural sheep scrapie (data from Parry, 1962): f/d analysis.

Subgroup	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
Females	930	Normality	0.900	.998	2.366	<0.001
Males	78	"	0.990	.984	0.498	0.063
Females	930	Log-	0.972	.998	1.143	<0.001
Males	78	normality	0.990	.984	-0.433	0.103

For females, f/d is incompatible with both normality and log-normality; both distributions are highly skewed to the right. Conversely, for males neither the hypothesis of normality nor log-normality can be rejected.

6.1.2 Data from Wooldridge (1991).

The histogram of a/o f/d for the 1334 cases (not shown) has an obvious right skew. The modal a/o, at 2 years (24-35 months) is very marked, and less than Parry's whole set modal a/o of 39 months. F/d analysis results are given in table 6.1.2.

Table 6.1.2 Natural sheep scrapie (Wooldridge, 1991); f/d analysis.

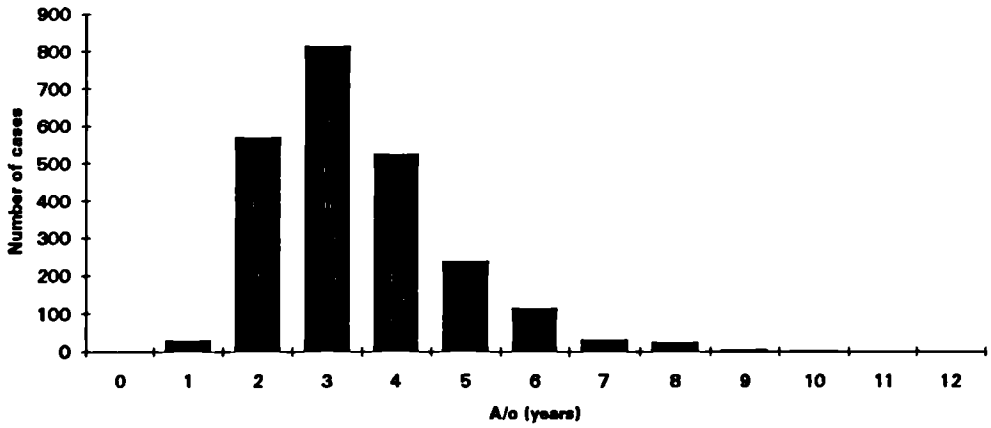
N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
1334	Normality	.987	.999	0.807	<0.001
1334	Log-normality	.999	.999	0.027	0.686

The hypothesis of normality is rejected, a right skew being present. However, f/d is compatible with log-normality; log-transformed data has no marked skew.

6.1.3 Data from VIDA.

Figure 6.1.3 shows the histogram of the a/o f/d (in years) for 2361 cases.

Figure 6.1.3 Natural sheep scrapie; VIDA data; a/o f/d.



A smoothly curved right skew is apparent. The mode is at 3 years, and cases are still being recorded at 10 years, very close to the maximum age to which any sheep are likely to be kept. A/o data (in years) for the 2361 cases:-

Mean :	**3.43**
SD:	**1.33**
Range:	**1-10**
Mode:	**3**
Median:	**3**

F/d analyses are given in table 6.1.3.

Table 6.1.3 Natural sheep scrapie (VIDA); f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
2361	Normality	0.979	.999	1.093	<0.001
2361	Log-normality	0.998	.999	-0.054	0.280

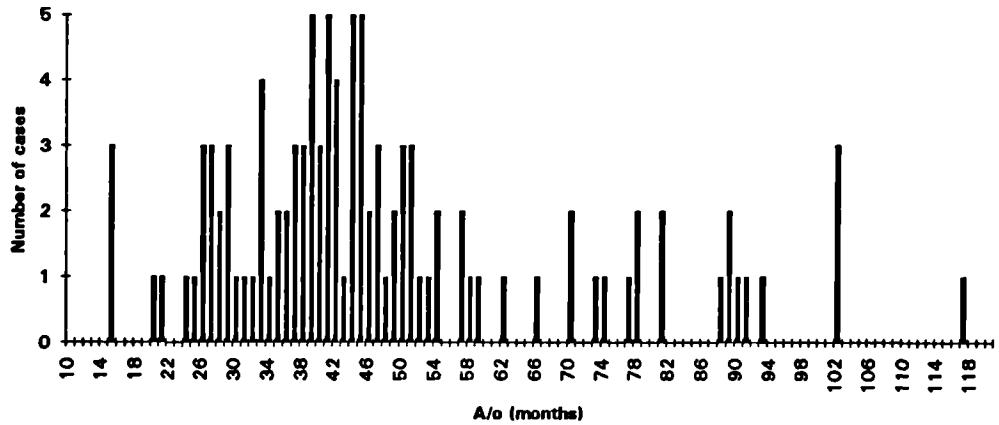
The hypothesis of normality is rejected; there is a marked right skew. Log-transformation of the data results in an insignificant skew, although data just fails to be compatible with log-normality. However, with large samples a small deviation may become significant.

6.1.4 Data from general literature search (GLS).

6.1.4.1 *Analysis of age at onset for full dataset.*

The histogram of a/o f/d (in months) for 'smoothed' data (see appendix 10 for smoothing method) for 107 cases is shown (figure 6.1.4.1).

Figure 6.1.4.1 Natural sheep scrapie; GLS data; a/o f/d.



The histogram is not symmetrical, and appears to be right skewed. A/o data (in months) for the 107 cases:-

Mean : 48.0
SD: 21.1
Range: 15-117
Modal region: 39-45
Median: 43

A/o f/d analysis results are given in table 6.1.4.1.

Table 6.1.4.1 Natural scrapie in sheep (GLS): f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
107	Normal	0.949	0 987	1.147	<0.001
107	Log-normal	0.989	0 987	0.015	0.948

The a/o f/d is incompatible with normality but compatible with log-normality. (Results and conclusions are similar for unsmoothed data, not shown).

6.1.4.2 Analysis of age at onset by exposure status of dam.

'Smoothed' a/o data by 'pregcode' status (see appendix 10 for definition of 'pregcode' categories) is given in table 6.1.4.2(i).

Table 6.1.4.2(i) Natural sheep scrapie (GLS); a/o data (in months) by exposure status (pregcode) of dam.

Status of dam	N	Mean a/o	SD	Range	Modal region	Median
Pregcode = 1	55	38.3	14.9	15-102	14-44	37
Pregcode = 0	52	58.3	21.9	27-117	39-45	50

For 'pregcode = 1', mean a/o is 20 months less than that for 'pregcode = 0' ($P < 0.001$, non-parametric test). There is also a difference in modal regions by pregcode. The histogram (not shown) indicates an extended right skew for both groups. A/o f/d analysis results are given in table 6.1.4.2(ii).

Table 6.1.4.2(ii). Natural sheep scrapie (GLS); a/o f/d by exposure status of dam.

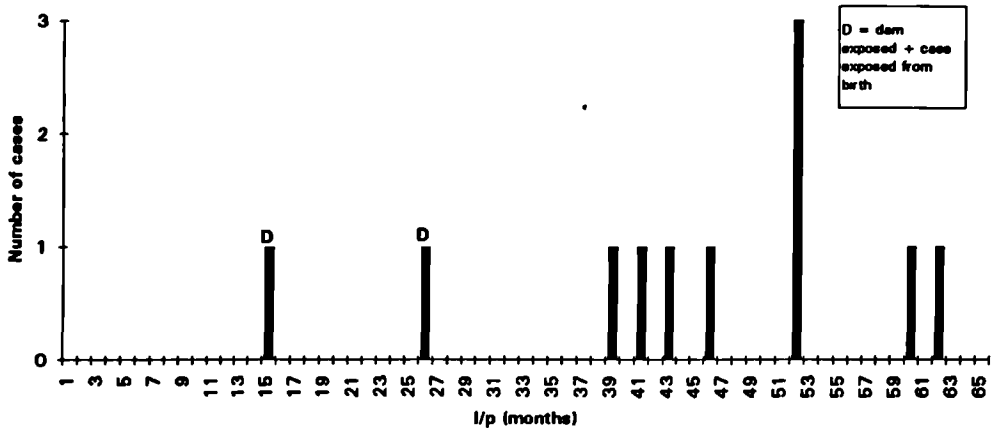
Status of dam	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
Pregcode = 1	55	Normal	0.938	0.978	1.627	<0.001
Pregcode = 0	52	"	0.949	0.977	0.818	0.015
Pregcode = 1	55	Log-	0.984	0.978	-0.143	0.633
Pregcode = 0	52	normal	0.976	0.977	0.356	0.254

For both groups the hypothesis of normality is rejected; c/s is large and positive. The hypothesis of log-normality is rejected only for 'pregcode = 0', although log-transformation of data results in an insignificant skew for both groups.

6.1.4.3 Analysis of incubation period.

The histogram of 'maximum i/p' (see appendix 10 for definition) f/d (in months) for 11 cases is shown (figure 6.1.4.3(a)).

Figure 6.1.4.3(a). Natural sheep scrapie; GLS data; 'maximum i/p' f/d.

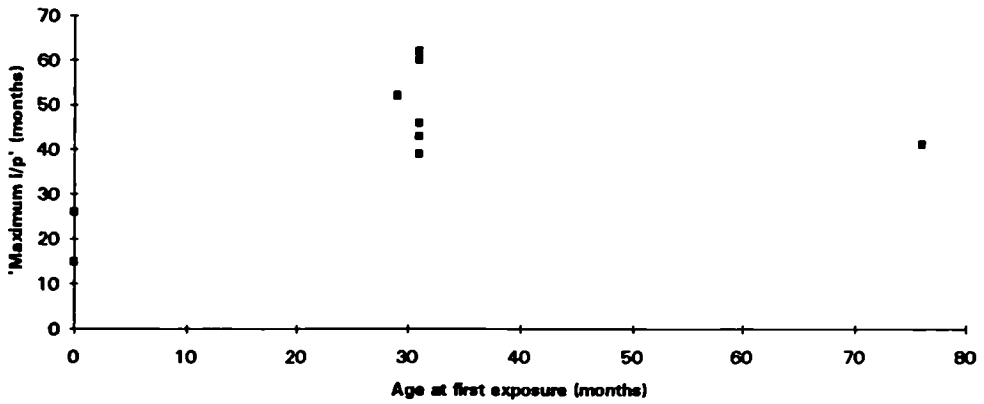


A left skew is apparent but small sample size makes interpretation difficult. 'Maximum i/p' data (in months):-

Mean :	**48.4**
SD:	**14.1**
Range:	**15-62**
Modal region:	**52**
Median:	**46**

The two cases with maternal exposure recorded (marked D in the histogram) had by far the shortest 'maximum' i/p's (and a/o's) of 15 and 26 months. The relationship between age at first recorded exposure and 'maximum i/p' is plotted in figure 6.1.4.3(b).

Figure 6.1.4.3(b). Natural sheep scrapie; GLS data; 'maximum i/p' by age at 1st recorded exposure.



The two cases exposed from birth had a shorter 'maximum' i/p (mean: 20.5 months) than cases exposed when older (mean: 49.7 months). However, the relationship is not obviously linear, and numbers are small. 'Maximum i/p' f/d analysis results are given in table 6.1.4.3.

Table 6.1.4.3. Natural sheep scrapie (GLS); 'maximum i/p' f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
11	Normal	0.968	0.940	-0.796	0.154
11	Log-normal	0.903	0.940	-1 497	0.012

In agreement with the graph, analysis gives a negative coefficient of skewness, although numbers are small and it is not possible to reject the hypothesis of normality. However, the hypothesis of log-normality is rejected; log-transformation of the data gives a marked left skew.

Ten cases with recorded exposure also had recorded a/o. In most flocks, age at exposure would not be known, therefore overall mean a/o (72.5 months; SD 30.5) would be the relevant statistic. Here, however, the two cases with exposure from birth have a/o's below 30 months (mean: 20.5, SD 15.0). The remainder were not exposed until after this age, and have a mean a/o of 85.5 months (SD 15.0).

6.2 Experimental TSE in sheep.

6.2.1 Full experimental dataset.

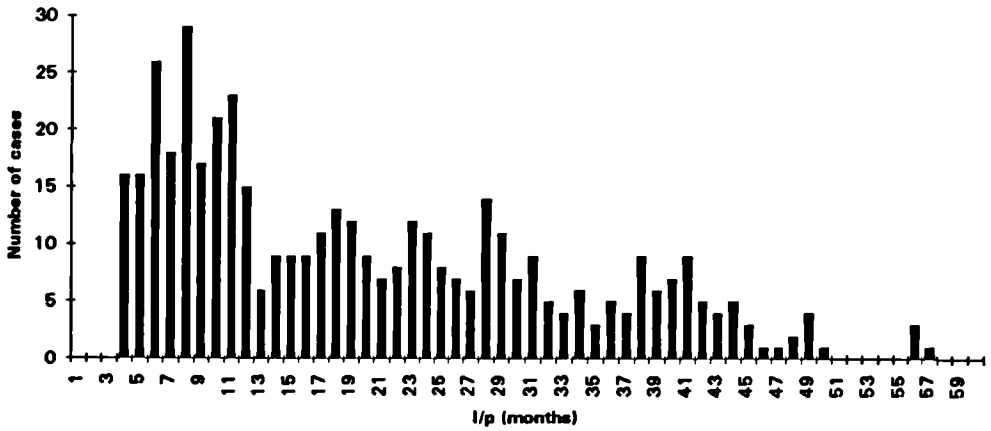
6.2.1.1 Analysis of incubation period.

I/p data (in months) for the total 477 cases:-

Mean i/p :	**20.0**
SD:	**12.7**
Range:	**4-57**
Mode:	**8**
Median:	**17**

The i/p f/d histogram (in months) is shown (figure 6.2.1.1).

Figure 6.2.1.1 Experimental TSE in sheep; full dataset; i/p f/d.



After a rapidly attained modal region, there is an extensive relatively high frequency, uneven right tail. The i/p f/d analysis results are given in table 6.2.1.1.

Table 6.2.1.1 Experimental TSE in sheep; full dataset; i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
447	Normal	.966	.997	0.682	<0.001
447	Log-normal	.984	.997	-0.200	0.082

Data is incompatible with normality, having a marked right skew. Although the log-transformed data is not statistically significantly skewed, CC is less than C/v, possibly due to kurtosis; therefore the dataset is also incompatible with log-normality.

6.2.1.2 Variables within the full experimental data set; analysis of variance (ANOVA).

Stratification by variables recorded reduces each stratum to a very small number, therefore further analysis uses ANOVA. After preliminary analyses (results not shown) four factors, passage status, preceding donor, route of infection, and infecting substance are analysed. Two cases from monkey donors, and three cases derived from dam inoculation are omitted; they can not reasonably be grouped with anything else. The final dataset comprises 442 cases, grouped as follows:-

Preceding donor: sheep (427), other ruminants (goats & cattle) (15)
Passage status: from natural case (256), from experimental case (186)
Infection route: CNS inoculation (119), s/c inoculation (98), other routes, including mixed (225).
Infecting substance: CNS (274), CNS + lymphoid mix (124), spleen (20), other tissues (24).

Log-transformation of i/p data is necessary to normalise residuals. Significance is attributed to each of the factors. Results, table 6.2.1.2.

Table 6.2.1.2 Experimental TSE in sheep; full dataset; ANOVA results.

Variable being examined	Variable grouping	N	F	P	Mean i/p (adj; geometric)	95% CI
Passage status	From natural case	256	99.91	<0.001	27.0	23.4-31.2
	From experimental case	186			13.9	11.8-16.4
Preceding donor	Sheep	427	5.60	0.018	16.8	15.5-18.2
	Other ruminant	15			22.3	17.0-29.1
Infection route	CNS inoculation	119	75.78	<0.001	13.2	11.4-15.4
	S/c inoculation	98			21.8	18.2-26.0
	Other routes/mixed	225			25.2	21.8-29.3
Infecting substance	CNS	274	14.36	<0.001	14.7	13.0-16.5
	CNS + lymphoid: mix	124			18.1	15.5-21.1
	Spleen	20			19.1	14.9-24.5
	Other tissues	24			27.9	22.2-35.2

There is no significant interaction between status and donor (F=0.05, P=0.832) or between status and infection route (F=0.36, P = 0.701). Analysis of other interactions failed due to lack of data.

Each factor examined appears to have a significant effect on the i/p; an assumption of constancy for each of the other factors gives differences in the mean (adjusted) i/ps as shown.

6.2.2. Subgroup 1: Experimental CNS inoculation of CNS tissue.

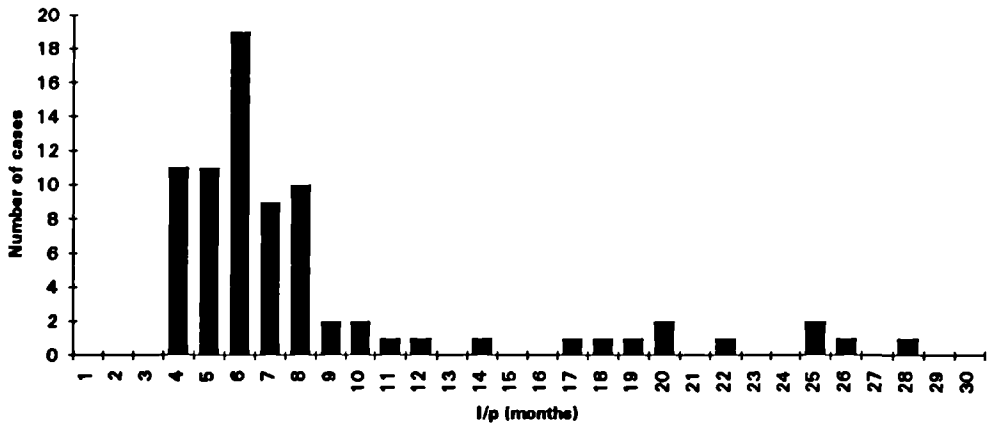
6.2.2.1 Analysis of incubation period.

Details of i/p's (in months) for 77 cases:-

Mean i/p :	**8.5**
SD:	**5.7**
Range:	**4-28**
Mode:	**6**
Median:	**6**

The histogram of i/p f/d is shown (figure 6.2.2.1) and f/d analysis results given in table 6.2.2.1.

Figure 6.2.2.1 Experimental scrapie in sheep; subgroup 1; i/p f/d.



A marked right skew is apparent, but the frequency drops sharply from the mode, hence the right tail is of low frequency.

Table 6.2.2.1 Experimental scrapie in sheep; subgroup 1; i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
77	Normal	0.853	0.983	2.008	<0.001
77	Log-normal	0.951	0.983	1.167	<0.001

The hypotheses of both normality and log-normality are rejected, the right skew persisting even in log-transformed data.

6.2.2.2 Variables within subgroup 1; further analysis.

Variable: Passage status.

The i/p data (in months) by passage status is given in table 6.2.2.2.

Table 6.2.2.2 Experimental scrapie in sheep; subgroup 1; i/p data (months) by passage status.

Passage status	N	Mean i/p	SD	Range	Mode	Median
From natural case	30	10.3	6.8	5-26	6	7
From experimental case	47	7.3	4.6	4-28	4	6

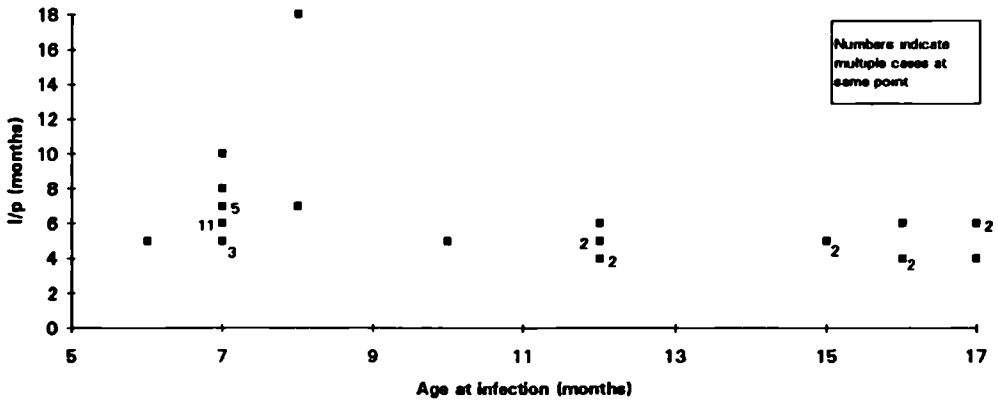
The mean i/p for cases infected directly from natural cases is longer than for those infected from experimental cases (P=0.020; non-parametric test).

On the histogram (not shown) there is no obvious difference between the two groups; both appear similar to complete subgroup 1. F/d analysis results (not shown) are also similar to subgroup 1; although there is a difference in mean i/p, there is no difference in conclusions regarding the f/d.

Variable: Age at Infection.

Figure 6.2.2.2 plots i/p and age at infection (all in months) for 38 cases.

Figure 6.2.2.2 Experimental scrapie in sheep; subgroup 1; i/p by age at infection



Although difficult to see graphically, regression analysis indicates a slight reduction of i/p with increasing age at infection, of 0.20 months per month (r = -0.32, F = 4.13, P = 0.050).

Removal of the outlier (at i/p = 18 months) gives a much reduced reduction of only 0.05 months per month (r=-0.18, F=0.46; P=0.511), but based on the skewed pattern of i/p data in this and other data sets, there is no biological rationale for removing this outlier. There is no reason to assume that it is an erroneous result

Variable: Sheep breeds.

For 25 cases where donor and case were recorded as the same breed, mean i/p was 7.5 months (SD = 5.8), and for the 50 cases where they were not so recorded, mean i/p was 8.8 months (SD = 5.8) (P=0.028, non-parametric analysis). Although one way ANOVA using the reciprocal of the i/p did not normalise the residuals, it gave the closest visual approximation (F=4.49, P= 0.038). However, neither test is particularly appropriate, as both make the assumption that the distributions are similar, which was not the case.

Analysis of variables: ANOVA

Earlier analyses indicated that age at infection (as a covariate), passage status and case-breed/donor-breed combination might affect i/p. Multiway ANOVA was therefore attempted using these variables, for 38 cases with the same dose and known age at infection. Using raw i/p data, residuals were not normally distributed. Neither log-transformation, nor taking the reciprocal of the i/p normalised the residuals; log-log transformation was slightly more successful. In all cases the data gave no statistical significance to any variable.

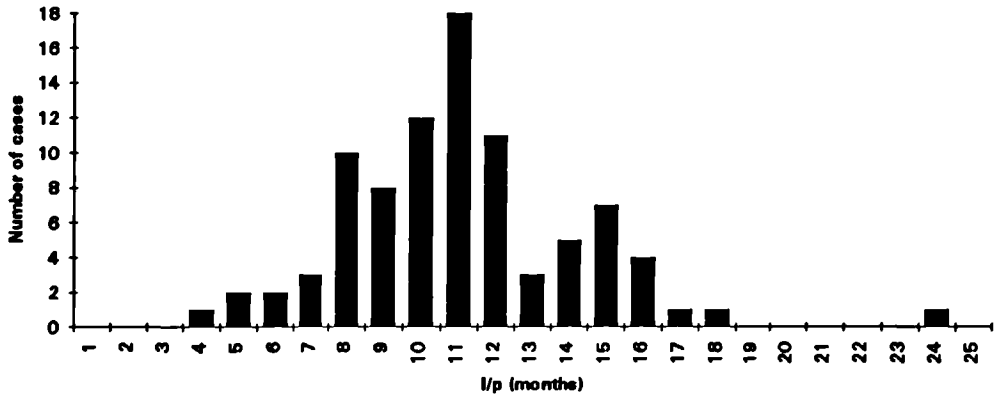
Attempts to reclassify variables for further analysis were unsuccessful, tending to remove the ability to discriminate between factors of interest. Interpretation of the above is therefore very difficult, incomplete data and small numbers being the main problems. Conclusions on the effect of these variables have therefore not been reached.

6.2.3 Subgroup 2: Infection by sub-cutaneous inoculation.

6.2.3.1 *Analysis of incubation period.*

Eighty-nine cases each receiving 2 mls of brain extract are used in i/p f/d analysis. The histogram of the i/p f/d is shown (figure 6.2.3.1) and f/d analysis results given in table 6.2.3.1.

Figure 6.2.3.1 Experimental scrapie in sheep; subgroup 2 (2 mls of extract); i/p f/d.



Inspection of the graph suggests a slight, but not pronounced, right skew.

Table 6.2.3.1 Experimental scrapie in sheep; subgroup 2 (2 mls of extract); i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
89	Normal	0.980	0.985	0.751	0.005
89	Log-normal	0.986	0.985	-0.511	0.044

Data is incompatible with normality but compatible with log-normality, although for the log-transformed data the c/s is just within the boundary of left-skew statistical significance.

6.2.3.2 Variables within subgroup 2; analysis of dose.

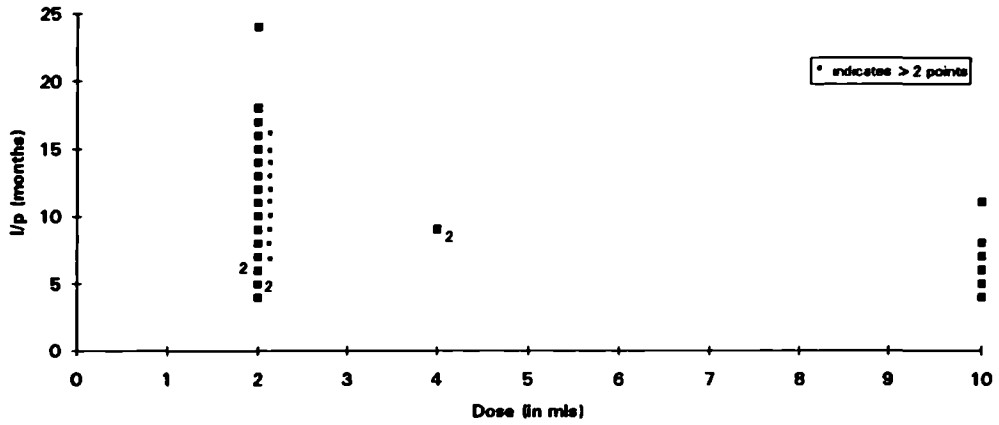
I/p details (in months) for 97 cases by dose are given in table 6.2.3.2.

Table 6.2.3.2. Experimental scrapie in sheep; subgroup 2; i/p data (in months) by volume of s/c inoculation.

Dosage (mls) of brain extract	N	Mean i/p	SD	Range	Mode	Median
2	89	11.1	3.2	4-24	11	11
4	2	9	0.0	-	-	-
10	6	6 8	2.5	4-11	4	6.5

Although the case numbers for doses of 4 and 10 mls are small the mean i/ps suggest that as dosage increases, i/p decreases. Figure 6.2.3.2 plots this relationship.

Figure 6.2.3.2 Experimental scrapie in sheep; subgroup 2; i/p by dose.



Regression analysis gives a decrease in i/p with increasing dosage of 0.54 months per ml ($r = -0.32$, $F=10.93$, $P = 0.001$), but sample size for 2 ml is much greater than for 10 mls, and the minimum i/p is the same in each case.

6.3 Discussion.

6.3.1 Age at onset and incubation period frequency distribution analysis findings.

6.3.1.1 Natural cases.

Summary of f/d analysis results is given in table 6.3.1.1(i).

Table 6.3.1.1(i). Natural sheep scrapie; summary of f/d analyses.

Data set	A/o or i/p	N	F/d compatible with:-		Skew for:-	
			Normality	Log-normality	Raw data	Logged data
Parry: females	a/o	930	No	No	R	R
Parry: males	"	78	Yes	Yes	NS	NS
Wooldridge	"	1334	No	Yes	R	NS
VIDA	"	2361	No	No**	R	NS
GLS: all cases	"	107	No	Yes	R	NS
GLS: pregcode = 1	"	55	No	Yes	R	NS
GLS: pregcode = 0	"	52	No	No**	R	NS
GLS: known exposure	i/p	11	Yes	No	NS	L

[Notes on table:- 1) Significance level for both f/d compatibility with normality and log-normality, and significance of skew is at the P=0.05 level.
2) ** indicates CC is equal to, or within 0.001 of c/v]

Table 6.3.1.1(ii) describes the f/d for each dataset. The cull guide is a very approximate estimate of how much culling (or artificial data truncation) is suspected, particularly regarding maximum longevity of the flock or maximum age recorded. A high number of asterisks suggests early, heavy culling or data truncation; no asterisks means no information is available.

Table 6.3.1.1(ii) Natural sheep scrapie; descriptions of f/d.

Data set	A/o or i/p	Description	Cull guide
Parry: females	a/o	Right skew; incompatible with log-normality	*
Parry: males	"	Compatible with normal and log-normal	****
Wooldridge	"	Compatible with log-normal	**
VIDA	"	Incompatible with log-normal (skew compatible)	***
GLS: all cases	"	Compatible with log-normal	
GLS: pregcode = 1	"	Compatible with log-normal	
GLS: pregcode = 0	"	Incompatible with log-normal (skew compatible)	
GLS: known exposure	i/p	Left skew	**?

Females comprise over 92% of Parry's dataset. They are the dataset most representative of 'natural' life-span and they have very marked right skew, more extensive than expected for log-normality. Nevertheless even here, culling has occurred. For males, with early heavy culling, right truncation is expected to reduce the right skew, and this is observed.

Wooldridge's data is automatically truncated at five years due to survey design. That this is less right skewed than Parry's females is not therefore surprising. A complete description of the right tail would probably increase right skewness. Since this data was all collected at the end of the survey period, unlike Parry's, reduction in modal a/o may reflect farmers recall bias, an effect of grouping animals by management factors, i.e. as adults from 2 years.

For VIDA recording, there is no maximum age, therefore no artificial data censoring. However, VIDA data is 'VIC submitted', more likely for males than females - and males generally suffer earlier culling. This, plus other factors (e.g. value, atypical age) bias VIC submissions (Wooldridge, 1991; Wooldridge et al, 1992). Therefore VIC cases may exclude the oldest animals, thus reducing right skew. Compared with Parry's females, this is observed. Converting VIDA data into months (i.e. x 12) gives a similar mean to Parry's full dataset, with a/o's of 41.2 and 41.6 months respectively. However, since Parry's data was recorded in 6 monthly intervals, and VIDA data grouped annually, methods are not identical.

One conclusion from this is that in the absence of heavy culling or data truncation of the right extremity of the f/d, and in the presence of good flock observation and recording (e.g. Parry's females), observation of scrapie onset is possible well into a sheep's old age. In these circumstances the a/o f/d may have a very elongated right skew, more extreme than that of log-normality, but this appears to be rarely observed in practice.

For cases aggregated from the literature, no overall comments regarding culling are applicable, although there is likely to be a reporting bias. Some cases, although 'naturally transmitted', were observed on experimental farms, where conditions may be atypical.

The small i/p dataset is left skewed; truncation of the right tail by culling or flock dispersal is possible (Greig, 1940). However, there is a slight impression that sample size may be associated with skewness. Only one of the large datasets is extremely right skewed, and only the smallest is left skewed. Other datasets examined (both a/o and i/p) but not included in this thesis tend to reinforce this impression.

6.3.1.2 Experimental cases.

Summary of i/p f/d analysis results are given in table 6.3.1.2(i)

Table 6.3.1.2(i). Experimental TSE in sheep; summary of f/d analyses.

Data set	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Logged data
Full set	447	No	No	R	NS
Subgroup 1	77	No	No	R	R
Subgroup 2	89	No	Yes**	R	L

[Notes on table:- 1) Significance level for both f/d compatibility with normality and log-normality, and significance of skew is at the P=0.05 level
2) ** indicates CC is equal to, or within 0.001 of c/v]

Table 6.3.1.2(ii) describes the f/d for each dataset.

Table 6.3.1.1(ii) Experimental TSE in sheep; descriptions of f/d.

Data set	Description
Full set	Incompatible with log-normal (skew compatible)
Subgroup 1	Right skew too elongated even for log-normality
Subgroup 2	Skew borderline between normal and log-normal

Most work was published before the mid-seventies. Greater danger of cross-contamination was possible in early work, since the scrapie agent's extreme resistance to inactivation became apparent only gradually.

Provided all animals in a given experiment are observed from infection to disease or death, culling should not be relevant, although a high incidence of death from other causes would have a similar effect. In fact there was evidence of this in one experiment (Dickinson et al, 1968) where i/c inoculation resulted in high death rates due to bacterial contamination. It is likely that all experimental animals were subjected to better observation and recording than the average flock, although not necessarily better than Parry's flocks.

In the natural TSE disease situation, for any particular case where infection is suspected, the specific infection source is usually unknown (for scrapie, it is usually assumed to be sheep but

other possibilities have not been disproven; for example goats, bovines, environmental, or other unsuspected sources). Infection route is also generally undetermined, although often assumed to be peripheral. The possibility of *de novo* spontaneous generation of agent in the brain is not disproven, and i/p may then approximate roughly to experimental CNS inoculation of agent. The proportions of natural cases due to each of these potential aetiologies is unknown. There is thus no legitimate reason for differentiating between experimental subsets in the first analysis.

However, the i/p f/d of the full dataset is very unlike any other f/d, either i/p or a/o, due to the high variable frequencies in the extensive right tail. It appears unlikely that this particular composite is a close reflection of any natural scrapie situation. Since the proportion infected intracerebrally, and the proportion infected by CNS tissue were high, lack of overall similarity to the findings in any natural scrapie disease situation is not surprising.

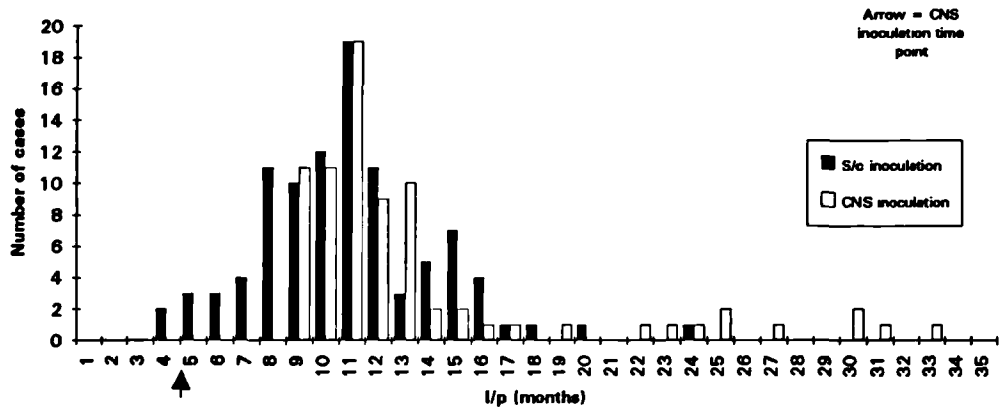
For subgroup 1 (CNS to CNS transmission), i/p f/d is too right skewed for compatibility even with log-normality, stratification by passage status giving similar results. This extensive skew implies that even when the TSE aetiological agent is introduced directly into the CNS, there is still extensive variation in the subsequent time to disease onset, despite the marked, rapidly attained modal i/p. Clearly, other variables are involved (perhaps particularly genotype). This suggests that by whatever route natural transmission occurs, once the agent reaches the CNS (or in some cases is perhaps spontaneously generated *de novo* in the CNS) a variable, and occasionally very lengthy i/p is still possible.

For s/c inoculation, a peripheral route perhaps more comparable to natural transmission (which might be via wounds, mucous membranes, cornea, naso-pharyngeal, or alimentary systems) the i/p f/d is less right skewed than that for CNS inoculation. It appears to lie between normal and log-normal distributions. Inspection of the histogram suggests that this is at least partly due to the longer duration and wider distribution of i/p values prior to the modal i/p, as compared with i/c inoculation. Again other sources of variation are involved, as they would be in the natural situation.

Peripheral transmission routes may therefore result in a wide variety of possible time intervals before the agent reaches the CNS, perhaps associated more with the particular entry point rather than genotype. If this variability in time taken for the agent to reach the CNS is superimposed upon the wide variety of possible times to disease once the agent is in the CNS (which as has been seen is right-skewed more extensively than log-normal), the result will be an even wider range of possible i/ps, reflected in a wide range for a/o, even if all infection occurs at the same age.

Looking back to the full experimental dataset, the modal region of the CNS to CNS inoculated cases can be identified, in advance of other transmission combinations. One possible hypothesis is that if the modal values for the s/c and CNS inoculated cases are aligned, the overall f/d corresponds more closely to the observed f/d in a natural disease situation. Here the assumption is that the majority of CNS inoculated cases would have taken on average an additional 5 months (the difference in modal values), to reach their modal value if they had been peripheral inoculations (perhaps the more probable natural transmission route), although a proportion may be considered to represent the possibility of *de novo* agent generation in the CNS. Figure 6.3.1.2 illustrates this manoeuvre.

Figure 6.3.1.2 Experimental scrapie in sheep; CNS and s/c inoculation: superimposition of modal points



There is an approximation to a/o f/d of the natural disease situation, although the artificial composite i/p mode, at 11 months, is much shorter than the natural a/o mode, generally at 3-4 years. One possible explanation would be that natural infection (whatever the route) is less efficient in terms of reaching the brain than s/c inoculation experimentally; another that natural infection occurs on average at 2 years of age.

The very great differences in experimental conditions preclude any meaningful comment on association between sample size and skew.

6.3.2 Effect of variables.

6.3.2.1 Natural cases.

The only additional variables which could be analysed for the natural cases were connected with probable age at exposure. Recorded exposure of the dam to scrapie reduces both mean a/o and mean i/p by about 20 months (see 6.1.4.2 and 6.1.4.3). Although numbers are small, figure 6.1.4.3(b) suggests a plateau for effect of exposure age on i/p. If all cases are transmission-induced, maternal contact cases possibly receive higher doses (i.e. from dam plus pasture), or are

infected pre-natally, and incubating at birth. Alternatively neonates may be more susceptible, contracting infection earlier in the exposure period; or once infected, pathological processes in neonates may evolve more rapidly, connected with immune system immaturity.

If there is an interaction between age at first exposure and subsequent i/p, then one also exists between a/o and length of preceding i/p, and hence time of infection. Unfortunately, no firm conclusions can be reached with such small numbers.

6.3.2.2 Experimental cases.

Examination of variables within the full dataset indicate that passage status, preceding donor, infection route, and infecting substance may have significant effects on mean i/p, although confounding by other variables is possible.

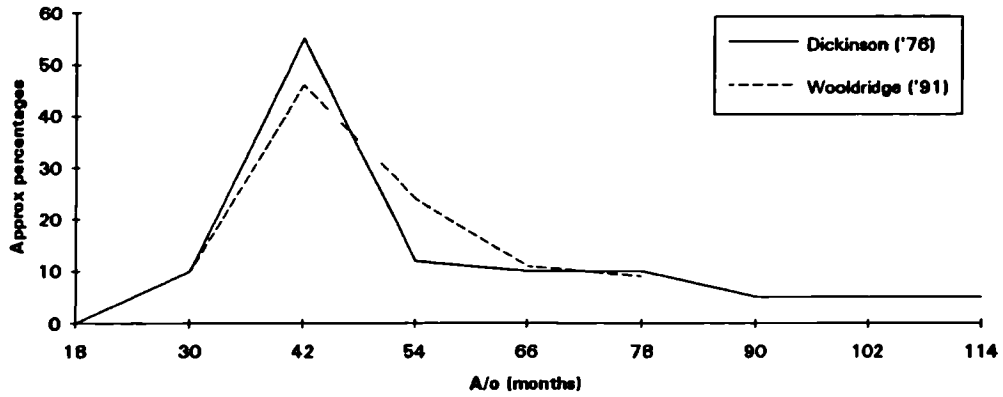
Within subgroup 1, passage status appears to affect mean i/p, perhaps due to confounding; perhaps because experimental passage in some way increases agent pathogenicity under certain circumstances. Results for age at infection, and breed combinations are somewhat equivocal. The effect of dose in shortening i/p (subgroup 2) appears more clear-cut, but with such a sample size disparity, the range of possible values in animals given 2 mls would be expected to be much greater, particularly in the longer i/p end of the 2 ml i/p f/d, as a right skew might be expected. The plot appearance tends to support this hypothesis, and may exaggerate any dose effect.

6.3.3 Culling.

Natural case data suggests culling is likely to reduce the a/o f/d right skew for observed scrapie in sheep. Anecdotal information is available on sheep culling, but numerical estimates are rare.

An average culling pattern can be estimated for the breeding flock up to five years of age based on a scrapie survey by Wooldridge (1991), and case numbers adjusted appropriately (see appendix 20). Dickinson (1976) plots a/o f/d in scrapie in Suffolk sheep, corrected for culling (method unknown), as proportions of scrapie cases by a/o in months. Figure 6.3.3 compares these, Wooldridge (1991) adjusted cases being converted to proportions by age, the a/o having been taken as the mid-point of each year of age, and converted to months.

Figure 6.3.3 Natural sheep scrapie; a/o f/d as percentage of cases by age, corrected for culling



Both estimates give a modal a/o at 3 to 4 years, the earliest a/o at 18 months, and a marked frequency reduction at around 54 months, although somewhat less for Wooldridge's data. Dickinson shows a sizeable (and incomplete) continuation of the tail beyond nine years, giving a coefficient of skewness of 1.143 for the 'truncated' graph. However, although scrapie cases may occur in older sheep, they appear to be much less likely after about 4-5 years of age. In conclusion, culling is likely to seriously affect only the right extremity of the f/d, but in doing so, may reduce the observed f/d right skewness.

6.3.4 Comparison with Sartwell's Conclusions.

For experimental TSE's in sheep, an i/p f/d with right skew too extreme for log-normality occurs in all 3 subgroup 1 datasets. Sartwell also noted the extreme right skew of some experimental datasets.

For a/o f/d of natural disease, this extreme right skew is also possible, as demonstrated by Parry's carefully recorded data. Incomplete case ascertainment, particularly of older cases, by culling, underestimating age for older animals, or biases in the data collected, may reduce observed skew such that f/d is frequently compatible with log-normality. However, although in an endemically infected flock all sheep may be exposed from soon after birth, the actual time of infection is unknown, making interpretation difficult. Nevertheless, an i/p f/d of log-normality cannot be assumed; it is possible that the underlying i/p f/d is actually considerably more skewed than that.

Whilst these findings broadly agree with those of Sartwell, the interpretation and conclusions differ.

CHAPTER 7

GOATS: DATA ANALYSIS AND DISCUSSION: Natural and experimental TSE/prion disease.

7. Summary.

7.1 Natural scrapie in goats.

7.1.1 Data from general literature search.

7.1.1.1 Analysis of age at onset for full dataset.

7.1.1.2 Analysis of age at onset by contact status.

7.1.1.3 Analysis of age at onset by age at last recorded exposure.

7.1.2 Data from VIDA.

7.2 Experimental TSE in goats.

7.2.1 Full experimental dataset.

7.2.1.1 Analysis of incubation period.

7.2.1.2 Variables within the full experimental data set; analysis of variance (ANOVA).

7.2.2 Subgroup 1: Experimental CNS inoculation of CNS tissue.

7.2.2.1 Analysis of incubation period.

7.2.2.2 Variables within subgroup 1; further analysis.

7.2.3 Subgroup 2: Cases with a plausible 'natural' infection route.

7.2.4 Subgroup 3: Cases infected directly from natural sheep scrapie.

7.3 Discussion.

7 3.1 Age at onset and incubation period frequency distribution analysis findings.

7.3.1.1 Natural cases.

7.3.1.2 Experimental cases.

7.3.1.3 Comparison of natural age at onset and experimental incubation period frequency distribution.

7.3.2 Effect of variables.

7.3.2.1 Natural cases.

7.3.2.2 Experimental cases.

7.3.3 Comparison with Sartwell's conclusions.

CHAPTER 7

GOATS: DATA ANALYSIS AND DISCUSSION

Natural and experimental TSE/prion disease.

Summary.

As for sheep, natural scrapie onset is possible well into a goat's old age; the age at onset (a/o) frequency distribution (f/d) may again have a very elongated right skew, more extreme than that of log-normality. Also as for sheep, there is an indication that sample size is associated with the observed f/d, with smaller datasets tending to be less right skewed than larger sets.

Similar conclusions are reached for the incubation period (i/p) f/d for experimental data. Examination of variables in experimental data indicate that, as for sheep, passage status, preceding donor, infection route and infecting substance effect mean i/p; age at infection, to 9 months of age, does not. Reduction of agent dose results in an i/p increase.

Overall, conclusions reached are therefore similar to those for sheep. Sartwell's hypothesis of underlying log-normality for i/p f/d is not confirmed; the underlying i/p f/d may be considerably more right skewed than that.

7.1 Natural scrapie in goats.

7.1.1 Data from general literature search.

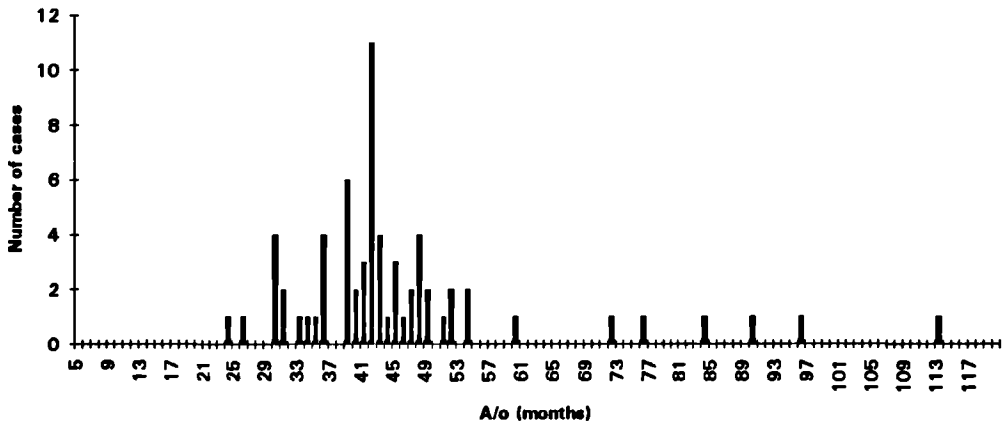
7.1.1.1 Analysis of age at onset for full dataset.

The a/o data (in months) for the 66 cases is as follows (see appendix 10 for all goat data source and categorisation details, unless otherwise indicated):-

Mean:	**45.6**
SD:	**15.8**
Range:	**24-113**
Mode:	**42**
Median:	**42**

There was no statistically significant difference in mean a/o by sex (P= 0.857, non-parametric test). The histogram of the a/o f/d (in months) for the 66 cases is shown (figure 7.1.1.1).

Figure 7.1.1.1. Natural goat scrapie; GLS data; a/o f/d.



An extensive marked right skew is apparent. The f/d analysis is given in table 7.1.1.1

Table 7.1.1.1 Natural goat scrapie (GLS); f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
66	Normality	0.865	0.980	2.267	<0.001
66	Log-normality	0.948	0.980	1.106	0.001

Both the hypotheses of normality and log normality are rejected. There is a significant right skew even for log-transformed data.

7.1.1.2 Analysis of age at onset by contact status.

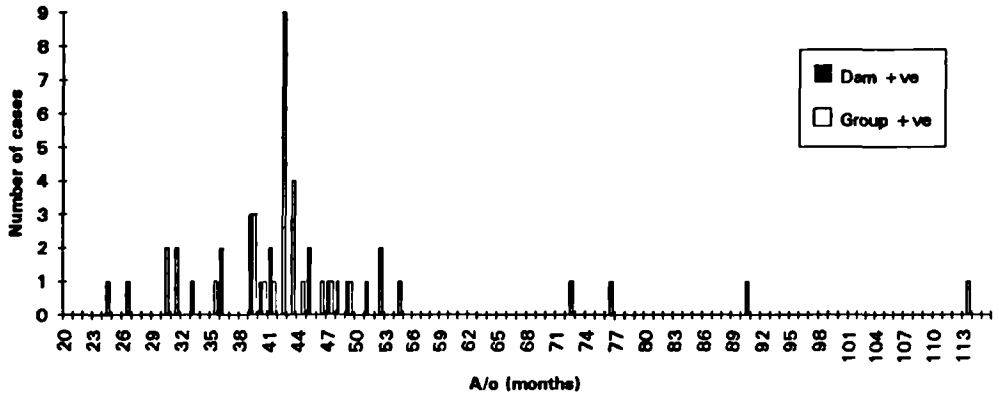
The a/o data (in months) by contact status given in table 7.1.1.2(a).

Table 7.1.2.2(i) Natural goat scrapie (GLS); a/o data (in months) by contact status.

Contact status	N	Mean a/o	SD	Range	Mode	Median
Dam scrapie +ve	15	41.9	5.3	33-51	39	41
Flock/herd scrapie +ve	36	45.8	17.3	24-113	42	42

There is no statistically significant difference in the mean a/o (P = 0.589 non-parametric test) by contact status, but ranges are different (see figure 7.1.1.2).

Figure 7.1.1.2 Natural goat scrapie; GLS data; a/o f/d by contact status recorded.



For dam-positive cases, a very small dataset, the f/d appears almost symmetrical whereas for flock/herd contact, it appears right skewed. The a/o f/d analyses are given in table 7.1.1.2(ii).

Table 7.1.1.2(ii) Natural goat scrapie (GLS); a/o f/d by contact status.

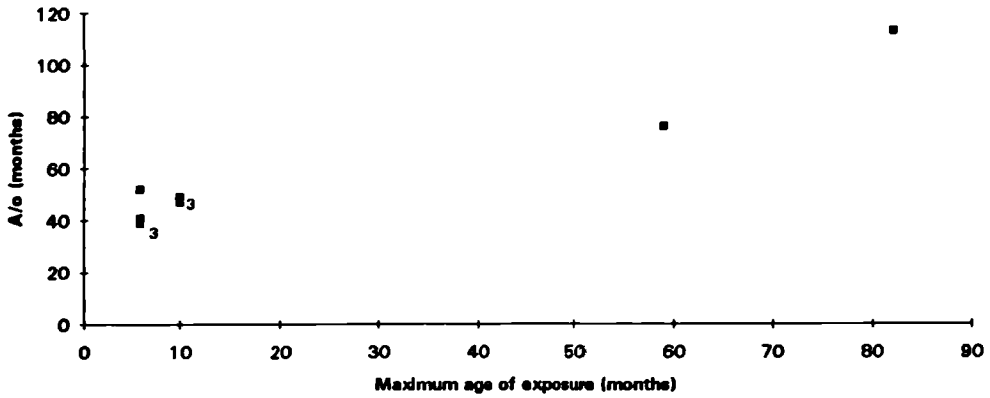
Contact status	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
Dam +ve	15	Normality	0.989	0.938	0.145	0.769
Flock/herd +ve	36	"	0.855	0.949	2.228	<0.001
Dam +ve	15	Log-normality	0.990	0.938	-0.395	0.427
Flock/herd +ve	36		0.939	0.949	1.007	0.012

For flock/herd-positive cases, the hypotheses of both normality and log-normality are rejected, with a right skew too elongated even for log-normality. Conversely, for dam-positive cases, neither hypotheses are rejected, and there is no marked skew.

7.1.1.3 Analysis of age at onset by age at last recorded exposure.

The relationship between a/o and maximum age of known exposure (in months) is plotted in figure 7.1.1.3.

Figure 7.1.1.3 Natural goat scrapie; GLS data; a/o by maximum age of exposure.



Regression analysis gives an additional 0.8 months at a/o per month of possible additional exposure (r = 0.97, F= 105.27, P=0.000), with an apparent a/o clustering (mean 45.1 months, range 39-52) for the seven animals removed from exposure by ten months of age.

7.1.2 Data from VIDA.

A right skew is apparent on the histogram of the 28 cases (not shown). The a/o data (in years), for the 28 cases:-

Mean:	**4.25**
SD:	**1.8**
Range:	**3-11**
Mode:	**3**
Median:	**4**

The a/o f/d analysis is given in table 7.1.2.

Table 7.1.2 Natural goat scrapie (VIDA); a/o f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
28	Normality	0.926	0.963	2.106	<0.001
28	Log-normality	0.985	0.963	1.188	0.008

The f/d is incompatible with the hypothesis of normality; there is a marked right skew. However, it is compatible with log-normality, although log-transformation of the data does not remove the asymmetry.

7.2 Experimental TSE in goats.

7.2.1 Full experimental dataset.

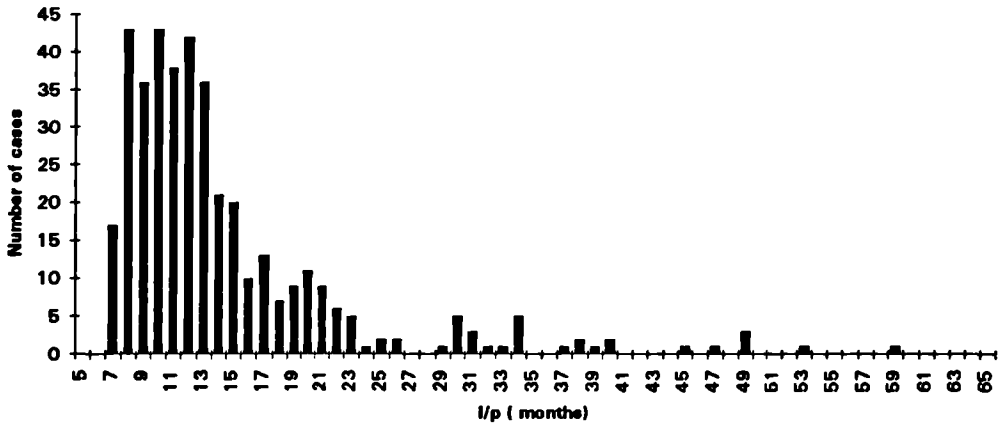
7.2.1.1 Analysis of incubation period.

The i/p data (in months) for the total of 400 cases:-

Mean i/p :	**14.4**
SD:	**8.0**
Range:	**7-59**
Mode:	**8-10**
Median:	**12**

There was no difference in mean i/p by sex (P=0.432, non-parametric test). The histogram of the i/p f/d (in months) is shown (figure 7.2.1.1).

Figure 7.2.1.1 Experimental TSE in goats; full dataset; i/p f/d.



After rapidly reaching the mode, the f/d has a very extended right skew. The i/p f/d analysis results are given in table 7.2.1.1.

Table 7.2.1.1 Experimental TSE in goats; full dataset; i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
400	Normality	0.866	0.995	2.476	<0.001
400	Log-normality	0.970	0.995	1.042	<0.001

Both the hypothesis of normality and log-normality can be rejected. Right skew is more extended than expected from log-normally distributed data.

7.2.1.2 Variables within the full experimental dataset; analysis of variance (ANOVA).

Stratification by variables recorded reduces each stratum to a very small number, as for sheep, therefore further analysis uses ANOVA. After preliminary analyses (results not shown), four factors (passage status, preceding donor, route of infection, and infecting substance) were used in the analysis. Twelve cases with an unknown infection route were omitted. The final dataset comprises 388 cases, grouped somewhat differently to sheep, as follows:-

Passage status: from natural case (21), from experimental case (367).
Preceding donor: goat (258), sheep and goat mixed (18), other ruminants (sheep or cattle) (29), unknown (but known to be either sheep or goats) (70), non-ruminants (13).
Infection route: CNS inoculation (361), s/c inoculation (27).
Infecting substance: CNS (337), any one of CSF, adrenal, sciatic nerve or spleen (39), any other tissue (12).

The i/p reciprocal was used, in order to normalise residuals (Kirkwood, 1988). Each factor was independently significant. Results are given in table 7.2.1.2.

Table 7.2.1.2 Experimental TSE in goats; full dataset; ANOVA results.

Variable being examined	Variable grouping	N	F	P	Mean i/p (adj)	95% CI
Passage status	From natural case	21	6.72	0.010	31.2	22.1-52.7
	From experimental case	367			20.6	17.9-24.3
Preceding donor	Goat	258	18.24	<0.001	16.1	14.1-18.9
	Sheep & goat mixed	18			28.4	20.2-47.4
	Sheep or cattle	29			42.5	28.8-81.7
	U/k (but either sheep or goat)	70			18.5	15.4-23.2
	Non-ruminants	13			37.2	22.8-100.2
Infection route	Inoculation into CNS	361	15.71	<0.001	19.1	16.5-22.9
	S/c (or i/d) inoculation	27			35.2	24.9-60.2
Infecting substance	CNS	337	12.90	<0.001	18.4	16.4-21.0
	CSF/adrenal/nerve/spleen	39			22.7	18.5-29.2
	Other tissues	12			44.4	25.7-165.7

There is no strong evidence of interaction between passage status and infecting substance (F=1.13, P=0.324) or passage status and infection route (F=1.77, P=0.185); analysis of other interactions failed due to insufficient data.

Each factor examined appears to have an independent effect on i/p; the assumption of constancy for each of the other factors gives differences in mean (adjusted) i/ps as shown. However, there is

wide variation in variance, reflected in differences for CI ranges. ANOVA assumes variances are similar so the validity of the test is reduced and the interpretation of results difficult.

7.2.2 Subgroup 1: Experimental CNS inoculation of CNS tissue.

7.2.2.1 Analysis of incubation period.

The i/p f/d data (in months) for 201 cases:-

Mean i/p :	**11.2**
SD:	**3.7**
Range:	**7-30**
Mode:	**8**
Median:	**10**

The i/p histogram (not shown) is similar to that for the full set, although the right tail is less extensive. The f/d analysis results are given in table 7.2.2.1.

Table 7.2.2.1 Experimental scrapie in goats; subgroup 1; i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
201	Normality	0.916	0.991	2.080	<0.001
201	Log-normality	0.982	0.991	0.875	<0.001

The hypotheses of normality and log-normality are both rejected. The right skew is greater than expected from a log-normal distribution.

7.2.2.2 Variables within subgroup 1; further analysis.

Passage number.

The i/p data (in months) by passage number is given in table 7.2.2.2.

Table 7.2.2.2 Experimental scrapie in goats; subgroup 1; i/p (in months) by passage number.

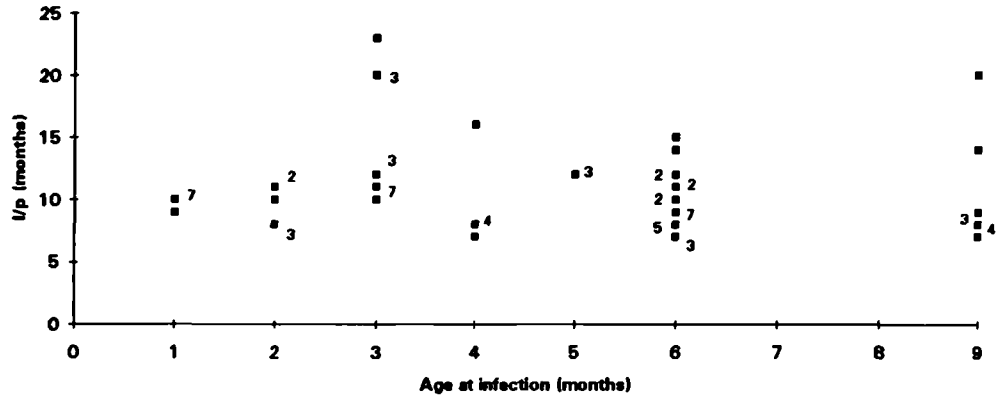
Passage number	Donor: species	N	Mean i/p	SD	Range (or values)	Mode	Median
2	Goat	31	12.7	3.6	7-21	14	13
3	"	7	9.9	1.9	7-12	9	9
4	"	10	10.3	2.9	7-16	8	10
5	"	8	9.5	1.3	8-11	8	10
6	"	4	9.3	1.5	8-11	8	10

There is a decrease in mean i/p with passage number (P= 0.021, non-parametric test). Regression analysis gives an average reduction of 1 month per passage (r=-0.41, F=11.55, P=0.001). Much of this decrease appears to be from P2 to P3; removing P2 results in a statistically insignificant reduction of 0.3 months per passage (r=-0.13, F=0.44, P=0.512).

Age at Infection.

Figure 7.2.2.2 (a) plots the relationship of age at infection with i/p (in months) for 71 cases

Figure 7.2.2.2(a) Experimental scrapie in goats; subgroup 1; i/p by age at infection

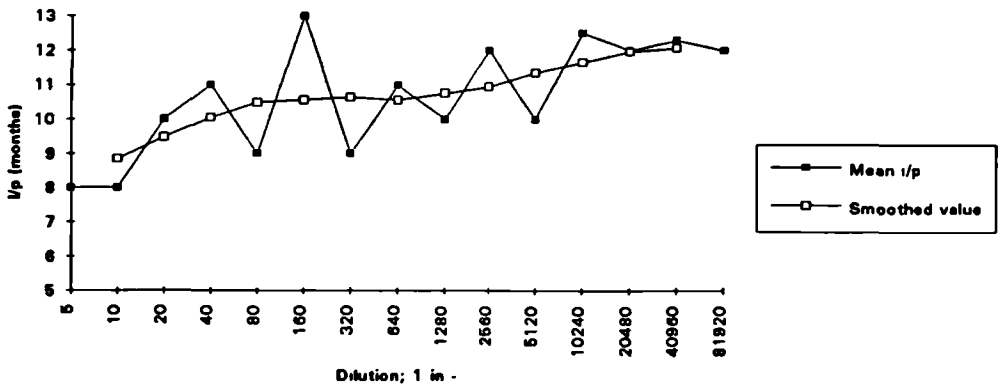


Regression analysis gives an insignificant decrease in mean i/p of 0.2 months per month as age at infection increases (r = -0.15, F=1.64, P=0 205).

Dose - quantified as dilution of brain extract.

Figure 7.2.2.2(b) plots the relationship of dose with i/p for 23 goats (smoothing was performed by taking a three stage moving average).

Figure 7.2.2.2(b) Experimental scrapie in goats: subgroup 1: i/p by dose



The mean i/p per dose ranges from 8 months to 13 months, observed i/p's from 8 to 15 months. There is an increase of mean i/p of 0.26 months per additional dilution (r=0.56, F=3.17 P= 0 090),

Other manipulations.

Formalin treatment (19 cases) increased the mean i/p to 12.5 months. Host pre-treatment with CNS inoculation of normal goat brain 2.5 months before agent inoculation (15 cases) decreased the mean i/p to 7.6 months (range 7-9; SD 0.63).

7.2.3 Subgroup 2: Cases with a plausible 'natural' infection route.

The i/p data (in months) for 34 cases:-

Mean i/p:	**27.5**
SD:	**9.8**
Range:	**11-49**
Modal region:	**21-34**
Median:	**28**

No marked skew is apparent from the histogram (not shown). There is an insignificant difference between mean i/p's of the two infection routes, with 27.0 (SD 10.5) and 29.4 (SD=10.5) months for s/c and oral routes respectively (P= 0.898, non-parametric test). I/p f/d analysis: table 7.2.3.

Table 7.2.3 Experimental TSE in goats; subgroup 2; i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
34	Normality	0.983	0.972	0.335	0.364
34	Log-normality	0.974	0.972	-0.613	0.108

Neither the hypothesis of normality, nor that of log-normality can be rejected. There is no marked skew in either raw or log-transformed data.

7.2.4 Subgroup 3: Cases infected directly from natural sheep scrapie.

The i/p data (in months) for the 14 cases:-

Mean i/p:	**36.4**
SD:	**10.9**
Range:	**23-59**
Mode:	**34**
Median:	**34**

The histogram (not shown) suggests the possibility of a right skew, but numbers are small. The f/d analysis results are given in table 7.2.4.

Table 7.2.4 Experimental scrapie in goats; subgroup 3; i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
14	Normality	0.948	0.938	0.864	0.102
14	Log-normality	0.974	0.938	0 499	0.330

Neither the hypothesis of normality, nor that of log-normality can be rejected. The right skew is not statistically significant.

7.3 Discussion

7.3.1 Age at onset and incubation period frequency distribution analysis findings.

7.3.1.1 Natural cases.

A summary of a/o f/d analyses results is given in table 7.3.1.1(i).

Table 7.3.1.1(i) Natural goat scrapie; summary of f/d analyses.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Logged data
GLS: all cases	66	No	No	R	R
GLS: dam scrapie +ve	15	Yes	Yes	NS	NS
GLS: flock/herd scrapie +ve	36	No	No	R	R
VIDA	28	No	Yes	R	R

[Notes on table:- Significance level of both f/d compatibility with normality and log-normality, and significance of skew is at the P=0.05 level.]

Table 7.3.1.1(ii) describes the f/d for each dataset. Although goats are known to be culled, no general data on culling patterns is available.

Table 7.3.1.1(ii) Natural goat scrapie; descriptions of f/d.

Dataset	Description
GLS: all cases	Right skew too elongated even for log-normality
GLS: dam scrapie +ve	Compatible with normal and log-normal
GLS: flock/herd scrapie +ve	Right skew too elongated even for log-normality
VIDA	Compatible with normality (but asymmetric)

Twenty-eight (42%) of the 66 natural transmission cases were under observation until at least 60 months of age (Hourrigan et al, 1979) during an investigation into indirect pasture transmission. The remainder were mainly kept singly or in small groups. No 'herd' information on culling is available.

The full dataset, and the 36 cases of flock/herd-positive contact both have an f/d too right skewed even for log-normality. The smaller dam-positive group a/o f/d is compatible with both normality and log-normality. Whether this group biologically represents a different group to the flock/herd contact group depends upon whether maternal transmission is possible by routes other than lateral contact. If maternal transmission occurs, it may result in a tighter a/o f/d due to a similar route and time of entry. However, lack of discrimination may simply reflect the small numbers involved.

The two tests for VIDA data give apparently conflicting results; i.e. compatible with log-normality, but with right skewed asymmetry in logged data. This may be a reflection of small sample size. Bias may be as for sheep VIDA data (Wooldridge, 1991) but, since many goats are kept singly, individual investigation may be more likely than for sheep thus reducing bias. Nevertheless, there may be some element of right censoring.

As for sheep, therefore, scrapie onset is possible well into a goat's old age; the a/o f/d may again have a very elongated right skew, more extreme than that of log-normality. Again as for sheep, there is a suggestion of an association of sample size and observed f/d, the smallest dataset being the least right skewed.

7.3.1.2 Experimental cases.

A summary of the i/p f/d analyses results is given in table 7.3.1.2(i)

Table 7.3.1.2(i) Experimental TSE in goats; summary of f/d analyses.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Logged data
Full set	400	No	No	R	R
Subgroup 1	201	No	No	R	R
Subgroup 2	34	Yes	Yes	NS	NS
Subgroup 3	14	Yes	Yes	NS	NS

[Notes on table:- Significance level of both f/d compatibility with normality and log-normality, and significance of skew is at the P=0.05 level.]

Table 7.3.1.1(ii) describes the f/d for each dataset.

Table 7.3.1.2(ii) Experimental TSE in goats; descriptions of f/d.

Dataset	Description
Full set	Right skew too elongated even for log-normality
Subgroup 1	"
Subgroup 2	Compatible with normal and log-normal
Subgroup 3	"

As with sheep, most work was published before the mid 1970's. Comments on possible contamination are therefore relevant also to goats.

For the complete set, 400 cases, the i/p f/d is right skewed more extensively than is compatible with log-normality. Maximum recorded i/p is 59 months, but some papers stated that animals were still under observation. It may be that some of these developed disease after even longer i/p's.

Subgroup 1, 201 cases of goat-to-goat CNS-to CNS transmission (the group chosen to be the most homogeneous possible) has a considerably reduced maximum i/p from that for the full set (30 months). Minimum value remains at 7 months. Consequently the mean is also reduced, by about 3 months. However, the f/d retains a right skewness too extensive for log-normality. It is hypothesised that if variability in all factors was genuinely reduced to negligible levels (in practice likely to be an impossible situation), the i/p range would comprise a very tight band.

For subgroups 2 (plausible natural infection routes) and 3 (transmission from sheep, suspected source of most goat scrapie), i/p f/d was compatible both with normality and log-normality. Both datasets are small, and again artificial truncation by cessation of observation is possible. Nevertheless, particularly for subgroup 2 (see figure 7.3.1.3), there does appear to be a more prominent left tail (comparable to that of the GLS natural dataset) than for subgroup 1. The modal region i/p for subgroup 2, 21-34 months, is considerably later than for subgroup 1, 8-12 months. This situation is similar to that for sheep, but the difference in time to reach the modal regions is greater. For sheep the difference was only 5 months.

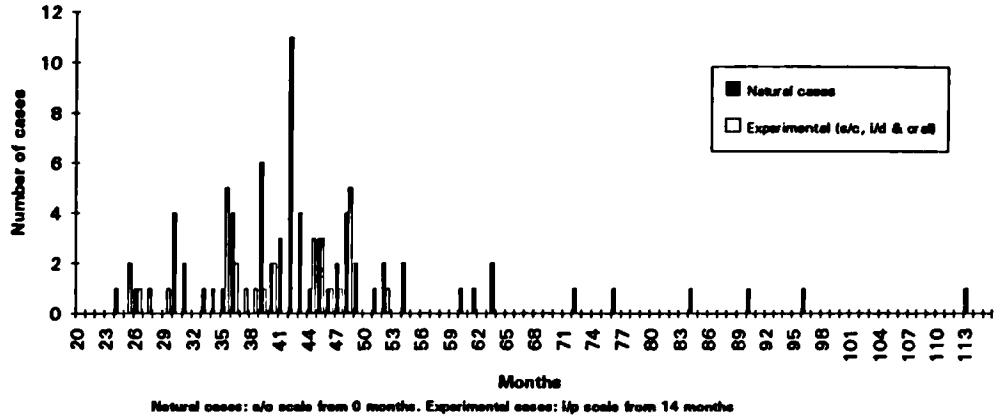
Conclusions are again therefore that i/p f/d may sometimes have a right skew more extreme than that of log-normality, and that smaller datasets tend to be less right skewed than larger sets. Homogeneity of conditions may affect i/p f/d but, with a given number of experimental cases, categorising to increase homogeneity automatically reduces case numbers which itself may also distort the i/p f/d. Complete separation of the two effects is therefore impossible in this study.

7.3.1.3 Comparison of natural age at onset and experimental incubation period frequency distribution.

As postulated for sheep, i/p f/d after CNS inoculation may mimic disease i/p f/d development from the point where natural peripheral infection has reached the CNS, as well as spontaneous *de novo* disease development if that does in fact occur, whilst s/c, i/d and oral inoculation i/p f/d's may mimic complete natural transmission i/p f/ds. It might therefore be expected that i/p f/d for the latter has some similarity to natural case a/o f/d if all natural cases are infected at approximately the same time.

Figure 7.3.1.3 compares f/d's of the GLS natural case dataset a/o's with experimental subgroup 2, the s/c, i/d and orally transmitted case i/p's. The modal regions are 'overlayed' at the same time point as it is distribution 'shape' which is being compared.

Figure 7.3.1.3 Goat TSE; comparison of natural a/o and experimental (s/c, i/d and oral) i/p f/d; modes overlayed



Natural cases: a/o scale from 0 months. Experimental cases: i/p scale from 14 months

If there is any validity in overlying the modal regions, and if the experimental transmission methods bear any resemblance to natural transmission, then the shift along the X-axis by 14 months of the i/p f/d might suggest that on average the natural cases were infected at around 14 months of age. Natural cases have a longer right tail. This may be connected with differences in sample size (66 versus 34), artificial truncation of experimental data, variation in infection age for natural cases or a genuinely different f/d. The left tails are broadly similar in extent.

7.3.2 Effect of variables.

7.3.2.1 Natural cases

Analysis of the a/o by the latest date of recorded exposure suggests the possibility of a/o clustering around 40-50 months for goats exposed from birth and infected within the first ten months of life, but with only seven cases no firm conclusions can be drawn.

7.3.2.2 Experimental cases.

The amount of detail given, and therefore variables which can be examined, differs greatly. The experimental method also varies; often only a small number of cases have been subjected to any one protocol. Many of the variables alter mean i/p in ways which might be expected. In all datasets lack of information means that some variables are unaccounted for. Inclusion of all unknown (and in many cases unknowable) variables might considerably alter the results.

Examination of the full dataset indicates that, as for sheep, passage status, preceding donor, infection route and infecting substance have independent statistically significant effects on the mean i/p. Variables not taken into account include breed, dose, actual passage number and

precise history ('strain') of agent. However, with respect to preceding donor, it is perhaps surprising that there is little difference in mean i/p for cases from other ruminants (cattle and sheep) and non-ruminants (mice and primates) although CI's are wide.

For subgroup 1, the analysis of the effect of passage number indicates a statistically significant i/p reduction only for P2, the first experimental passage. Age at infection has no statistically significant effect up to 9 months of age.

The reduction of agent dose by the dilution of brain extract results in a gradual i/p increase to 13 months. Pattison (1965b) reports 3 cases in goats where donor brain inoculum was diluted to 1/10 (10%), and 1ml inoculated, however, there the donor goats were healthy, presumably incubating disease, and i/p's are much longer at 23, 30 and 30 months. This extended i/p, if caused by a reduction comparable in dose to that plotted in figure 7.2.2.2(b), suggests that under certain circumstances (e.g. brain-to-brain, same species) incredibly small amounts of agent might be infectious, a possibility supported by the existence of electrode-infected human cases of CJD.

Pre-treatment by inoculation with supposedly uninfected goat brain reduces mean i/p after subsequent infectious inoculation, and was carried out 10 weeks before agent inoculation, i.e. approximately 10 months before disease onset. Pattison (1965a) suggests an immune mechanism 'sensitises' the recipients. Contamination of the 'normal' brain preparation is another possibility. If this was an accurate observation, unaffected by contamination, a parallel in natural cases could be hypothesised, whereby haematogenous or neuronal pathway entry to the CNS target area by 'priming' material of some sort, either organic or maybe even inorganic, results in increased susceptibility to disease.

7.3.3 Comparison with Sartwell's Conclusions.

For experimental TSE's in goats, i/p f/d with right skew too extreme for log-normality occurs in the two largest of four datasets. For a/o f/d of natural disease, this extreme skew occurs in the three largest of four datasets, although as with sheep, time of infection is unknown. Overall, however, an i/p f/d of log-normality cannot be assumed; experimental results indicate the possibility of a considerably greater right skew in natural disease.

Therefore, conclusions reached are broadly similar to those for sheep; Sartwell's hypothesis of underlying log-normality for i/p f/d is not confirmed.

CHAPTER 8

PRIMATES: DATA ANALYSIS AND DISCUSSION: Experimental TSE/prion disease.

8. Summary.

8.1 Full dataset of each primate group.

8.1.1 Analysis of incubation period.
8.1.2 Association between dataset size and frequency distribution skewness

8.2 Datasets with greater homogeneity.

8.2.1 Analysis of incubation period.
8.2.2 Association between dataset size and frequency distribution skewness.

8.3 Squirrel monkeys; analysis of incubation period by inoculation route.

8.4 Comparison of incubation period with some life history measures, by taxonomic grouping.

8.5 Discussion.

8.5.1 Incubation period frequency distribution analysis findings.
8.5.2 Association of incubation period with certain life history measures.
8.5.3 Comparison with Sartwell's conclusions.

CHAPTER 8

PRIMATES: DATA ANALYSIS AND DISCUSSION.

Experimental TSE/prion disease.

Summary.

For experimental TSEs in primates, the right skew of the incubation period (i/p) frequency distribution (f/d) appears to increase with increasing size of the dataset.

Datasets with more than about 50 cases are frequently found to have greater right skews than compatible with log-normality; any artificial right truncation (e.g. early cessation of experimental observation) would tend to mask this skew. Focusing on a particular 'homogenous' subset of cases does not necessarily reduce the skew. Some individual i/p's in such a dataset are extremely long compared with the modal i/p value.

It is concluded that the i/p f/d's of TSE's in primates are unlikely to be log-normally distributed, having a more extensive right skew; Sartwell's hypothesis is not confirmed.

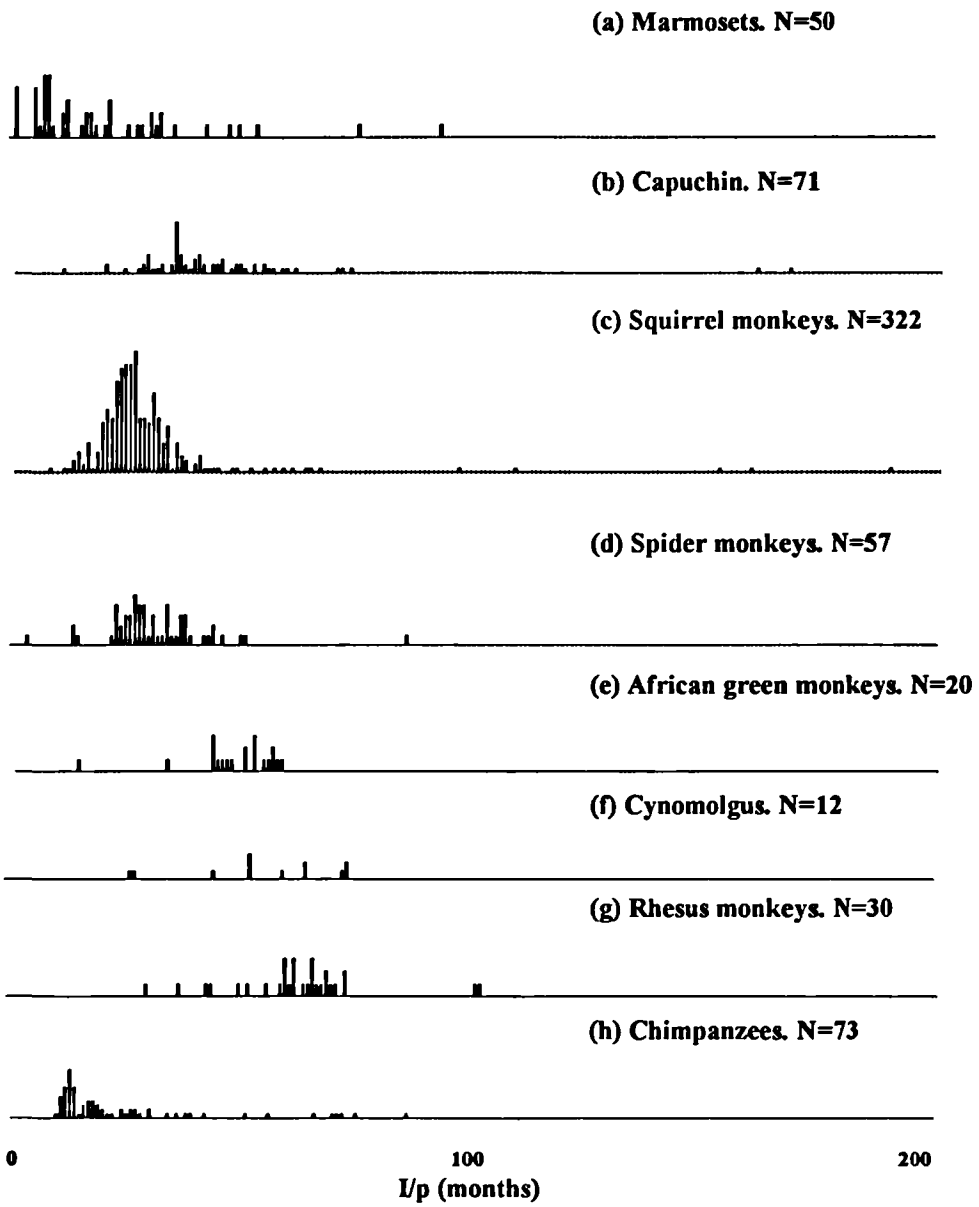
A non-significant decrease in f/d right skewness with increasing modal (and mean) i/p was noted; this observation may be confounded by species differences, sample size, and infection age. If all infection occurs at a similar proportion of species lifespan, then the longer the duration to modal i/p, the less subsequent mean lifespan remains in which to observe a right skew. Examination of f/d c/s for diseases with different modal i/p's, but the same age at infection, in a single species using 'large' samples, might clarify this observation.

8.1 Full dataset of each primate group

8.1.1 Analysis of incubation period

Figures 8.1.1 (a)-(h) illustrates the i/p f/d (in months) for each group (see appendix 10 for data source details, and appendix 21 for relevant primate taxonomy).

Figure 8.1.1(a-h) Experimental TSE in primates: i/p f/d by primate group.
[Note: Vertical axes give an indication of case numbers, but are minimised and not to scale].



The i/p data (in months), and a visual estimate of the f/d skew for each primate group are given in table 8.1.1(i).

Table 8.1.1(i) Experimental TSE in primates; i/p data (in months) plus visual estimate of f/d skew, by primate group.

Group	N	Mean i/p	SD	Range	Mode	Median	Skew
Marmoset	50	20.7	19	2-94	8	15	R
Capuchin	71	43.8	24	11-168	35	39	R
Squirrel M.	322	28.3	17	8-189	26	25	R
Spider M.	57	31.4	11	4-86	27	29	R
African Green	20	47.5	10	14-58	43	50	L
Cynomolgus	12	55.8	16	27-74	53	57	None
Rhesus	30	62.9	15	30-102	60	63	None
Chimp	73	23.8	18	10-86	13	17	R

The mean i/p's for different primates are clearly different ($P<0.001$, non-parametric test). The modal values also vary greatly, from 8 to 60 months.

Visually, the five largest datasets (N= 50 to 322) appear right skewed. Of the remaining three datasets (N= 12 to 30) one appears to be left skewed, and two have no clear skewness, suggesting an association between sample size and skewness. The i/p f/d analyses are given in table 8 1.1(ii).

Table 8.1.1(ii). Experimental TSE in primates; i/p f/d analysis by primate group.

Group	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
Marmoset	50	Normality	0.902*	0.977	1.919	<0.001	R
Capuchin	71	"	0.770*	0.982	3.762	<0.001	R
Squirrel M.	322	"	0.680*	0.995	6.110	<0.001	R
Spider M.	57	"	0.916*	0.979	1.774	<0.001	R
African green	20	"	0.889*	0.950	-1.936	0.001	L
Cynomolgus	12	"	0.970	0.926	-0.630	0.244	NS
Rhesus	30	"	0.942*	0.964	0.506	0.199	NS
Chimp	73	"	0.841*	0.983	1.991	<0.001	R
Marmoset	50	Log-normality	0.988	0.977	-0.314	0.320	NS
Capuchin	71	"	0.938**	0.982	0.771	0.009	R
Squirrel M.	322	"	0.926**	0.995	1.706	<0.001	R
Spider M.	57	"	0.897**	0.979	-1.901	<0.001	L
African green	20	"	0.776**	0.950	-2.970	<0.001	L
Cynomolgus	12	"	0.933	0.926	-1.058	0.058	NS
Rhesus	30	"	0.941**	0.964	-0.706	0.081	NS
Chimp	73	"	0.937**	0.983	1.116	<0.001	R

[Notes on table:- * indicates hypothesis of normality can be rejected at P=0.05 level
** indicates hypothesis of log-normality can be rejected at P=0 05 level]

The hypothesis of normality can be rejected for all except the cynomolgus group (the smallest). The skew in every group confirms the visual impression.

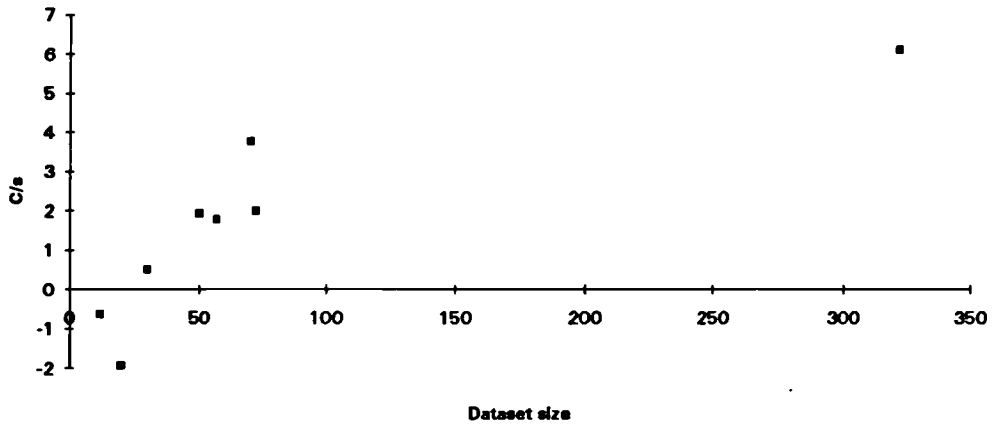
The hypothesis of log-normality can be rejected for six of the eight groups. After log-transformation the three largest sets, capuchin, chimp and squirrel monkey, are still right skewed (N=71, 73, 322), whereas for spider and African green monkeys (N=57, 20), the skew is left. The Rhesus f/d (N=30) is not compatible with log-normality, probably due to kurtosis, since there is no skew after transformation.

Two datasets are compatible with log-normality, cynomolgus and marmoset (N=12, 50), having no significant skew in the transformed data. The cynomolgus is a particularly small dataset, which may be the reason for inability to reject either description.

8.1.2 Association between dataset size and frequency distribution skewness.

Regression analysis gives an increase in c/s of 0.021 units per additional case in the dataset (r=0.84, F=13.93, P=0.010). Figure 8.1.2 plots this association.

Figure 8.1.2 Experimental TSE in primates; c/s by dataset size.



8.2 Datasets with greater homogeneity; subgroups 1 to 6.

8.2.1 Analysis of incubation period.

Analyses were performed for all subgroups (see appendix 10 for subgroup definitions) as in section 8.1.1. Results for subgroup 6, the most homogeneous, are given. Table 8.2.1(i) gives sample sizes, mean i/p and a visual estimate of skew taken from the histograms (not shown).

Table 8.2.1(i) Experimental CJD in primates; subgroup 6; sample size, mean i/p (months), and visual estimate of i/p f/d skew by primate group.

Group	N	Mean i/p	Skew: by visual estimation
Marmoset	8	22.1	R
Capuchin	28	44.7	R
Squirrel M.	129	27.6	R
Spider M.	20	31.9	None
African green	13	49.2	None
Cynomolgus	4	56.0	None
Rhesus	17	63.3	L
Chimp	24	27.1	R

The mean i/p is virtually unchanged for all groups, compared with the full set. All subgroups with over 20 cases still appear right skewed. Smaller groups generally have no clear shape. Table 8.2.1(ii) gives the i/p f/d analyses.

Table 8.2.1(ii) Experimental CJD in primates; subgroup 6; i/p f/d analysis by primate group.

Group	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
Marmoset	8	Normality	0.821*	.905	1.974	0.002	R
Capuchin	28	"	0.641*	.962	4.379	<0.001	R
Squirrel m.	129	"	0.555*	.988	6.909	<0.001	R
Spider m.	20	"	0.987	.950	0.406	0.369	NS
African green	13	"	0.988	.931	0.249	0.631	NS
Cynomolgus	4	"	1.000	.868	1.155	nc	nc
Rhesus	17	"	0.963	.942	-0.990	0.051	NS
Chimp	24	"	0.885*	.957	1.328	0.006	R
Marmoset	8	Log-normality	0.916	.905	1.402	0.022	R
Capuchin	28	"	0.818**	.962	2.964	<0.001	R
Squirrel m.	129	"	0.844**	.988	3.046	<0.001	R
Spider m.	20	"	0.975	.950	-0.403	0.373	NS
African green	13	"	0.989	.931	0.138	0.789	NS
Cynomolgus	4	"	1.000	.868	1.155	nc	nc
Rhesus	17	"	0.941**	.942	-1.285	0.015	L
Chimp	24	"	0.945**	.957	0.719	0.101	NS

[Notes on table:- * indicates hypothesis of normality can be rejected at P=0.05 level
** indicates hypothesis of log-normality can be rejected at P=0.05 level
nc indicates no calculation was completed; programme would not run on this small data set.]

The two largest subgroups have right skews too extensive for compatibility with log-normality. The smaller datasets tend to be less right skewed. The calculated skew in every subgroup except Rhesus monkeys confirms the visual estimate No overall f/d pattern emerges. Subgroups 1 to 5 give broadly similar results (not shown); again no overall f/d pattern emerges.

8.2.2 Association between dataset size and frequency distribution skewness.

Regression analysis gives a c/s increase of 0.051 units per additional case in the dataset (r=0.82, F=12.12, P=0.013).

8.3 Squirrel monkeys; analysis of incubation period by inoculation route.

Squirrel monkeys are the largest group, therefore further analyses are possible. The i/p data (in months) for subgroup 6 is given by inoculation route: table 8.3(i).

Table 8.3(i) Experimental CJD in squirrel monkeys; subgroup 6; i/p data (in months) by inoculation route.

Inoculation route	N	Mean i/p	SD	Range	Mode	Median
CNS only	53	29.9	23.4	16-189	26	26
Multiple site	76	26.0	16.1	11-159	24	25

The difference in mean i/p by inoculation route is almost 4 months, (P=0.037, non-parametric test), multiple site inoculation giving disease onset more rapidly. However, there appears to be little difference in the very right skewed i/p f/d histograms (not shown). The ranges show that both inoculation protocols occasionally give a very long i/p. The f/d analyses are given in table 8 3(ii).

Table 8.3 (ii) Experimental CJD in squirrel monkeys; subgroup 6; i/p f/d by inoculation route.

Inoculation route	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Ske w
CNS only	53	Normality	0.559	0.978	6.137	<0.001	R
Multiple site	76	"	0.540	0.983	7.476	<0.001	R
CNS only	53	Log-normality	0.831	0.978	3.299	<0.001	R
Multiple site	76		0.836	0.983	2.784	<0.001	R

The two inoculation routes have similar i/p f/d's, both more right skewed than compatible with log-normality.

8.4 Comparison of incubation period with some life history measures, by taxonomic grouping.

Table 8.4 compares maximum recorded i/p (as a percentage of maximum recorded longevity) and mid-point of the adult weight range for the genus (see appendix 21 for relevant taxonomy, and appendix 22 for selected life history measures) with both the mean and the modal i/p and the c/s for the full dataset groups (i.e. those described in section 8.1).

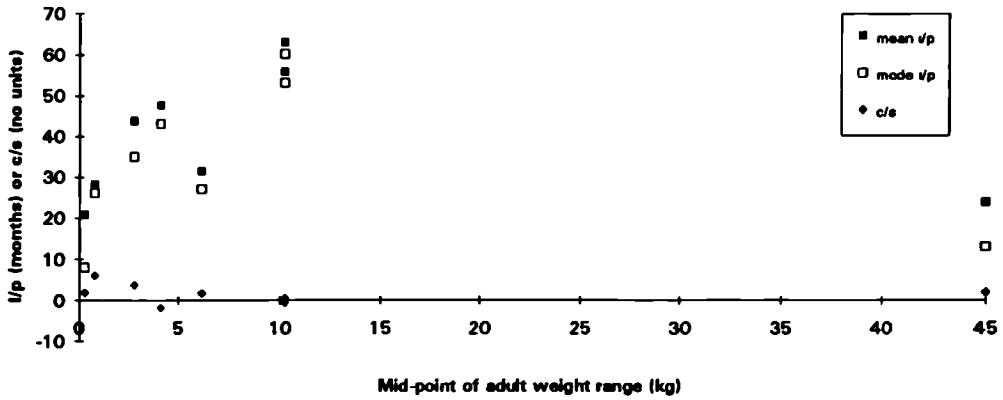
Table 8.4 Experimental TSE in primates; full dataset; comparison of analysis results with life history measures, stratified by data representative of genus.

Primate group	Max i/p as % of max longevity	Adult weight: mid-range (kg)	Mean i/p	Modal i/p	C/s (full dataset)
Marmosets	63	0.25	20.7	8	1.9
Capuchin	35	2.75	43.8	35	3.8
Squirrel monkey	75	0.75	28.3	26	6.1
Spider monkey	36	6.15	31.4	27	1.8
African green	15	4.10	47.5	43	-1.9
Cynomolgus	21	10.25	55.8	53	-0.6
Rhesus monkey	29	"	62.9	60	0.5
Chimpanzee	16	45	23.8	13	2.0

Squirrel monkeys, with a large dataset, and thus the potential for a more complete i/p range, at 75%, had a maximum i/p which most closely approached the maximum recorded lifespan. Regression analysis gives an increase in c/s of 0.9 units per 10% increase in maximum i/p as a percentage of maximum recorded lifespan (r=0.75, F=7.73, P=0.032).

Figure 8.4 plots the i/p relationship with the mid-point of the adult weight range, for each primate group.

Figure 8.4 Experimental TSE in primates; comparison of certain i/p measures with mid-point of adult weight range, by primate group.



With the exception of chimps, which are far heavier than the other groups, there appear to be associations between weight-range mid-point, and i/p mean, mode and c/s. Excluding chimps, regression analysis gives an increase of 3.7 months in mean i/p (r=0.86, F=13.66, P=0.014), and of 3.1 months in modal i/p (r=0.84, F=13.66, P=0.014), and an insignificant decrease in c/s of 0.038 units (r=0.59, F=2.66, P=0.164) per kilogram increase in weight range mid-point.

These results suggest an association between the mean (and modal) i/p and the c/s. Regression analysis gives similar, insignificant results with or without chimps; including chimps, c/s

decreases by 0.09 per additional month for mean i/p (r=0.53, F=2.35, P=0.176), and by 0.07 per additional month for modal i/p (r=0.48, F=1.77, P=0.232).

8.5 Discussion.

8.5.1. Incubation period frequency distribution analysis findings.

Table 8.5.1(i) summarises the f/d analyses for the full datasets.

Table 8.5.1(i) Experimental TSE in primates; full dataset, descriptions of f/d.

Group	N	Description
Marmoset	50	Compatible with log-normality
Capuchin	71	Right skew too elongated even for log-normality
Squirrel M.	322	"
Spider M.	57	Right skew between normal & log-normal
African green	20	Left skew too elongated even for normality
Cynomolgus	12	Compatible with normality and log-normality; small data set
Rhesus	30	Not compatible with log-normality (kurtosis?)
Chimp	73	Right skew too elongated even for log-normality

Two possible associations are present; one between skew and dataset size, the other between skew and whether categorisation is by species (members within a species being closely related genetically) or genus (greater genetic heterogeneity, if different species involved). However, the five largest datasets are also the five groups categorised by genera, making investigation difficult. In addition, three different superfamilies are involved (see appendix 21), and these segregate with the genus/species categorisation, further complicating interpretation.

Table 8.5.1(ii) summarises the f/d analysis for subgroup 6.

Table 8.5.1(ii) Experimental CJD in primates; subgroup 6; descriptions of f/d.

Group	N	Description
Marmoset	8	Right skew too elongated even for log-normality
Capuchin	28	"
Squirrel m.	129	"
Spider m.	20	Compatible with both normality & log-normality
African green	13	"
Cynomolgus	4	"
Rhesus	17	Compatible with normality
Chimp	24	Not compatible with log-normality (kurtosis?)

More homogenous experimental conditions did not result in a similar i/p f/d across primate categories, but subgroup sizes are very dissimilar; some are extremely small. More subgroup i/p f/d's are compatible with the hypothesis of normality and log-normality than for full groups. Since

numbers have also been reduced, it will, due to lack of power, be more difficult to reject these hypotheses, therefore this may be the effect of reduction of sample size. Table 8.5.1(iii) compares the f/d skew for the full datasets of primates with that for the 'homogenous' subgroup 6

Table 8.5.1(iii) Experimental TSE in primates, homogeneity, size and f/s skew comparisons.

Group	Full set		Subgroup 6	
	N	Skew	N	Skew
Marmoset	50	R	8	R
Capuchin	71	R	28	R
Squirrel m.	322	R	129	R
Spider m.	57	R	20	None
African green	20	L	13	None
Cynomolgus	12	None	4	None
Rhesus	30	None	17	L
Chimp	73	R	24	R

All datasets with over 30 cases are right skewed. There is a tendency for a loss of right skew as the datasets reduce in size. It is therefore surprising that the marmoset (N=8) f/d is still right skewed, but this is heavily influenced by one case (histogram not shown). If a decrease of right skewness results from reduction in case numbers, and thus on the probability of observing particular values rather than on a biological phenomenon, one would occasionally expect such results.

Subdivision of squirrel monkeys into two inoculation route groups confirms the occasional very long i/p for both inoculation routes. Although both routes included a CNS component, multiple site inoculation shortened the mean i/p by about 4 months. Possible explanations include confounding, for example by dose, and the possibility that the non-CNS component produces disease more rapidly than CNS inoculation.

Analysis of association between dataset size and c/s gave statistical significance for both the full datasets and for subgroup 6, with an increase for c/s of 0 021 and 0.051 units per additional case in the respective datasets.

Therefore, and as with sheep and goats, 'large' datasets are often more right skewed than log-normality, and there is an association between right skewness with dataset size. It is not possible to confirm an underlying i/p f/d pattern for primates, neither has its existence been disproved. Difficulties lie in the disparate and often small sizes of datasets.

8.5.2 Association of incubation period with certain life history measures.

Some of the primate groups represent genera, and some species (see appendix 21); group genetic diversity will therefore vary. The African green monkey is a single species with no sub-species, therefore it may be the most genetically homogenous group examined. However other groups may, although described by their generic name, comprise only one experimental species or subspecies.

In examination of the i/p f/d, the relationship between age at inoculation and longevity is clearly relevant. Death from other causes during incubation will effect the observed f/d. Little information is available on the mean lifespan for these primates, therefore maximum recorded lifespan was used.

The right skewness of the i/p f/d increases with an increase in maximum recorded i/p as a percentage of maximum recorded lifespan. Although the mean lifespan may be much less, if all animals were young when inoculated this suggests there was no lack of natural lifespan remaining in which to exhibit an i/p f/d right skewness.

However, age at inoculation was unknown. If inoculation in a large proportion of the animals was towards the end of their natural lifespan, truncation of the maximum observed i/p values would be predicted. Although using elderly-looking animals to investigate disease with known likelihood of long i/p is unlikely, using even 'middle-aged' animals might artificially truncate the results, resulting in a reduced right f/d skew. The possibility of truncation cannot be eliminated.

Variation in most life history measures, including longevity, is highly correlated with variation in body size (Harvey and Clutton-Brock; 1985). For captive primates, since the age of a captured animal is generally unknown, adult weight is more reliably estimated. The weight (and probably therefore longevity) ranges for some genera are wide, and for some genera sex also makes a large difference. Genus pan (chimps) are much heavier than any other genera used. The mid-point of the given genus weight range was arbitrarily taken for analyses. It was therefore surprising to find, excluding chimps, a graphically visible and statistically significant increase of over 3 months in the mean (and modal) i/p per additional kilogram weight.

Although not statistically significant, there was a tendency for a lower body weight to result in a greater f/d right skew but this also correlated with sample size. As small primates are cheaper to keep (less space, less food etc.), they may be used more extensively, resulting in larger samples.

Sacher and Staffeldt (1974, cited by Harvey & Clutton-Brock, 1987) argue that species differences are better predicted by absolute brain size than body size. They describe a 'postnatal brain growth index' and primate subfamilies with relatively large brains at birth have relatively less postnatal brain growth. This is an exception to the usual pattern whereby growth rates are positively correlated with birth weight across species. It is conceivable that brain growth rates may be relevant to TSE i/p's in primates.

Examination of life-measures led to the observation that, although not statistically significant, the f/d right skewness decreased with increasing modal (and mean) i/p. This may be related to longevity; if all infection occurs at a similar proportion of species lifespan, then the longer the duration to modal i/p, the less subsequent mean lifespan remains in which to observe a right skew. However, interpretation is difficult; different species/genera are involved, which affects mean (and modal) i/p; age at infection may vary, and (as for all relationships) there may be confounding by sample size.

Examination of the f/d c/s for diseases with different mean i/p's, but the same age at infection (or initiating-event, for non-infectious disease), in a single species using 'large' samples, might be informative.

8.5.3 Comparison with Sartwell's conclusions.

As for sheep and goats, the i/p f/d for experimental TSE datasets in primates often gives a more extended right-skewed distribution than is compatible with log-normality, particularly if the sample size is over 30. Sartwell's hypothesis is not confirmed.

CHAPTER 9

RODENTS: DATA ANALYSIS AND DISCUSSION: Experimental TSE/prion disease.

9. Summary

9.1 Experimental BSE in mice.

9.1.1 Full dataset.
- 9.1.1.1 Analysis of incubation period for full dataset.
- 9.1.1.2 Analysis of incubation period by age at inoculation.

9 1.2 Analysis by originating BSE case ('BSE group').
- 9.1.2.1 Analysis of incubation period.
- 9.9.1.2 Analysis of incubation period by age at inoculation.

9.2 Experimental scrapie in mice.

9 2.1 Analysis of full dataset.
9 2.2 Analysis by agent strain.
9.2.3 Analysis of strain 22A by 'sinc' genotype.
9 2.4 Analysis of strain ME7 by 'sinc' genotype.
9.2.5 Interaction of 'sinc' genotype and agent strain.
9.2.6 Association of mean and modal incubation period with frequency distribution coefficient of skewness.

9.3 Experimental CJD in guinea-pigs.

9 3.1 Analysis of full dataset.
9 3.2 Effect of passage number on incubation period.
9.3.3 Analysis of incubation period frequency distribution for passage numbers with 10 or more cases.

9.4 Discussion.

9.4.1 BSE in mice.
- 9.4.1.1 Incubation period frequency distribution analysis findings.
- 9.4.1.2 Association between inoculation age and incubation period.

9.4.2 Scrapie in mice.

9.4.2.1 Incubation period frequency distribution analysis findings.

9.4.2.2 Association between incubation period, agent strain, and mouse strain.

9.4.2.3 Association between modal and mean incubation period and coefficient of skewness.

9.4.3 CJD in guinea-pigs

9.4.3.1 Incubation period frequency distribution analysis findings.

9.4.3.2 Effect of passage on incubation period.

9.4.4 Comparison with Sartwell's conclusions.

CHAPTER 9

RODENTS: DATA ANALYSIS AND DISCUSSION.

Experimental TSE/prion disease.

Summary

For both experimental BSE in mice, and CJD in guinea-pigs, with one strain of host and a reduced number of other variables, an incubation period (i/p) frequency distribution (f/d) with right skew more extensive than that compatible with log-normality was found; Sartwell's hypothesis is therefore not confirmed. For BSE, skewness is associated with mean inoculation age, right skewness decreasing as mean inoculation age increases.

With experimental scrapie in mice, two highly stable scrapie agent strains were used, inoculated into two strains of host mouse, plus their crosses (giving three sets of mice with known 'sinc' gene alleles; s7s7, s7p7 and p7p7); so that both host and agent were tightly defined, giving six possible combinations. Overall i/p f/d appearance was multi-modal; within this, the i/p f/d of each combination was tightly clustered, giving rise to the possibility of prediction of the host genotype, if agent strain and i/p are known. Clusters were compatible with normality and log-normality, but sample sizes were all fewer than 50. C/s decreases exponentially with increasing modal (and mean) i/p.

Age at inoculation and passage number have significant effects on i/p.

9.1 Experimental BSE in mice.

9.1.1 Full dataset.

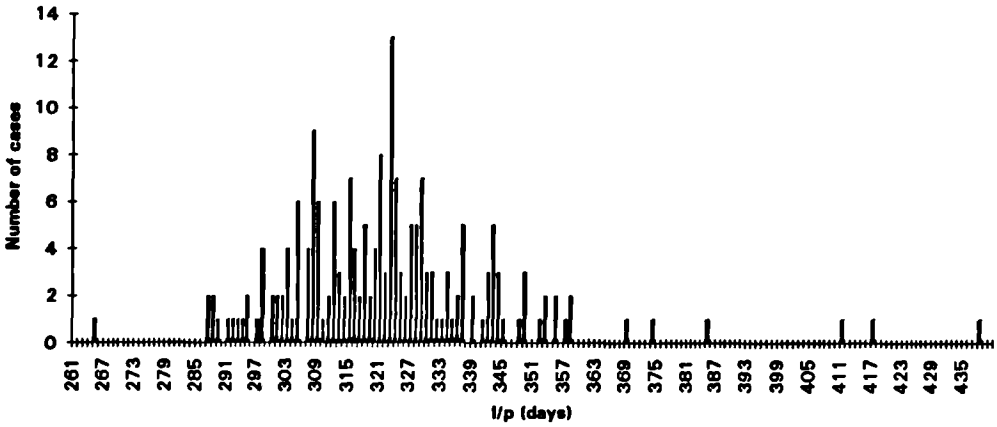
9.1.1.1 Analysis of incubation period for full dataset.

RIII mice were used (see appendix 10 for details). I/p details (in days) for 196 cases :-

Mean i/p:	323
SD:	21
Range:	265-438
Mode:	323
Median:	321

There is no difference in mean i/p by sex (females, 325 days; males, 320; P= 0.303, non-parametric test). The i/p f/d histogram is shown in figure 9.1.1.1.

Figure 9.1.1.1 BSE in mice; full dataset; i/p f/d.



Only one 'early' case is present, in comparison with several 'late' cases distant from the mode which contribute to a right skewed appearance. I/p f/d analysis results, table 9.1.1.1.

Table 9.1.1.1 BSE in mice; full dataset; i/p f/d analysis.

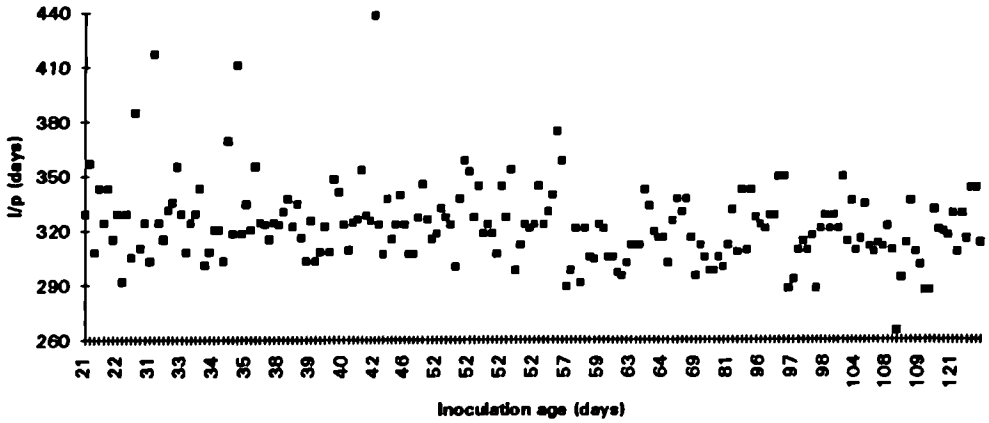
N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
196	Normality	0.941	0.992	1.660	<0.001
196	Log-normality	0.960	0.992	1.165	<0.001

Data are incompatible with both normality and log-normality. The right skew is too extensive.

9.1.1.2 Analysis of incubation period by age at inoculation.

The mean inoculation age of the mice is 62.7 days (SD 29.2, range 21-124). Regression analysis gives an i/p reduction of 0.16 days per additional day of age at inoculation (r=-0.22, F=9.79, P=0.002), plotted in figure 9 1.1.2.

Figure 9.1.1.2 BSE in mice; full dataset; i/p by inoculation age



9.1.2 Analysis by originating BSE case ('BSE group').

9.1.2.1 Analysis of incubation period.

Table 9.1.2.1(i) gives details for each 'BSE group' (see appendix 10, data used, for details), one group having been inoculated from each of the 7 cattle donors.

Table 9.1.2.1(i) BSE in mice: inoculation age and i/p data (in days) by BSE group.

BSE group	N	Mean infection age	Age range	Mean i/p	SD	Range	Mode	Median
1	35	51.7	44-65	328	17	300-374	323	325
2	32	29.2	21-41	327	24	292-417	324	324
3	30	110.1	94-124	316	18	265-343	308	313
4	31	63.2	41-113	314	15	289-353	312	312
5	23	97.0	96-98	321	17	288-349	321	321
6	22	54.3	38-69	319	14	295-344	323	322
7	23	37.0	32-42	335	34	303-438	303	322

As apparent from the table, there is a difference in mean i/p by BSE group (P = 0.009, non-parametric test). From inspection of individual histograms (not shown), f/d differences are apparent. Groups 2 and 7 have marked right skews, whilst the remainder appear more symmetrical. Although group sizes vary all are small, but groups 2 and 7 are at opposite ends of the range, with sizes of 32 and 23 respectively. I/p f/d analyses are given in table 9.1.2.1 (ii).

Table 9.1.2.1(ii). BSE in mice; i/p f/d analysis by BSE group.

BSE group	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
1	35	Normality	0.983	0 968	0.607	0.107	NS
2	32	"	0.898*	0 996	1.951	<0.001	R
3	30	"	0.975	0.964	-0.588	0.140	NS
4	31	"	0.984	0.958	0.655	0.099	NS
5	23	"	0.985	0.955	-0.155	0.716	NS
6	22	"	0.989	0.954	-0.219	0.612	NS
7	23	"	0.898*	0 955	1.763	0.001	R
1	35	Log-normality	0.988	0.968	0.484	0.192	NS
2	32	"	0.922*	0.996	1.638	<0.001	R
3	30	"	0.968	0.964	-0.781	0.057	NS
4	31	"	0.988	0.958	0.539	0.168	NS
5	23	"	0.983	0.955	-0.291	0.495	NS
6	22	"	0.987	0.954	-0.295	0.496	NS
7	23	"	0.920*	0.955	1.545	0.002	R

BSE groups 2 and 7 (*) are too right skewed even for compatibility with log-normality. All other groups are compatible with both normality and log-normality. Regression analysis gives no evidence for a relationship between group size and c/s (r=0.02, F=0.18, P=0.687).

9.1.2.2 Analysis of incubation period by age at inoculation.

Particularly with skewed f/d's, mean i/p is sensitive to extremes, and so dependant on range of observed data, whilst modal i/p is not. Regression of modal i/p on BSE group mean inoculation age gives a slight, non-significant reduction of 0.05 days per additional day of age at inoculation (r=0.18, F=0.17, P=0.695). For mean i/p, the equivalent estimate gives a decrease of 0.16 days per additional day of age at inoculation (r=0.65, F=3.58, P=0.117).

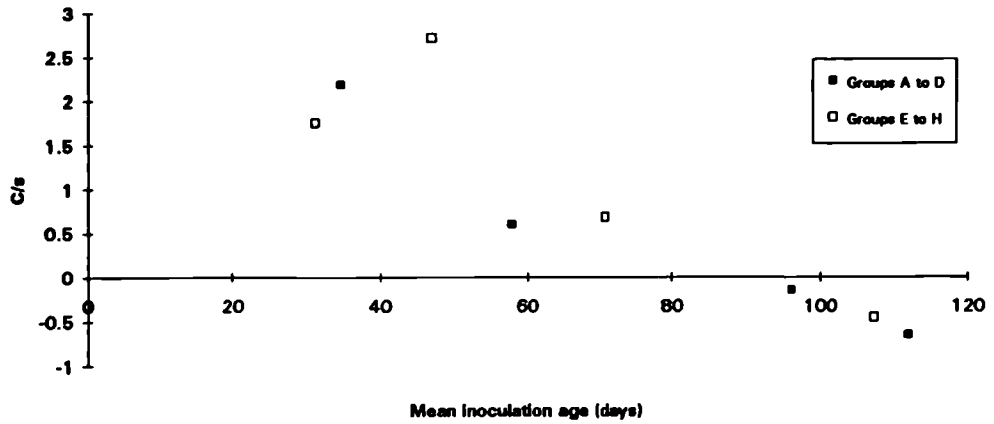
Regression analysis gives a c/s reduction in of 0.03 units for each day's increase in BSE group mean inoculation age (r=0.85, F= 12.83, P = 0.016). However, two way ANOVA gives significance to the effect of BSE group (F=2.03, P=0.063) on i/p, but not to inoculation age once BSE group is taken into account (F=0.48, P=0.490). Data was rearranged into groups A to H (see appendix 10 for details) to partially separate effects of inoculation age from BSE group; table 9.1.2.2 gives inoculation age details, plus c/s data.

Table 9.1.2.2 BSE in mice; inoculation age data (in days) and f/d c/s analysis; groups A to H.

Group	N	Mean inoculat-ion age	Inoculat-ion age range	Range span (days)	C/s	P for c/s	Skew
A	71	34.6	21-46	26	2.186	<0.001	R
B	69	57.8	47-72	26	0.597	0.038	R
C	27	96.1	73-98	26	-0.144	0.718	NS
D	29	111.9	99-124	26	-0.652	0.109	NS
E	47	31.1	21-38	18	1.752	<0.001	R
F	54	46.9	39-52	14	2.713	<0.001	R
G	52	70.7	54-96	43	0.687	0.038	R
H	43	107.3	97-124	28	-0.457	0.180	NS

For both regrouping methods, there is a tendency for c/s to decrease as group mean inoculation age increases (see figure 9.1.2.2). This is similar to that in analysis by BSE group.

Figure 9.1.2.2 BSE in mice; c/s age at infection for groups A to H



Regression analyses for groups A to D gives a decrease in c/s of 0.04 units per 'group mean day older' at inoculation (r=-0.96, F=24.17, P=0.039), and for groups E to H similarly a decrease of 0 04 units per day (r=-0.88, F=6.71, P=0.122).

9.2 Experimental scrapie in mice.

9.2.1 Analysis of full dataset.

Appendix 10 gives details of scrapie strains and mouse genotypes used. I/p data (in days) for all 407 cases:-

Mean i/p:	**400**
SD:	**157**
Range:	**165-645**
Mode:	**215**
Median:	**455**

By inspection of the histogram (not shown; similar to figure 9.2.2 shown later) neither normality nor log-normality is applicable; there is no obvious single peak. Appearance is of multi-modality.

9.2.2 Analysis by agent strain.

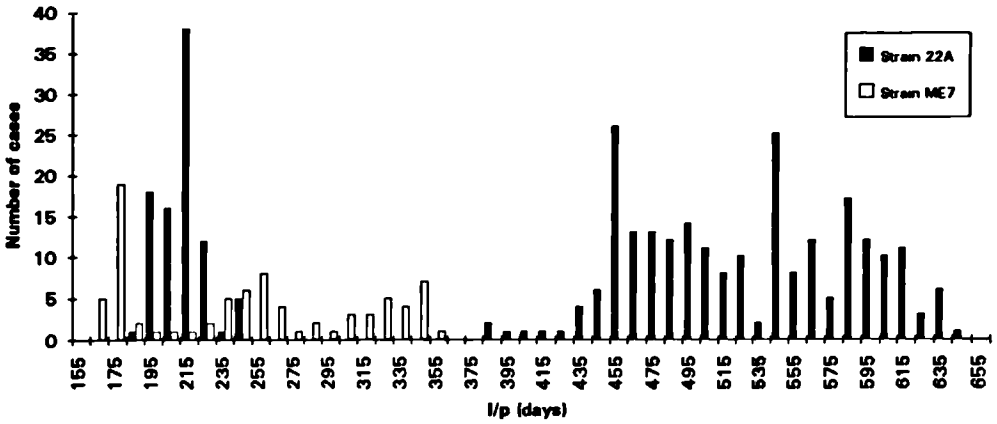
Table 9.2.2 gives i/p data by agent strain.

Table 9.2.2 Scrapie in mice: i/p data (in days) by scrapie agent strain.

Strain	N	Mean i/p	SD	Range	Mode	Median
22A	326	438	150	185-645	215	485
ME7	81	248	63	165-355	175	245

Mean i/ps are very different, that of strain 22A being almost 200 days more than that of strain ME7 (P<0.001, non-parametric test). The i/p f/d histogram is given below (figure 9.2.2).

Figure 9.2.2 Scrapie in mice; full dataset; i/p f/d by agent strain



Although strain ME7 has fewer cases, and is compacted into a smaller range, both strains appear tri-modal. Neither normality nor log-normality is applicable.

9.2.3 Analysis of Strain 22A by 'sinc' genotype.

Table 9.2.3(i)gives i/p details by 'sinc' genotype.

Table 9.2.3(i) Scrapie in mice; strain 22A: i/p data (in days) by 'sinc' genotype.

'Sinc' type recorded	N	Mean i/p	SD	Range	Mode	Median
s7s7	27	462	17	415-495	455	465
s7p7	47	568	37	485-635	585	575
p7p7	29	210	14	185-245	205	205
Unrecorded	223	438	149	195-645	215	485

Mean i/ps for the three 'sinc' genotypes are very different, ranging from 210 days to 568 days (P<0.001, non-parametric test); the mean for unrecorded genotypes is within this range. Figure 9 2.3 (a)-(b) illustrates the data.

Figure 9.2.3(a) Scrapie in mice; strain 22A; i/p f/d by 'sinc' genotype

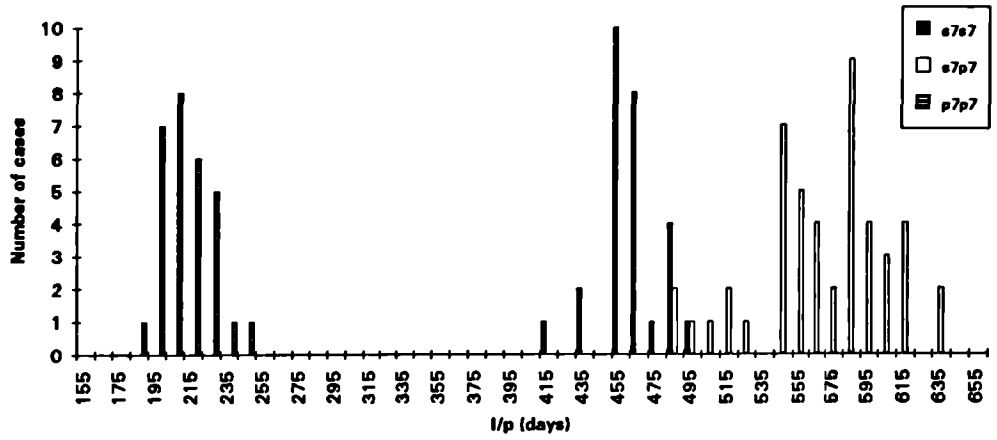
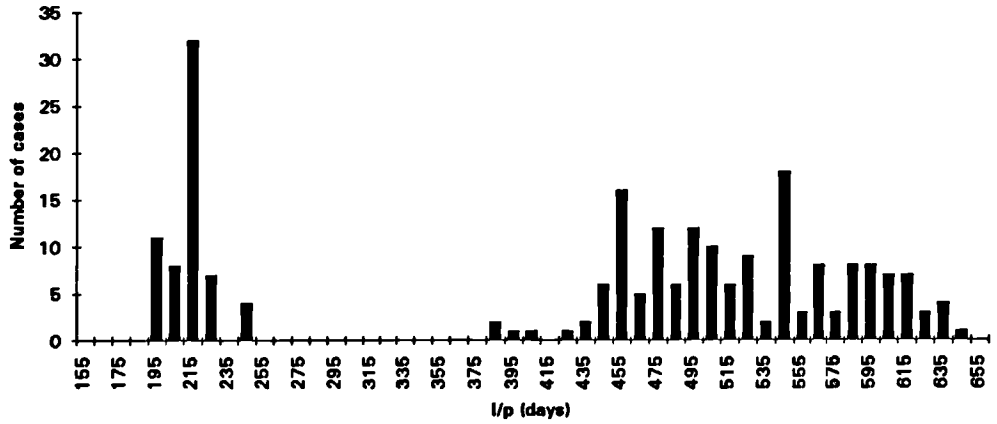


Figure 9.2.3(b). Scrapie in mice; strain 22A; i/p f/d, genotype unrecorded



The three genotypes are each closely grouped. The unrecorded genotype appears to be a mixture of all three. The i/p f/d analysis is given in table 9.2.2(ii).

Table 9.2.3(ii) Scrapie in mice; strain 22A; i/p f/d analysis by 'sinc' genotype.

'Sinc' genotype recorded	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
s7s7	27	Normality	0.982	0.960	-0.408	0.317	NS
s7p7	47	"	0.991	0.974	-0.454	0.168	NS
p7p7	29	"	0.993	0.962	0.508	0.203	NS
s7s7	27	Log-normality	0.979	0.960	-0.559	0.175	NS
s7p7	47		0.985	0.974	-0.607	0.071	NS
p7p7	29	"	0.996	0.962	0.364	0.355	NS

All genotype datasets are compatible with both normality and log-normality; none has a significant skew.

9.2.4 Analysis of strain ME7 by 'sinc' genotype.

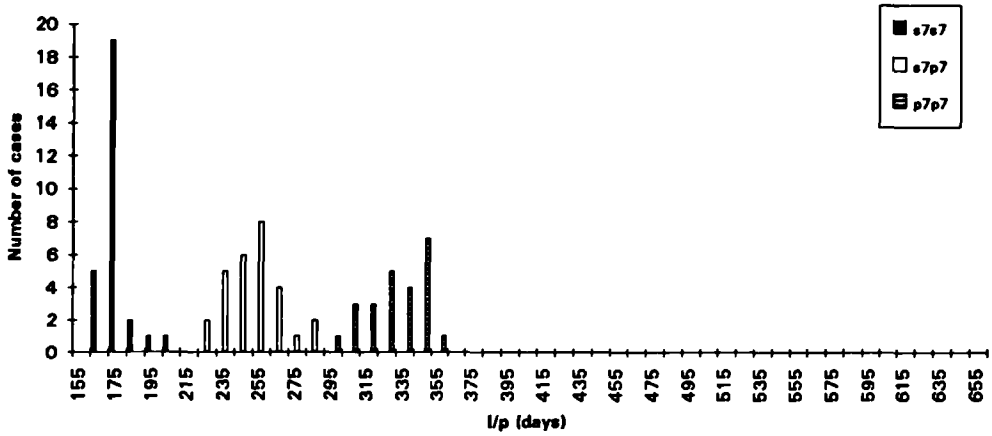
Table 9.2.4(i), i/p data by 'sinc' genotype. One unrecorded genotype case had an i/p of 215 days.

Table 9.2.4(i). Scrapie in mice; strain ME7; i/p data (in days) by 'sinc' genotype.

'Sinc' genotype recorded	N	Mean i/p	SD	Range	Mode	Median
s7s7	28	176	9	165-205	175	175
s7p7	28	251	16	225-285	255	255
p7p7	24	329	16	295-355	345	330

Mean i/ps for the three genotypes are again very different (P<0.001, non-parametric test). There is also a different order of increasing mean i/p magnitude. Histogram, figure 9.2.4.

Figure 9.2.4.Scrapie in mice; strain ME7; i/p f/d by 'sinc' genotype



Again all genotypes are closely grouped and the different sequence is clear. I/p f/d analysis is given in table 9.2.4 (ii).

Table 9.2.4(ii). Scrapie in mice; strain ME7; i/p f/d analysis by 'sinc' genotype

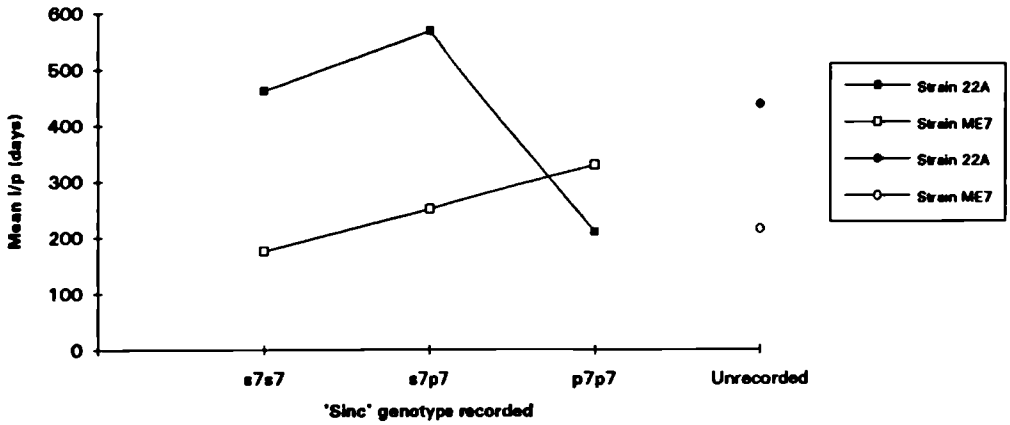
'Sinc' genotype recorded	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
s7s7	28	Normality	0.956*	0.962	1.659	0.001	R
s7p7	28	"	0.995	0.962	0.376	0.346	NS
p7p7	24	"	0.991	0.957	-0.381	0.367	NS
s7s7	28	Log-normality	0.967	0.962	1.447	0.002	R
s7p7	28		0.998	0.962	0.231	0.559	NS
p7p7	24	"	0.989	0.957	-0.452	0.288	NS

The s7s7 group is incompatible with normality (*), but just compatible with log-normality. Both raw and log-transformed data are right skewed, the extent of which is not immediately apparent visually. The analysis result may be influenced by the exaggerated mode. Both other genotypes are compatible with both normality and log-normality, with no significant skew present.

9.2.5 Interaction of 'sinc' genotype and agent strain.

Mean i/ps by genotype vary differentially with agent strain as shown in figure 9.2.5.

Figure 9.2.5 Scrapie in mice: mean i/p by 'sinc' genotype and strain



Results of ANOVA for i/p using variables strain and 'sinc' genotype are given in table 9.2.5; cases with unrecorded genotype are not included. No data transformation was necessary.

Table 9.2.5 Scrapie in mice; analysis of strain & 'sinc' genotype by ANOVA.

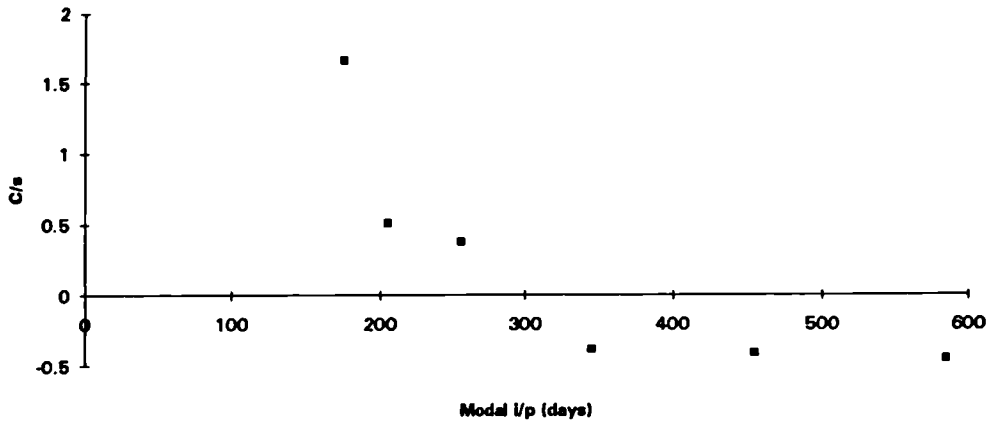
Variable	Stratum	N	F	P	Mean i/p(adj)	95% CI
Strain	22A	103	2223.56	<0.001	413	408-418
	ME7	80			252	247-257
'Sinc' genotype	s7s7	55	616.29	<0.001	319	313-325
	s7p7	75			410	405-415
	p7p7	53			269	263-275
Interaction term	Strain & 'sinc'	-	1618.96	<0.001	-	-
Strain 22A, by genotype	s7s7	27			462	453-471
	s7p7	47			568	562-574
	p7p7	29			210	202-218
Strain ME7, by genotype	s7s7	28			176	168-184
	s7p7	28			251	243-259
	p7p7	24			329	320-338

Strain and genotype are both statistically highly significant, as is the interaction between them. The adjusted mean i/p for strain 22A is 161 days longer than that for ME7, and the adjusted mean i/p for the heterozygous allele genotype s7p7 is around 100 or more days longer than for either homozygous genotype. The mean i/p categorised by strain and genotype is as plotted earlier.

9.2.6 Association of mean and modal incubation period with frequency distribution coefficient of skewness.

Figure 9.2.6 is a plot of modal i/p and c/s for each 'sinc' gene/agent strain combination. Mean i/p's (not shown) give very similar results.

Figure 9.2.6 Scrapie in mice: association between f/d c/s and modal i/p



Regression analysis gives a c/s reduction of 0.04 units per additional 10 days for modal i/p (r= 0.81, F=7.72, P=0.050). Comparable results for mean i/p are 0.04 c/s units (r=0.81, F=7.48, P=0.52). However, both plots appear curved; equations fitted (by R Sayers) to take this into account give the following relationships:-

For modal i/p:- C/s = -(0.4307) + [(35.1452) x exp(-0.0164 x i/p)] (r=0.97)

For mean i/p:- C/s = -(0.4446) + [(38.9451) x exp(-0.0167 x i/p)] (r=0.98)

9.3 Experimental CJD in guinea-pigs.

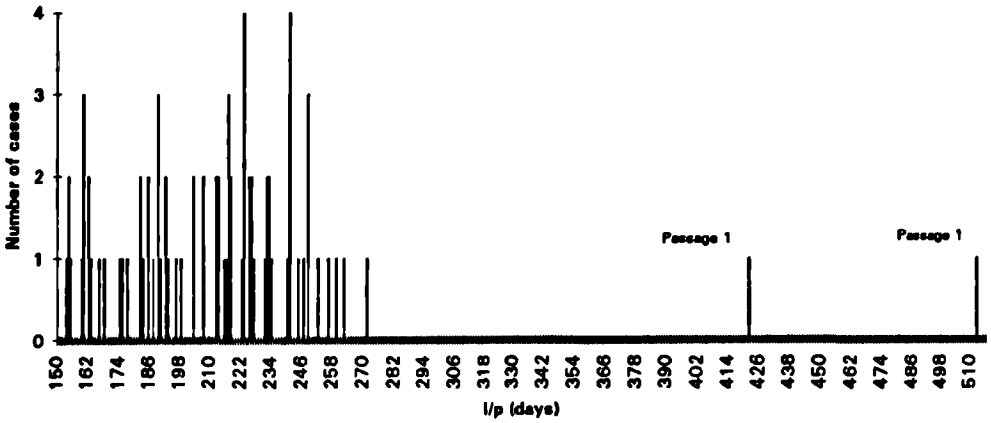
9.3.1 Analysis of full dataset.

I/p data (in days), for 79 cases:-

Mean:	214
SD:	52
Range:	153-512
Mode:	223
Median:	215

A histogram of i/p f/d is given (figure 9.3.1) with the only two 'first passage' cases (see appendix 10) indicated.

Figure 9.3.1 CJD in guinea-pigs; full dataset; i/p f/d.



Right skewness appears to result from the extended i/p of the two first passage cases; the remainder are relatively clustered. Table 9.3.1 gives the i/p f/d analysis.

Table 9.3.1 CJD in guinea-pigs; full dataset; f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
79	Normality	0.840	0.984	3.133	<0.001
79	Log-normality	0.935	0.984	1.355	<0.001

The data is incompatible with both normality and log-normality, with an extensive right skew.

9.3.2 Effect of passage number on incubation period.

I/p data (in days) by passage number is given in table 9.3.2 (i).

Table 9.3.2(i) CJD in guinea-pigs; i/p data (in days) by passage number.

Passage No	N	Mean i/p	SD	Range	Mode	Median
1	2	467	64	422, 512	-	467
2	10	217	21	187-259	226	220
3	27	231	16	193-256	241	233
4	14	222	24	192-271	223	218
5	6	189	14	177-215	182	185
6	5	160	6	153-168	-	160
7	3	158	4	154-162	-	159
8	3	179	11	166-185	185	185
9	3	161	2	160-163	160	160
10	6	179	15	154-196	-	179

[For some passage numbers, no mode was identified]

First passage cases have much longer i/p's than other 'passage number' cases. Regression analysis gives a reduction in mean i/p of 12.4 days per passage ($r=-0.58$ $F=39.60$, $P<0.001$). Much of this occurs at first passage, although a lesser decrease is apparent over subsequent passages. Table 9.3.2(ii) gives regression analysis results sequentially omitting one passage, up to passage 5.

Table 9.3.2(ii). CJD in guinea-pigs; effect of passage number on i/p.

Passage from number:-	Mean i/p decrease per passage (days)	r	F	P
2	9.0	-0.69	69.76	<0.001
3	10.1	-0.74	77.44	<0.001
4	8.0	-0.60	21.03	<0.001
5	0.8	-0.10	0.25	0.621

Passage of agent continues to reduce the subsequent mean i/p up to and including inoculation into the fourth experimental animal; after this, mean i/p appears to stabilise.

9.3.3 Analysis of incubation period frequency distribution for passage numbers with 10 or more cases.

From the f/d histograms (not shown), no marked skew is apparent for any set. Table 9.3.3 gives the i/p f/d analysis.

Table 9.3.3 CJD in guinea-pigs; f/d analyses for passage numbers 2-4.

Passage No	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
2	10	Normality	0.955	0.918	0.324	0.564	NS
3	27	"	0.986	0.961	-0.488	0.233	NS
4	14	"	0.973	0.934	0.737	0.158	NS
2	10	Log-	0.959	0.918	0.092	0.869	NS
3	27	normality	0.980	0.961	-0.629	0.130	NS
4	14	"	0.982	0.934	0.568	0.269	NS

These small data sets are all compatible with both normality and log-normality.

9.4 Discussion.

9.4.1 BSE in mice.

9.4.1.1 Incubation period frequency distribution analysis findings.

Table 9.4.1.1(i) gives f/d analyses findings, table 9.4.1.1(ii), the f/d description.

Table 9.4.1.1(i). BSE in mice; summary of f/d analyses.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
Full dataset	196	No	No	R	R
BSE group 1	35	Yes	Yes	NS	NS
BSE group 2	32	No	No	R	R
BSE group 3	30	Yes	Yes	NS	NS
BSE group 4	31	Yes	Yes	NS	NS
BSE group 5	23	Yes	Yes	NS	NS
BSE group 6	22	Yes	Yes	NS	NS
BSE group 7	23	No	No	R	R

[Note:- All significance levels set at P=0.05]

Table 9.4.1.1(ii). BSE in mice; descriptions of f/d.

Data set	Description
Full dataset	Right skew too elongated even for log-normality
BSE group 1	Compatible with both normality and log-normality
BSE group 2	Right skew too extensive even for log-normality
BSE group 3	Compatible with both normality and log-normality
BSE group 4	"
BSE group 5	"
BSE group 6	"
BSE group 7	Right skew too extensive even for log-normality

The full dataset demonstrates that an i/p f/d more right skewed than log-normality is possible for BSE in mice. When categorised by BSE group, i/p f/d for 5 of 7 groups are compatible with both normality and log-normality, probably reflecting lack of statistical power at small sample sizes; for these groups there was no correlation between c/s and sample size, but all comprise 35 or fewer cases.

9.4.1.2 Association between inoculation age and incubation period.

Increase of inoculation age results in an i/p reduction of 0.16 days per day; this is of broadly similar order to estimated i/p decrease of 0.05 +/- 0.01 days per day for inoculation of mice using a scrapie agent strain (Dickinson & Mieckle, 1971).

However, both inoculation age and i/p are associated with BSE group. As group mean inoculation age increases, so group mean i/p (dependant on observed data range) tends to decrease, whereas for modal i/p (not dependant on observed range) the effect is much less marked (although neither associations are statistically significant). Brain titre of agent may differ considerably resulting in different mean i/p's by BSE group, but this is unlikely to correlate with inoculation age, and Bruce et al (1994) conclude that only a single 'strain' of BSE agent exists. It therefore seems likely that observed effects are due to inoculation age rather than BSE group.

Since BSE groups 2 and 7 (the most right skewed groups) have youngest mean inoculation ages, data was regrouped to examine the association between inoculation age and i/p f/d with reduced BSE group effect (although complete removal is impossible). Both regrouping by similar inoculation age range (A-D) and by similar sample size (E-H) confirms that younger mean inoculation age in a particular group results in a tendency for i/p f/d to be more right skewed.

One explanation is that groups inoculated at younger mean ages are able to manifest a much greater range of i/p's, whereas for groups inoculated at later ages, too little time remains to observe the longest possible i/p's. This might imply that incubating mice, inoculated later in life, will die from other causes but, although 17 (8%) mice did die from other causes, these were generally shortly after inoculation, implicating trauma or bacterial infection at inoculation.

Additionally, the mean lifespan of RIII mice at the laboratory concerned is considerably longer than the oldest age at death in this data (H Fraser, personal communication). Totalling the longest i/p and oldest inoculation age gives 562 days whereas, in one survival analysis, mean age at 'natural' death for 78 mice was approximately 590 days (SD = 222; range approx 175-1025). One 'normal' RIII mouse survived to over 1100 days of age.

It therefore seems unlikely that not enough 'time' was left for older-inoculated mice to manifest an i/p f/d which was more right-skewed. Perhaps infection later in life speeds the incubation process in some way, but there is no evidence that the modal i/p is shortened; maybe interaction between inoculation age and host-response shortens i/p in only a particular subset of cases.

9.4.2 Scrapie in mice.

9.4.2.1 Incubation period frequency distribution analysis findings.

Table 9.4.2.1(i)gives the f/d analyses findings, table 9.4.2.1(ii), the f/d description.

Table 9.4.2.1(i). Scrapie in mice; summary of f/d analyses.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
Strain 22A: s7s7	27	Yes	Yes	NS	NS
Strain 22A: s7p7	47	Yes	Yes	NS	NS
Strain 22A: p7p7	29	Yes	Yes	NS	NS
Strain ME7: s7s7	28	No	Yes	R	R
Strain ME7: s7p7	28	Yes	Yes	NS	NS
Strain ME7: p7p7	24	Yes	Yes	NS	NS

Table 9.4.2.1(ii). Scrapie in mice; descriptions of f/d.

Dataset	Description
Strain 22A: s7s7	Compatible with both normality and log-normality
Strain 22A: s7p7	"
Strain 22A: p7p7	"
Strain ME7: s7s7	Exaggerated mode; right skewed
Strain ME7: s7p7	Compatible with both normality and log-normality
Strain ME7: p7p7	"

Of the ten datasets examined, all four which are not single-agent and single-mouse strain combined appear tri-modal; clearly they are neither normal nor log-normal.

The remaining 6 single-agent plus single-mouse strain datasets all have fewer than 50 cases. All except one are compatible with both normality and log-normality. For that which is not, although analysis gives a right skew, there are no distant right outliers in the f/d; this right skewness is therefore likely to be a reflection of the effect of the very exaggerated mode. In summary, there are no datasets with very extended right 'tails'.

9.4.2.2 Association between incubation period, agent strain, and mouse strain.

Scrapie agent strains are 'isolated' and stabilised by passage through experimental animals until reproducibility of i/p is achieved.

With such well characterised agent strains, inoculated into two genetically different but highly inbred strains of mice and their crosses, the six combinations give characteristic, tightly-clustered, uni-modal i/p ranges; mixed datasets reflect this by assuming multi-modality. The unrecorded genotype, 22A strain combination, has an appearance of tri-modality from which probable genotype can be predicted, using cluster characteristics.

Overall genetic uniformity, as well as specific 'sinc' genotype may be important in development of characteristic clustered i/p ranges. If other variables (e.g. age at inoculation) are also standardised, it may be that clustering would be tightened further. Perhaps when multiple values per variable exist (generally so in 'natural' situations), characteristic i/p clustering is lost resulting in a wide i/p range, with occasional extremely long i/p's giving an f/d more right-skewed than log-normality. However, even in genetically mixed populations, an appropriate equivalent of the 'sinc' gene might result in less readily apparent clustering tendencies about various modal values.

However, although all six combinations showed compactness of i/p range, all were small samples. Large samples for each specific combination might result in occasional extreme right outliers, as in other datasets, which would alter mean, but not modal i/p.

9.4.2.3 Association between modal and mean incubation period and coefficient of skewness.

Although the c/s was not statistically significant for five of the six mouse/agent combinations, there was a statistically significant reduction in c/s with increasing modal i/p, and almost so for mean i/p. However, inoculation age was stated to vary from 22-152 days, therefore this result cannot be directly equated to a reduction of c/s with increasing mean a/o, unless mean inoculation age was similar for each group, and this is not known. (Since sample sizes are all small and fairly similar, any effect of sample size on c/s would be expected to affect each group similarly.)

9.4.3 CJD in guinea-pigs

9.4.3.1 Incubation period frequency distribution analysis findings.

Table 9.4.3.1(i) gives the f/d analyses findings, table 9.4.3.1(ii) the f/d description.

Table 9.4.3.1 CJD in guinea-pigs; summary of f/d analyses.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
Full set	79	No	No	R	R
Passage 2	10	Yes	Yes	NS	NS
Passage 3	27	Yes	Yes	NS	NS
Passage 4	14	Yes	Yes	NS	NS

Table 9.4.3.1(ii). CJD in guinea-pigs; descriptions of f/d.

Dataset	Description.
Full set	Right skew too extensive even for log-normality
Passage number 2	Compatible with both normality and log-normality
Passage number 3	"
Passage number 4	"

The full dataset was very right skewed, due almost entirely to the two 'first passage' cases. These are also different from each other, at 422 and 512 days; it is assumed that one is derived from each of the two human cases mentioned in the source paper. On subsequent guinea-pig to guinea-pig passage the three small dataset f/d's were all compatible with both normality and log-normality.

9.4.3.2 Effect of passage on incubation period.

Passage of CJD agent reduces the subsequent mean i/p up to and including inoculation into the fourth guinea-pig and after this it appears to stabilise. The greatest reduction is on passage from first to second guinea-pig (46.5%), after which reduction per passage is much less marked, reaching 40.5% of original i/p by the fifth guinea-pig.

9.4.4 Comparison with Sartwell's conclusions.

Both mice inoculated with BSE agent, and guinea-pigs inoculated with CJD agent demonstrate that, with experimental variation present (age at inoculation, and passage number respectively), a dataset with a right skew too extended for compatibility with log-normality is possible.

Reduction of variability by using highly characterised mouse and scrapie agent strains resulted in substantial narrowing of i/p ranges, which were then generally compatible both with normality and log-normality. However, these data sets were also small; although they do not conflict with Sartwell's hypothesis, it is not confirmed.

CHAPTER 10

HUMANS: DATA ANALYSIS AND DISCUSSION: Natural and iatrogenic TSE/prion disease.

10. Summary.

10.1 Analyses of sporadic cases of TSE.

10.2 Analyses of cases of TSE with familial history or PrP gene mutation; NIH and GLS data.

10.2.1 Analysis of all cases with known familial history.

10.2.2 Analysis of cases with known PrP gene mutation, either in the case or another affected family member.

- 10.2.2.1 Mean age at onset by genetic mutation recorded.
- 10.2.2.2 Analysis of age at onset frequency distribution by genetic mutation.
- 10.2.2.3 Association between coefficient of skewness, and mean and modal age at onset.
- 10.2.2.4 The 'met/val' heterogeneity factor.

10.3 Analysis of transmitted cases.

10.3.1 Kuru.

10.3.2 Analysis of pituitary hormone injection cases.

- 10.3.2.1 Analysis of data supplied by Dr R Will.
- 10.3.2.2 Analysis of data on French cases supplied by Dr P Brown.

10.3.3. Analysis of other iatrogenic cases.

10.4 Discussion

10.4.1 Age at onset and incubation period frequency distribution analysis findings.

- 10.4.1.1 Comparison of age at onset frequency distribution in sporadic and familial TSE datasets.
- 10.4.1.2 Cases with recognised transmission routes: age at onset frequency distribution findings.
- 10.4.1.3 Cases with recognised transmission routes: incubation period frequency distribution findings.

10.4.1.4 Cases with mutation of PrP gene recorded: association between age at onset frequency distribution, and mean and modal age at onset.

10.4.2 Summary of the findings from the analysis of the other variables.

10.4.3 Comparison with Sartwell's conclusions.

CHAPTER 10

HUMANS: DATA ANALYSIS AND DISCUSSION.

Natural and iatrogenic TSE/prion disease.

Summary.

The frequency distribution (f/d) of age at onset (a/o) for cases with no known risk factor (sporadic disease) has an extended left skew incompatible with normality, and mean a/o of around 60 years. For non-sporadic cases (those with identified risk factors of familiarity or PrP gene mutation) a/o f/d is also left skewed, although mean a/o reduces to 45-50 years. If a/o approximates to incubation period (i/p), Sartwell's hypothesis of log-normality does not apply.

Cases grouped by PrP gene mutation have characteristic a/o means and modal regions; mean a/o varies from 33 to 55 years. There is evidence that f/d coefficient of skewness (c/s) is associated with modal (and mean) a/o, right skewness decreasing as modal a/o increases. The complete non-sporadic f/d thus comprises discrete modal regions and f/d's, each associated with a particular mutation. Further genetic analysis should allow further discrimination into individual f/d's.

Datasets for recognised infectious transmission route cases (kuru, iatrogenic) are small, but both i/p f/d and a/o f/d tend to right skewness; modal a/o's are earlier than for other groups and again may be associated with c/s. Small sample sizes makes confirmation of Sartwell's hypothesis impossible.

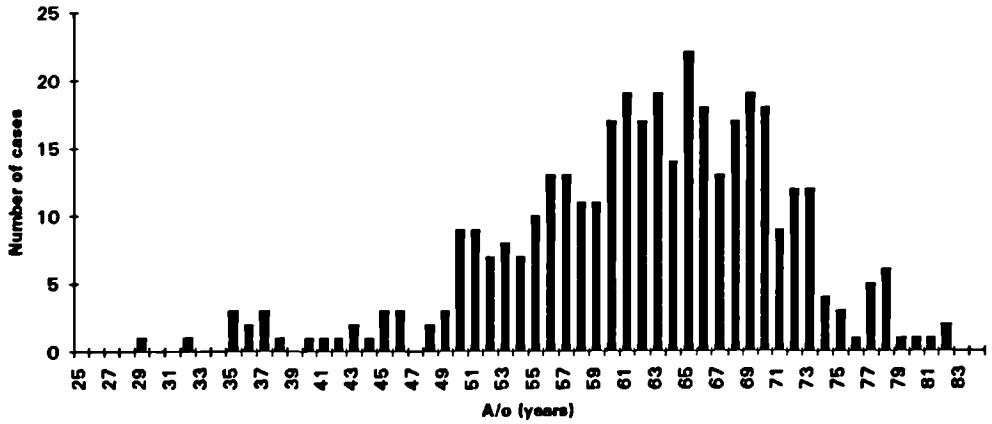
10.1 Analyses of sporadic cases of TSE.

The a/o data (in years) for a/o for the three data sources (see appendix 10 for all human data source and categorisation details, unless otherwise indicated):-

	UK data	NIH data	GLS data
Mean a/o:	62.0	61.6	56.3
SD:	9.2	9.3	10.7
Range:	29-82	16-83	23-80
Mode/modal region:	65	59-67	53-70
Median:	63	62	57
N:	376	501	97

There is no difference in mean a/o by sex for any dataset (UK data, P = 0.162; NIH data, P = 0 499; GLS data, P= 0.706; non-parametric test). The mean GLS a/o is younger than the other two dataset a/o's. The histogram for a/o f/d for UK dataset is shown (figure 10.1). The f/d histograms for NIH and GLS data (not shown) are broadly similar.

Figure 10.1 Sporadic CJD in humans; UK dataset; a/o f/d.



A left skew is apparent, the mode being at 65 years. There maybe a suggestion of bi-modality, at around 35 to 37 years. Results of the a/o f/d analysis for all three datasets are given (table 10.1).

Table 10.1 Sporadic TSE in humans: a/o f/d analysis by data source.

Dataset	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
UK	376	Normality	0.984	0.995	-0.728	<0.001
NIH	501	"	0.985	0.997	-0.754	<0.001
GLS	97	"	0.984	0.986	-0.676	0.007
UK	376	Log-	0.954	0.995	-1.372	<0.001
NIH	501	normality	0.930	0.997	-2.205	<0.001
GLS	97	"	0.948	0.986	-1.394	<0.001

As predicted from the histograms, none of these datasets have an f/d compatible with log-normality. In each dataset there is a marked left skew too extended even for compatibility with normality.

10.2 Analyses of cases of TSE with familial history or PrP gene mutation; NIH and GLS data.

10.2.1 Analysis of all cases with known familial history.

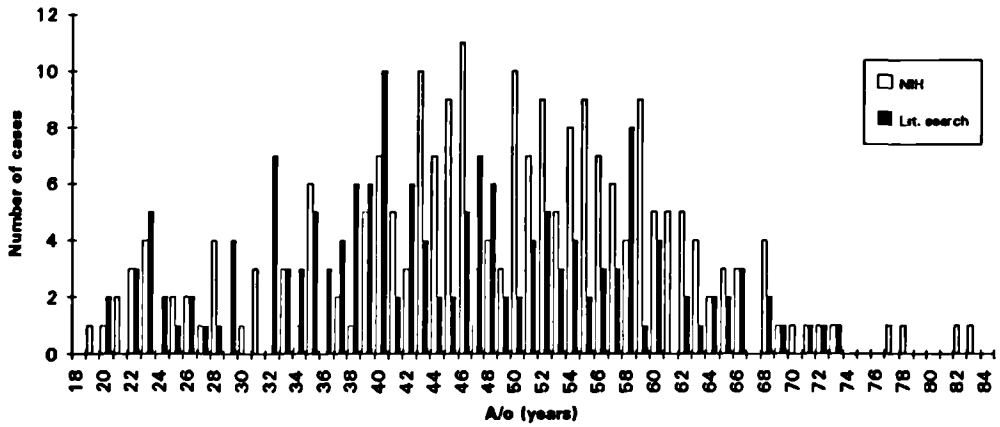
The a/o data (in years) for a/o for the two datasets is given in table 10 2.1(i).

Table 10.2.1(i). Familial TSE in humans: a/o data (in years) by data source.

Data source	N	Mean a/o	SD	Range	Mode	Median
NIH	218	48.8	12.6	19-83	46	50
GLS	159	44.5	12.5	20-73	40	43

Within neither dataset is there any difference in the mean a/o categorised by parent-positive disease status, whether for mother, father, or both parents (P values ranged from 0.118 to 0.436, non-parametric tests). However, there is a difference in mean a/o between the two datasets (P=0.001, non-parametric test). The histogram of a/o f/d by data source is shown (figure 10.2.1).

Figure 10.2.1 Familial TSE in humans; a/o f/d by data source



The f/d's are not markedly dissimilar and no particular skew is obvious. The a/o f/d analysis results are given in table 10.2.1(ii).

Table 10.2.1 (ii) Familial TSE in humans: a/o f/d analysis by data source.

Dataset	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
NIH	218	Normality	0.994	0.993	-0.207	0.202
GLS	159	"	0.995	0.989	0.102	0.585
NIH	218	Log-normality	0.964	0.993	-1.012	<0.001
GLS	159		0.986	0.989	-0.517	0.008

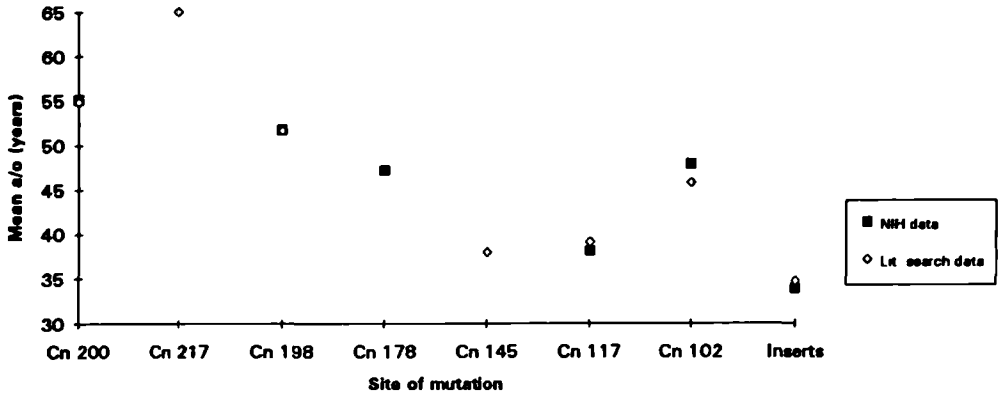
Both datasets are compatible with normality, but not log-normality. There is no statistically significant skew in the raw data.

10.2.2 Analysis of cases with known PrP gene mutation, either in the case or another affected family member.

10.2.2.1 Mean age at onset by genetic mutation recorded.

The mean a/o for each codon (Cn) mutation by data source is plotted (figure 10.2.2 1).

Figure 10.2.2.1 Familial TSE in humans; mean a/o by PrP gene mutation and data source.



The two data sources give very similar mean a/o's. Other values for the two data sources (not shown) are also similar (see appendix 23 for method of estimating amount of data overlap by source). Because the pairs are so similar, only the larger dataset in each pair is used in further analysis. The a/o data (in years) by mutation recorded is given in table 10.2.2.1.

Table 10.2.2.1 TSE in humans: a/o data (in years) by PrP gene mutation.

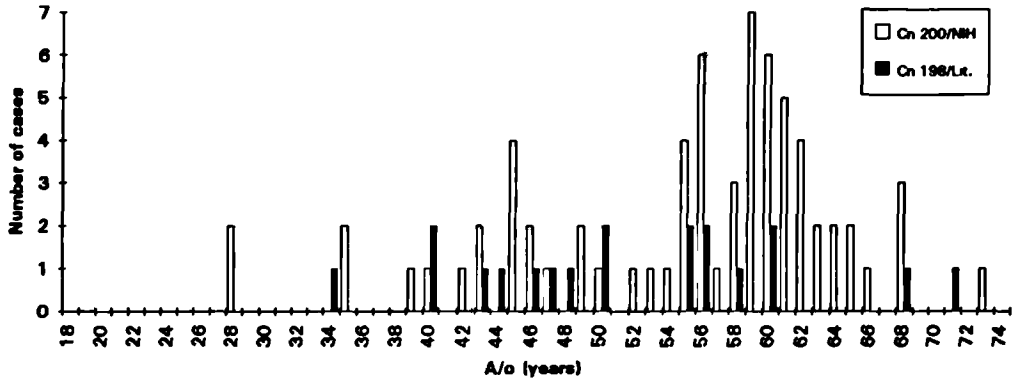
Mutation	Dataset	N	Mean a/o	SD	Range	Mode/al region	Median
Codon 217	GLS	1	65	-	-	-	-
Codon 200	NIH	69	55.1	9.4	28-73	59	58
Codon 198	GLS	19	51.6	9.6	34-71	40-60	50
Codon 178	NIH	73	47.2	9.2	19-72	46	48
Codon 145	GLS	1	38	-	-	-	-
Codon 117	GLS	25	39.1	11.3	20-64	23-42	38
Codon 102	GLS	29	45.8	12.0	22-66	39	48
Insert	GLS	29	34.7	7.9	22-53	32	35

10.2.2.2 Analysis of age at onset frequency distribution by genetic mutation.

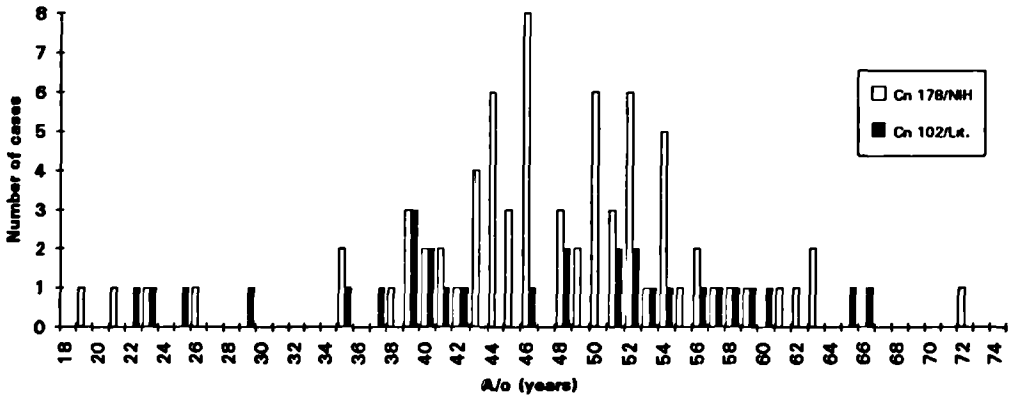
The histograms of the a/o f/d for the six mutations with more than one case are shown (figure 10.2.2.2 (a)-(c)).

Figure 10.2.2.2 TSE in humans; a/o f/d by PrP gene mutation.

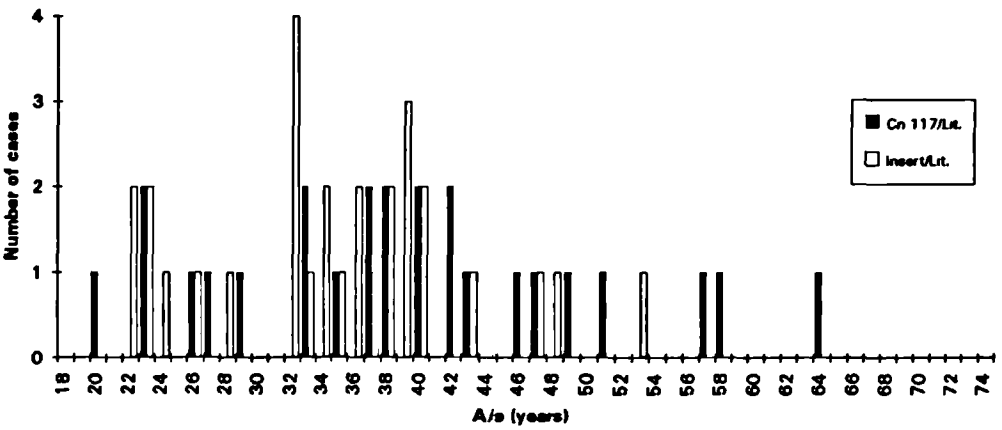
(a) Mutations associated with a 'late' mean a/o; at codons 200 and 198



(b) Mutations associated with an 'intermediate' mean a/o; at codons 178 and 102



(c) Mutations associated with an 'early' mean a/o; at codon 117, and inserts



Inspection of the histograms suggests an association between the mean a/o and f/d skewness, particularly for the larger samples. Table 10.2.2.2 gives a/o f/d analysis results.

Table 10.2.2.2 TSE in humans; a/o f/d analysis by PrP gene mutation category.

Mutation	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
Codon 200	69	Normality	0.962*	0.982	-0.955	0.002
Codon 198	19	"	0.993	0.948	0.209	0.646
Codon 178	73	"	0.971*	0.983	-0.643	0.023
Codon 102	29	"	0.987	0.963	-0.327	0.404
Codon 117	25	"	0.992	0.958	0.304	0.463
Insert	29	"	0.986	0.963	0.164	0.673
Codon 200	69	Log-normality	0.925*	0.982	-1.491	<0.001
Codon 198	19	"	0.994	0.948	-0.198	0.663
Codon 178	73	"	0.915*	0.983	-1.783	<0.001
Codon 102	29	"	0.960*	0.963	-0.880	0.037
Codon 117	25	"	0.991	0.958	-0.319	0.442
Insert	29	"	0.979	0.963	-0.351	0.372

[Note: asterisks indicate rejection of the hypothesis of normality or log-normality at this level]

Analysis confirms the visual impression that for codon 200 mutations, a comparatively large dataset, the a/o f/d is left skewed and incompatible with both normality and log-normality. Codon 198 data, a very small dataset, is compatible with both.

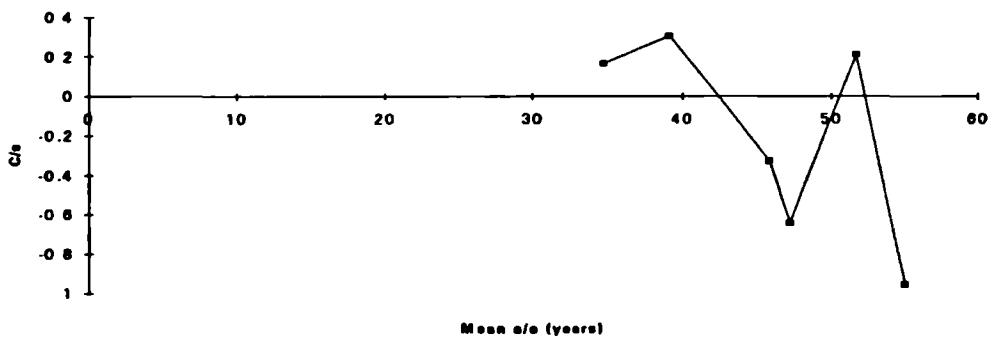
For codon 178 mutations, also a large sample, the a/o f/d is again left skewed and incompatible with both normality and log-normality. Codon 102 mutation data is compatible with normality, but not log-normality.

For codon 117 mutations, and the group of inserts, both of which are small datasets, results are compatible with both normality and log-normality.

10.2.2.3 Association between coefficient of skewness and mean and modal age at onset.

Figure 10.2.2.3 is a plot of the mean a/o by the c/s for each mutation.

Figure 10.2.23 TSE in humans; f/d c/s by mean a/o and PrP gene mutations.



Regression analysis gives a non-significant reduction in c/s of 0.043 units for each 'additional year to mean a/o' (r=0.64; F=2.85, P=0 166). Although for small samples, mode identification

proved difficult, the mode, or mid-point of the observed 'modal region' is compared with c/s. The plot (not shown) is similar to that shown above. Regression analysis again gives a non-significant reduction, 0.035 units per additional year (r=0.70, F=3.91, P=0.119).

10.2.2.4 The 'met/val' heterogeneity factor.

For 36 cases, codon 129 met/val status was known. Table 10.2.2.4 gives a comparison of the a/o for these cases with the overall a/o for the primary mutation. Case numbers are given in brackets.

Table 10.2.2.4 TSE in humans: effect of met/val heterogeneity on mean a/o (data in years).

Primary mutation	Mean a/o for primary mutation	Codon 129 alleles; mean a/o plus case numbers		
		met/met	met/val	val/val
Codon 178	47.2	46.3 (3)	48.2 (10)	41.0 (7)
Codon 200	54.9	63.5 (2)	-	-
Codon 198	51.6	-	56.0 (1)	44.0 (2)
Codon 117	39.1	-	40.4 (11)	-

With such small samples, further analysis is not indicated.

10.3 Analysis of transmitted cases.

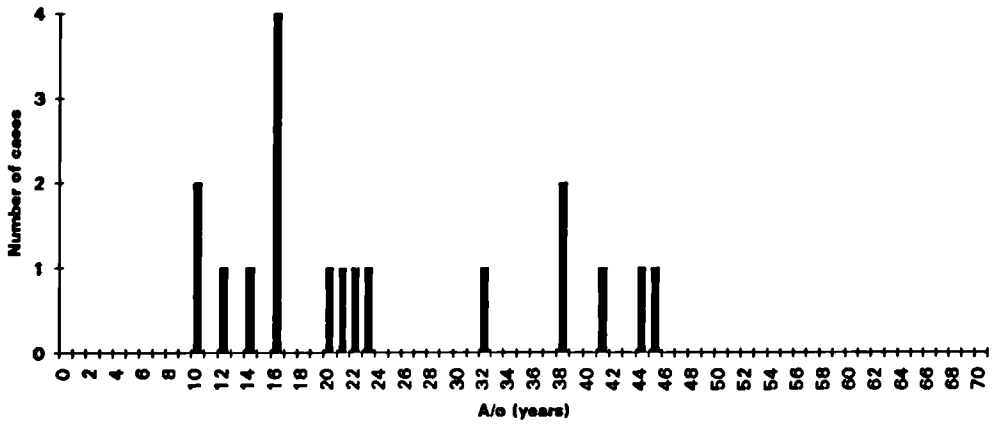
10.3.1 Kuru.

The a/o details (in years) for 18 cases:-

Mean a/o:	24.1
SD:	12.1
Range:	10-45
Mode:	16
Median:	21

There is no difference in the mean a/o by sex (9 of each; P= 0 859, non-parametric test). The a/o f/d histogram is shown (figure 10.3.1).

Figure 10.3.1 Kuru in humans; a/o f/d.



The modal a/o, at 16 years, is much younger than for previous human datasets. The histogram appears right skewed. The f/d analysis are given in table 10.3.1.

Table 10.3.1 Kuru in humans: a/o f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
18	Normality	0.952	0.946	0.573	0.223
18	Log-normality	0.978	0.946	0.123	0.790

Neither the hypotheses of normality nor log-normality can be rejected. There is a statistically insignificant right skew, however the dataset is small.

10.3.2 Analysis of pituitary hormone injection cases.

10.3.2.1 Analysis of data supplied by Dr R Will.

No significant differences were found by sex, for a/o or i/p analyses. The histogram for the a/o (not shown) has no obvious asymmetry. The modal region is 20-32 and the median is at 22. The f/d is compatible both with normality and log-normality, and is thus compatible with the visual impression of symmetry (c/s=0.178; P=0.675).

The i/p (estimated as years from first treatment to disease onset) data for 23 cases:-

Mean:	15.7
SD:	6.9
Range:	6-35
Modal region:	9-16
Median:	15

The i/p f/d histogram (not shown) has a right skewed appearance. The i/p f/d analysis results are given in table 10.3.2.1.

Table 10.3.2.1 Pituitary injection cases in humans, first set: i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
23	Normality	0.965	0.955	1.010	0.029
23	Log-normality	0.998	0.955	-0.050	0.906

The i/p f/d is compatible with both normality and log-normality. However, the CC value for the log-normal transformation is much larger than for the raw data, and the raw data is right skewed.

No association is apparent between the age at first injection and the estimated i/p. Regression analysis gives an i/p increase of 0.05 years, per additional year of age at commencement of treatment (r=0.05, F=0.05, P=0.821).

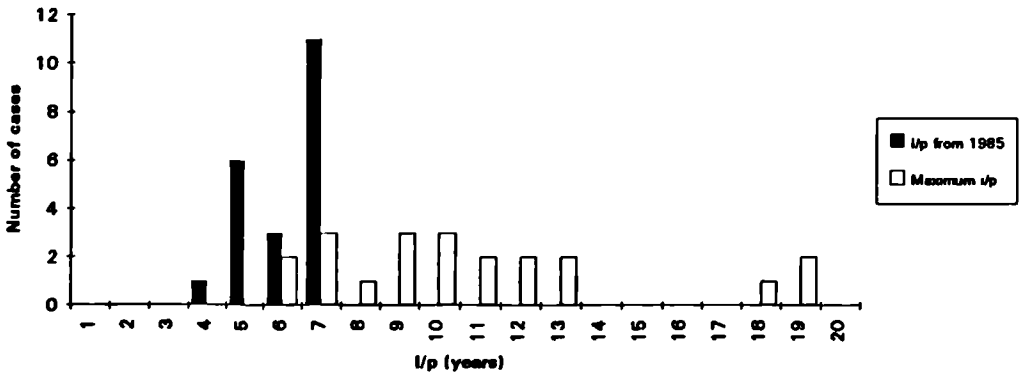
10.3.2.2 Analysis of data on French cases supplied by Dr P Brown.

The i/p is estimated both from first treatment (maximum possible) and from 1985 (probable contamination date; see appendix 10). The i/p (estimated as years from 1985 to disease onset) data for 21 cases:-

Mean:	6.1
SD:	1.0
Range:	4-7
Modal region:	7
Median:	7

The i/p f/d histogram, showing i/p estimate from 1985, plus time from first treatment is shown (figure 10.3.2.2).

Figure 10.3.2.2 Pituitary injection cases in humans, French dataset: i/p f/d assuming infection in 1985, plus f/d of time from first treatment ('maximum i/p').



Assuming infection occurred in 1985, the f/d appears left skewed, with a very marked mode and a maximum at 7 years. However, data was collected up to mid-1993, eight years after 1985. Allowing for a reporting lag, 7 years is probably the longest observable i/p duration under this assumption. Table 10.3.2.2 gives i/p f/d analysis, using the i/p estimated from 1985.

Table 10.3.2.2 Pituitary injection cases in humans, French set: i/p f/d analysis.

N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
21	Normality	0.997	0.952	-0.583	0.197
21	Log-normality	0.997	0.952	-0.782	0.091

Neither the hypotheses of normality nor log-normality can be rejected, although the left skew approaches significance in the log-transformed data.

10.3.3. Analysis of other iatrogenic cases.

The data comprises 9 dura mata grafts, one corneal graft (i/p 18 months), and one transmission by contaminated electrode (i/p 20 months). Table 10.3.3 gives the 9 dura graft details.

Table 10.3.3 Dura graft cases in humans; i/p and a/o data.

Factor examined	Mean	SD	Range	Modal region	Median
I/p (months**)	63.1	36.1	19-120	30-96	45
Age at graft (years)	26.1	9.5	10-44	23	23
A/o (years)	31.3	10.1	18-51	26	28

The minimum i/p is less than 2 years, the maximum 10. The f/d histograms for both age at graft and a/o (not shown) appear right skewed. Six (66%) of the grafts occur from ages 22-26. For the i/p f/d histogram (not shown) for dura grafts alone, no particular skew is apparent. However, inclusion of corneal and electrode transmission visually suggests the possibility of a right skew (c/s = 0.577; P=0.296) for iatrogenic CNS infection route cases.

10.4 Discussion

10.4.1 Age at onset and incubation period frequency distribution analysis findings.

10.4.1.1 Comparison of age at onset frequency distribution in sporadic and familial TSE datasets.

Table 10.4.1.1(i) gives the a/o f/d analyses summary.

Table 10.4.1.1(i) Sporadic and familial TSE in humans; summary of a/o f/d analysis.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
Sporadic:					
UK	376	No	No	L	L
NIH	501	No	No	L	L
GLS	97	No	No	L	L
Familial :					
NIH	218	Yes	No	NS	L
GLS	159	Yes	No	NS	L

[Note:- Significance level for all analyses is at the P=0.05 level.]

Table 10.4.1.1(ii) gives the f/d description for each dataset, plus the mean a/o.

Table 10.4.1.1(ii) Sporadic and familial TSE in humans; mean a/o, and f/d descriptions.

Dataset	Description	Mean a/o (years)
Sporadic:		
UK	Left skew too extended for normality	62.0
NIH	"	61.6
GLS	"	56.3
Familial :		
NIH	Compatible with normality	48.8
GLS	"	44.5

No dataset is particularly small, the size ranging from 97 to 501. The a/o ranges observed might therefore be expected to be representative, and were certainly wide for all datasets, from around 20 to 80 years. In both categories (sporadic and familial) GLS data has the youngest mean a/o. This may reflect a publication bias, which may also include 'time bias'; i.e. older publications classifying cases as non-familial due to lack of evidence to the contrary, plus a possible bias introduced by deletion of known duplications.

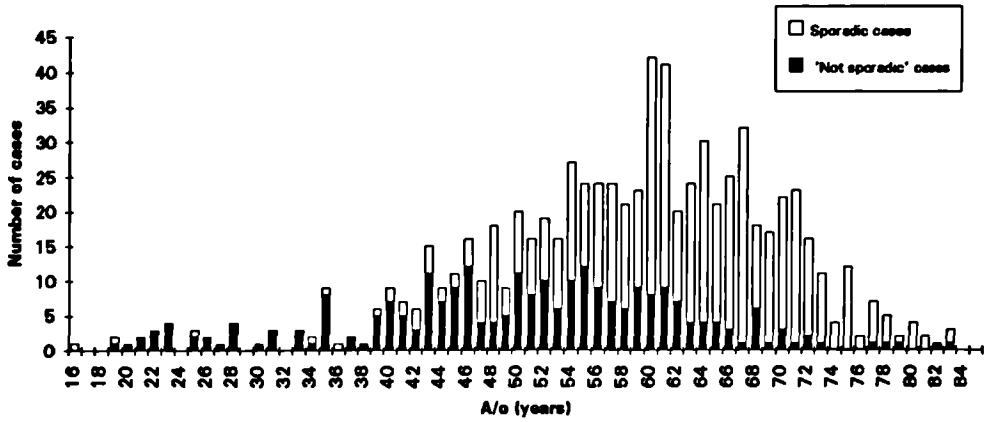
There is a clear difference in the f/d for the two categories, sporadic and familial, as well as a separation of mean a/o values.

Sporadic datasets all have modal regions around 60-70 years, followed by a sharp frequency drop, which may be partly explained by death from competing causes, and partly by older case under-ascertainment (CJD SU Report, 1994) and all have very extended, low frequency left 'tails' incompatible with normality.

Both familial datasets have earlier modal regions, around 40-50 years, and both are compatible with normality, with a higher frequency than sporadic cases in the 20-30 age group.

Figure 10.4.1.1 is a composite (stacked) of NIH sporadic and non-sporadic (defined as having any known risk factor; familial, ethnic, or a known genetic mutation) cases, showing the tendency to earlier onset in cases with a known risk factor, and the overall f/d left skew for TSE/prion disease in humans.

Figure 10.4.1.1 Sporadic and non-sporadic TSE in humans; NIH data; a/o f/d; stacked composite.



None of these datasets demonstrates a right skew, in contrast to the situation with natural scrapie in sheep and goats, and many experimental datasets.

10.4.1.2 Cases with recognised transmission routes: age at onset frequency distribution findings.

Data on a/o was available for three of these datasets. Table 10.4.1.2(i) summarises the a/o f/d analysis findings for the two larger datasets.

Table 10.4.1.2(i) Recognised transmission route cases in humans; summary of a/o f/d analyses.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
Kuru	18	Yes	Yes	NS	NS
Pituitary cases	23	Yes	Yes	NS	NS

[Note:- Significance level for all analyses is at the P=0.05 level.]

These are both small datasets, and from statistical analysis, compatible with both normality and log normality. However, all have much younger mean and modal a/o's than both sporadic and familial TSE's, and, although not statistically significant, all are right skewed. Table 10.4.1.2(ii) summarises this finding.

Table 10.4.1.2(ii) Recognised transmission route cases in humans; a/o f/d descriptions.

Dataset	C/s	Mode/al region a/o	Mean a/o
Kuru	0.573	16	24.1
Pituitary cases	0.178	20-32	25.0
Dura mata cases	Not analysed: right skewed appearance	23	31.3

Whilst the actual a/o for any particular case obviously depends, at least partly, on infection age, it is suggested that the a/o f/d is dependant both on the mean, and particularly the modal a/o of a

given dataset, with an increase in modal and mean a/o associated with a decrease in c/s. With such small samples, this is difficult to demonstrate conclusively.

The oldest-onset case of kuru, at about 45 years, may approach the extent of natural lifespan for remote dwelling natives of Papua New Guinea at the time these cases occurred, in the 1950's.

10.4.1.3 Cases with recognised transmission routes: incubation period frequency distribution findings.

Three such datasets have i/p information. Two comprise pituitary injection cases, treatment generally being given subcutaneously. The third comprises cases where infection site closely involves the CNS; dura grafts, corneal grafts, and brain implantation of contaminated electrodes.

For the first (international) set of pituitary hormone cases (N=23), the estimated i/p f/d is right skewed, and the CC suggests closer compatible with log-normality than normality.

The small dataset of iatrogenic CNS-proximate transmissions (N=11) is also right skewed, though non-significantly. The modal i/p, around 1-2 years, is much shorter than peripherally-inoculated pituitary cases.

The i/p f/d in those cases known to be transmitted therefore appears to resemble experimental datasets, in that the tendency is to right skewness, and an association is suggested between infection route, particularly proximity to the CNS, and mean i/p, which would also be compatible with experimental evidence.

For the second (French) pituitary dataset (N=21) the relevant f/d is probably that estimated from 1985. The rapid frequency increase and sharp truncation at 7 years suggests an epidemic only part-way through. If the f/d is similar to that of many experimental datasets, a marked right skewness is to be expected. If total numbers treated in 1985 were large, such an f/d indicates that many more cases are likely to develop.

10.4.1.4 Cases with mutation of PrP gene recorded: association between age at onset frequency distribution, and mean and modal age at onset.

Table 10.4.1.4 (i) gives a summary of the a/o f/d analyses.

Table 10.4.1.4(i) TSE with known PrP gene mutation; summary of a/o f/d analyse.

Dataset	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
Codon 200	69	No	No	L	L
Codon 198	19	Yes	Yes	NS	NS
Codon 178	73	No	No	L	L
Codon 102	29	Yes	No	NS	L
Codon 117	25	Yes	Yes	NS	NS
Inserts	29	Yes	Yes	NS	NS

Each of the six main PrP mutations described have a different mean a/o ranging from around 33 to 55 years. Table 10.4.1.4(ii) describes the f/d for each dataset (divided into 3 arbitrary groups, 'late', 'intermediate' and 'early' mean a/o), and gives the modal and mean a/o values.

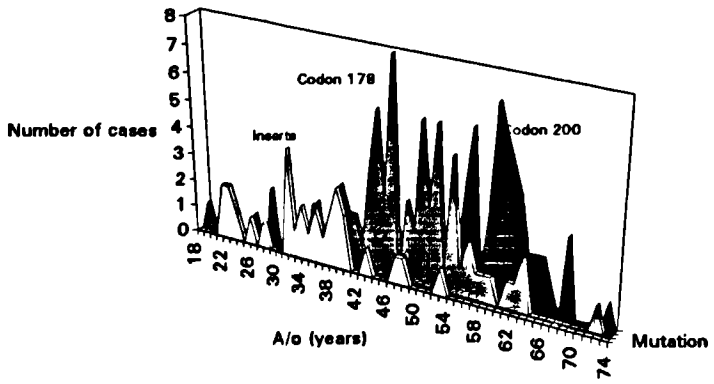
Table 10.4.1.4(ii) TSE with known PrP gene mutation; modal and mean a/o (in years), and f/d descriptions.

Dataset	Description	Mod/al region a/o	Mean a/o
Codon 200	Left skew too elongated for normality	59	55.1
Codon 198	Compatible with normality and log-normality	40-60	51.6
Codon 178	Left skew too elongated for normality	46	47.2
Codon 102	Compatible with normality	39-52	45.8
Codon 117	Compatible with normality and log-normality	23-42	39.1
Inserts	Compatible with normality and log-normality	32	34.7

Even using the largest dataset for each mutation, only two have sample sizes over 30, and of those smaller than this, three are compatible with both normality and log-normality, most probably a reflection of small sample size. The two largest sets, one 'late' and one 'intermediate' mean a/o both have a left skew too elongated for normality, and the remaining set, of 'intermediate' a/o, is compatible with normality only.

Allowing for the difficulties of interpretation with variable size datasets, both this analysis, and the non-significant reduction of the c/s with increasing modal and mean a/o (section 10.2.2.3) suggest that modal and mean a/o's are associated with f/d skewness, an hypothesis already put forward in sections 10.4.1.1 and 10.4.1.2. Figure 10.4.1.4 demonstrates this for three mutation groups with well separated modal a/o's.

Figure 10.4.1.4 TSE in humans; f/d skewness and modal a/o; a comparison of three PrP gene mutations.



Different modal regions are clear. Dataset size variation may affect mean a/o results - the a/o range observed (and thus the mean a/o) may be linked with sample size - but sample mode should not be. Theoretically, a 'late' modal a/o implies enough time to observe unusually early cases, and an 'early' modal a/o gives time for unusually late cases, resulting in left and right skews respectively, and this tendency is apparent. Mean lifespan, and competing causes of death are also relevant factors, particularly for 'late' modal a/o groups. A similar pattern exists for the remaining mutation groups (not shown), but less clearly; case numbers are much smaller.

There is some indication that mean a/o is affected by codon 129 met/val combination. If so, the extent of this effect will depend on the proportions of each combination present in the sample. If case numbers allowed, further subdivision for example by this classification, and separation of inserts by type, might enable further multi-modality to be recognised. The total set of genetic/familial cases, whilst appearing originally as a uni-modal continuum with a wide a/o range, appears to comprise many groups, each with a characteristic a/o mode and f/d. In addition, further PrP gene mutations have now been found (Prusiner, 1993).

Whether PrP gene mutation is sufficient cause for disease, or merely increases susceptibility, it seems probable that the final prion protein shape will differ, dependant upon the particular genetic mutation involved. With an extensive genetic insert, the final protein conformation may be disrupted more than with a single codon change. Assuming protein conformation is significant in disease aetiology, then the more abnormal it is, the less likely that particular protein is, perhaps, to be enzymatically degraded, resulting in accelerated accumulation of an abnormal protein and an associated earlier a/o. This suggestion, that insert mutations might be the most damaging, is supported by this data; such cases do have the earliest modal a/o.

10.4.2 Summary of the findings from the analysis of the other variables.

No differences were found by gender. For familial cases, there was no evidence of a difference in the a/o between cases where the mother was also a case, and those where the father was also a case. This may argue against maternal transmission of infectious agent (possibly concurrent with a genetic susceptibility factor). However, case numbers with known parent disease-status were small.

For NIH sporadic data, the mean a/o for experimentally transmitted cases (data not shown) is younger than that with other diagnostic methods. This may reflect bias in selecting cases for experimental transmission. Atypical cases, including younger cases, are more likely to have been selected, particularly in later years, partly as a method of confirming diagnosis. Eight iatrogenic cases were successfully transmitted.

For pituitary hormone cases, there is no association between i/p and age at first treatment; however, actual infection date is unknown.

10.4.3 Comparison with Sartwell's conclusions.

Both sporadic and familial datasets were left skewed. If a/o is taken as an estimate of i/p, these were clearly incompatible with Sartwell's hypothesis. For f/d by specific PrP gene mutation, those mutations associated with an 'early' modal a/o tended to be more right skewed than those with a later modal a/o. However, small sample sizes made discrimination between normality and log-normality (or even perhaps a more extended right skew) impossible.

For transmitted cases, datasets were small but the i/p f/d were right skewed, and for one set of pituitary injection cases at least, compatible with log-normality. For all such datasets where both the i/p and the mean a/o was known, mean a/o was considerably less than for sporadic, familial, or even individual PrP gene mutation groups.

For many categories of human TSE/prion disease, taking the a/o as an estimate of the i/p, the large left skew present means that Sartwell's hypothesis of log-normality does not apply. Even in the iatrogenic transmission dataset where i/p f/d is log-normal, the small sample size makes confirmation of this hypothesis impossible.

CHAPTER 11

CATTLE: DATA ANALYSIS AND DISCUSSION: BSE.

11 Summary.

11.1 Age at onset frequency distribution.

11.1.1 Age at onset frequency distribution by group.
11.1.2 The effect of censoring.
11.1.3 Comparison of left censoring in groups 1 and 3Y.

11.2 Association of other variables with age at onset.

11.2.1 Analysis of age at onset by maximum possible age of exposure: Group 2.
11.2.2 Analysis of age at onset by feedflag status: Group 3

11.3 Effect of Culling.

11.4 Application of AIDS methodology to these datasets: Back-calculation to overcome censoring.

11.5 Discussion.

11.5.1 Age at onset frequency distribution analysis findings.
    11.5.1.1 Frequency distribution for complete groups.
    11.5.1.2 Effect of censoring on frequency distribution.
11.5.2 Censoring: association of birth cohort with age at onset and coefficient of skewness.
11.5.3 Effect of age at infection on mean age at onset.
11.5.4 Culling.
11.5.5 Back-calculation applied to BSE data.
11.5.6 Comparison with Sartwell's conclusions.

CHAPTER 11

CATTLE: DATA ANALYSIS AND DISCUSSION

BSE.

Summary.

Large datasets extending in time over the whole of the BSE epidemic have an age at onset (a/o) frequency distribution (f/d) which is too right skewed for compatibility with log-normality.

The effect of censoring over the duration of the epidemic is marked, resulting in a decrease in mean a/o in successive birth cohorts, with a concurrent decrease in coefficient of skewness (c/s); f/d of successive cohorts changes from right skewness, through symmetry, to left skewness.

Back-calculation (as used in AIDS analysis) utilises an assumed incubation period (i/p) f/d in analysis; of the three distributions tested (Weibull, gamma, log-normal), log-normal gave the best fit to observed data.

It is suggested that, for BSE, log-normality may be a reasonable approximation to a/o f/d over much of the a/o range. However, even if a/o is a good estimate of i/p, Sartwell's hypothesis is not confirmed; low probability, long i/p (a/o) cases may make large datasets too right skewed.

11.1 Age at onset frequency distribution.

11.1.1 Age at onset frequency distribution by group.

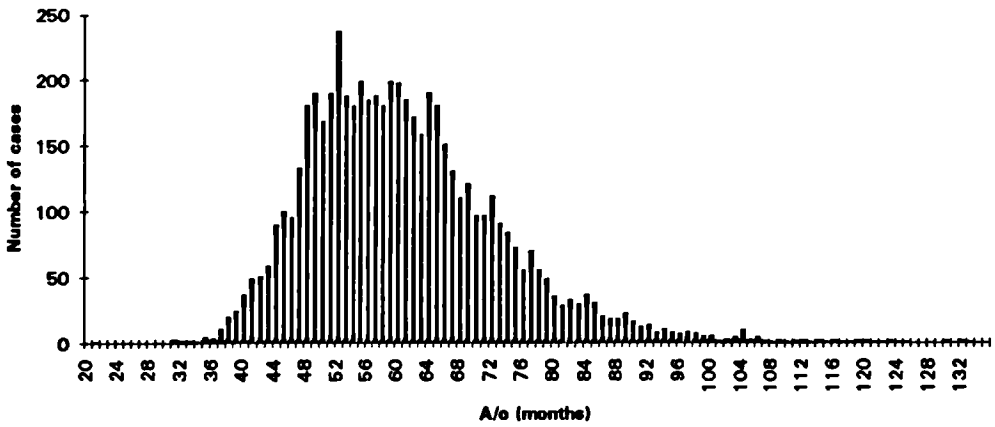
The a/o data (in months) by group (see appendix 10 for group definitions) is given in table 11.1.1(i).

Table 11.1.1(i) BSE in cattle; a/o data (in months) by group.

Group	N	Mean	SD	Range	Modal region	Median
1	488	60.9	12.0	33-113	59-60	60
2	204	61.8	11.1	38-122	55	61
3	5810	60.8	12.9	21-133	49-65	59

At just over five years (i.e. 61 months), there is no difference between the mean a/o by group (P = 0 143, non-parametric test). Group 3, the largest sample, has the widest range. Figure 11.1.1 illustrates a/o f/d for group 3; all groups were similar.

Figure 11.1.1 BSE in cattle; group 3; a/o f/d.



A right skew is apparent. Table 11.1.1(ii) gives results of a/o f/d analyses for all groups.

Table 11.1.1(ii) BSE in cattle; a/o f/d analysis by group.

Group	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
1	488	Normality	0.975	0.997	0.973	<0.001
2	204	"	0.973	0.992	1.040	<0.001
3	5801	"	0.973	0.999	1.066	<0.001
1	488	Log-	0.996	0.997	0.289	0.010
2	204	normality	0.995	0.992	0.231	0.169
3	5801	"	0.997	0.999	0.309	<0.001

None are comparable with normality; all are right skewed. The two larger groups, 1 and 3, are also too right skewed for compatibility with log-normality whereas group 2, the smallest group, is compatible with log-normality.

11.1.2 The effect of censoring.

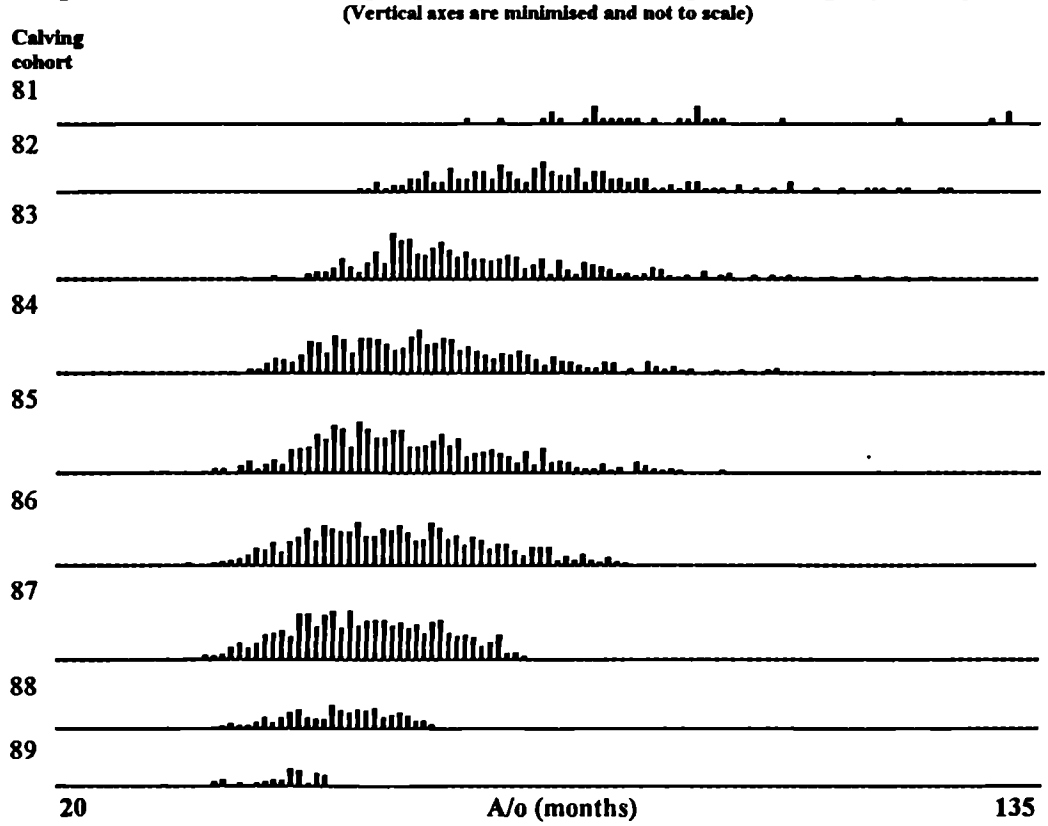
Groups 1 and 3Y were examined and gave similar results. Results for group 3Y results are shown. Table 11.1.2 (i) gives a/o data (in months) by calving cohort.

Table 11.1.2(i) BSE in cattle; group 3Y; a/o data (in months) by calving cohort.

Calving cohort	N	Mean a/o	SD	Range	Mode	Median
81	29	92.6	16.8	68-132	83	88
82	185	78.6	14.1	55-125	77	77
83	454	69.9	14.2	41-123	59	67
84	713	63.7	12.6	39-105	62	62
85	712	61.5	11.8	32-98	55	60
86	932	60.2	10.8	31-87	55	59
87	1281	56.1	8.8	34-76	52	56
88	306	52.8	6.2	31-65	52	53
89	65	46.2	4.3	32-53	47	47
90	1	21.0	-	-	-	-

Sample size varies (1 to 1281), the 87 cohort directly before the 'feedban' (on 18th July, 1988) being largest. Mean a/o reduces significantly ($P<0.001$, non-parametric test) with cohort, as do medians. The a/o appears stable for cohorts 85-86 ($P=0.106$, non-parametric test). Inclusion of either cohort 84 or 87 alters this ($P<0.001$ in both cases, non-parametric test). Medians for 84-86 are the same. Regression of a/o on calving cohort gives a reduction of 3.6 months per cohort ($r=-0.48$, $F=1384.80$, $P<0.001$). Figure 11.1.2 shows a/o histograms by calving cohort (90 excluded).

Figure 11.1.2 BSE in cattle; group 3Y: effect of censoring on a/o range by calving cohort.
(Vertical axes are minimised and not to scale)



The one case for cohort 90 is atypically early. Excluding this, left censoring is only marked for cohorts 81-82. Right censoring occurs through the whole period. Table 11.1.2. (ii) gives a/o f/d analyses by calving cohort.

Table 11.1.2(ii) BSE in cattle; group 3Y; a/o f/d analysis by calving cohort.

Calving cohort	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s	Skew
81	29	Normality	0.931*	0.964	1.178	0.008	R
82	185	"	0.969*	0.991	1.004	<0.001	R
83	454	"	0.964*	0.997	1.117	<0.001	R
84	713	"	0.980*	0.998	0.755	<0.001	R
85	712	"	0.987*	0.998	0.616	<0.001	R
86	932	"	0.994*	0.998	0.249	0.002	R
87	1281	"	0.994*	0.999	0.049	0.472	NS
88	306	"	0.992*	0.995	-0.362	0.010	L
89	65	"	0.962*	0.981	-1.000	0.002	L
81	29	Log-	0.960*	0.964	0.812	0.051	R
82	185	normality	0.990*	0.991	0.505	0.006	R
83	454	"	0.989*	0.997	0.534	<0.001	R
84	713	"	0.996*	0.998	0.289	0.002	R
85	712	"	0.998	0.998	0.125	0.171	NS
86	932	"	0.996*	0.998	-0.130	0.104	NS
87	1281	"	0.993*	0.999	-0.260	<0.001	L
88	306	"	0.983*	0.995	-0.686	<0.001	L
89	65	"	0.945*	0.981	-1.271	<0 001	L

[Note: Asterisks indicate incompatibility with normality/log-normality; significance of skew is at 0.05 level, programme will not calculate probability for N<5]

With large sample sizes, the trend from right to left skewness with successive cohorts is clear. None are compatible with normality. Only cohort 85, with demonstrated right censoring, is compatible with log-normality. Cohorts 88-89 have left skews.

Regression of the c/s on calving cohort gives a reduction of 0.26 units per cohort (r=0.96, F=76.50, P<0/001). Conversely, in the presence of censoring, regression of c/s on mean a/o gives an increase of 0.05 units per additional month (r=0.86, F=19.94, P=0.003).

11.1.3 Comparison of left censoring in groups 1 and 3Y.

Comparison of the histograms (not shown) suggests that left censoring may be slightly less in group 3Y than in group 1. Differences in the amount of left censoring in a cohort would be expected to alter minimum and mean a/o. For group 1, for cohorts 81-82 and 89, small sample size makes comparisons unrewarding. The remaining cohorts are compared (table 11.1.3).

Table 11.1.3 BSE in cattle; groups 1 and 3Y; comparison of minimum and mean a/o's (in months) per cohort.

Group	Measurement	83	84	85	86	87	88
1	Minimum a/o	51	43	40	33	37	36
3Y	"	41	39	32	31	34	31
1-3Y	Difference	10	4	8	2	3	5
1	Mean a/o	70.3	65.0	63.6	61.7	57.9	54.3
3Y	"	69.9	63.7	61.5	60.2	56.1	52 8
1-3Y	Difference	0.4	1.3	2.1	1.5	1.8	1.5
	P for difference	0.841	0.941	0.192	0.162	0 008	0.123

For all cohorts, the mean a/o is less for group 3Y than 1, but only statistically significantly so for 87, the largest cohort. Minimum a/o is again less for all 3Y cohorts, more markedly so in the early cohorts. Data are consistent with the hypothesis that there is less left censoring in group 3Y.

11.2 Association of other variables with age at onset.

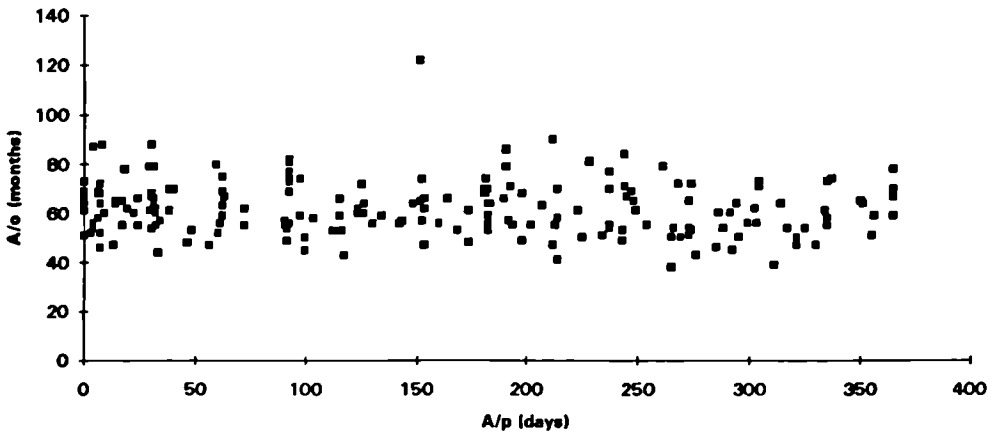
11.2.1 Analysis of age at onset by maximum possible age of exposure: Group 2.

Maximum possible age of exposure is estimated as age at purchase (a/p). Details (in days):-

Mean a/p:	**152**
SD:	**111**
Range:	**0-365**
Mode:	**0**
Median:	**151**

A minimum a/p of zero implies transfer to the new farm on the day of birth; this is also the modal a/p (9 cases). Figure 11.2.1 plots a/p (in days) against a/o (in months).

Figure 11.2.1 BSE in cattle; a/o (months) by a/p (days).



No marked association is apparent. Regression analysis gives a reduction of 0.01 months in the a/o per additional day of possible exposure (r=0.011, F=2.29, P=0.132). The mean a/p does not vary with cohort (P=0.225, non-parametric test), and no correlation is apparent (r=-0.03), therefore confounding by calving cohort is unlikely.

11.2.2 Analysis of age at onset by feedflag status: group 3

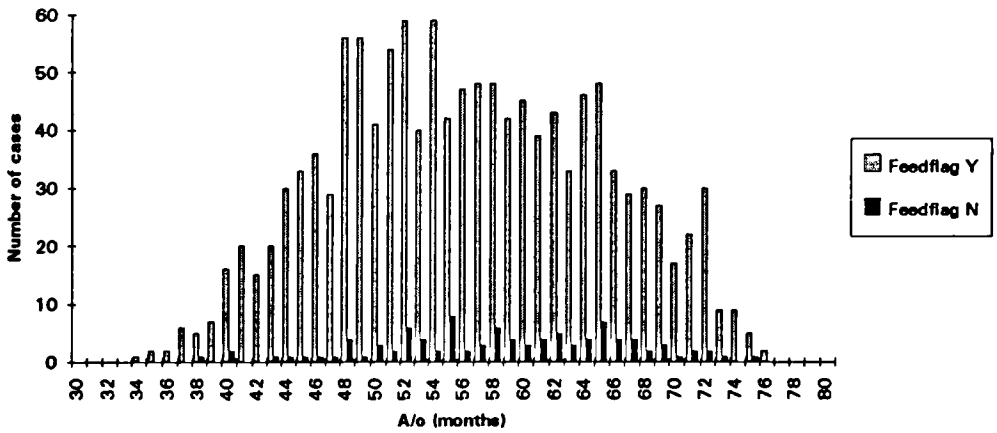
The large 87 cohort is used to examine the a/o by feedflag status (see appendix 10 for feedflag definition). Table 11.2.2(i) gives a/o data (in months) by feedflag status.

Table 11.2.2(i). BSE in cattle; group 3/87; a/o data (in months) by feedflag status.

Feedflag	N	Mean a/o	SD	Range	Mode	Median
Y	1281	56.1	8.8	34-76	52	56
N	99	58.4	8.0	38-75	55	59

Although the mean a/o by feedflag is different (P=0.006, non-parametric test), a/o ranges are broadly similar. Figure 11.2.2 compares a/o f/d by feedflag status.

Figure 11.2.2 BSE in cattle: Group 3/87: a/o f/d by feedflag status



There is an impression that, for feedflag N, f/d may be slightly left skewed. Feedflag Y appears roughly symmetrical. Table 11.2.2(ii) gives f/d analysis results by feedflag status.

Table 11.2.2(ii). BSE in cattle; group 3/87; a/o f/d by feedflag status.

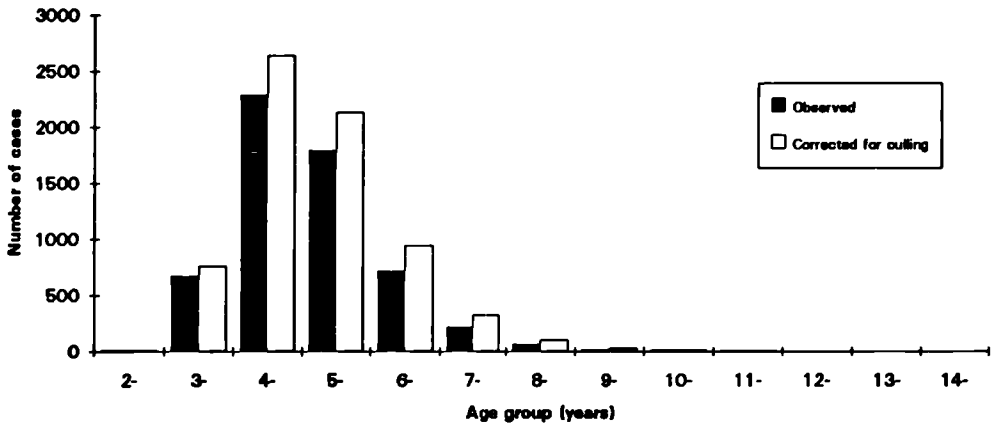
Feed -flag	N	F/d under analysis	CC	C/v at P= 0.05	C/s	P for c/s
Y	1281	Normality	0.994	0.998	0.049	0.472
N	99	"	0.995	0.987	-0.278	0.236
Y	1281	Log-	0.993	0.998	-0.260	0.002
N	99	normality	0.986	0.987	-0.622	0.012

Feedflag Y f/d, although symmetrical, is incompatible with both normality and log-normality. Rejection of normality is therefore on other grounds, for example kurtosis; large sample size makes compatibility more difficult to achieve. For feedflag N, a much smaller sample, data is compatible with normality; left skewness is not marked.

11.3 Effect of Culling.

Figure 11.3 shows observed the a/o f/d (in years) for group 3 as observed, compared with the same group adjusted for culling (method, appendix 24). Results for group 1(not shown) were similar.

Figure 11.3 BSE in cattle: group 3: effect of culling correction on a/o f/d



The culling correction has increased the sample size by 20% to 6962 cases. The 'unculled' mean a/o is 4.7 years (SD 1.17), as against 4.6 years (SD 1.11), for the observed, culled sample. The a/o f/d appears similarly right skewed for both datasets, and a/o f/d analysis conclusions are similar (see section 11.1.1). Both have a right skew which persists even in log transformed data. No values for cull adjustment are calculable for ages 12 and 13.

11.4 Application of AIDS methodology to these datasets: Back-calculation to overcome censoring.

Back-calculation (kindly undertaken by R Sayers; see appendix 25) was performed for three different a/o f/d assumptions, Weibull, gamma and log-normal, each for groups 1, 2, 3 and 3Y. All groups gave similar results. Table 11.4 gives results for group 3, both unadjusted and adjusted (in brackets) for culling (excluding the 1 case < 2 years old); N=5809.

Table 11.4 BSE in cattle; group 3; estimates of a/o data (in months) from back-calculation: unadjusted and (adjusted) for culling.

Distribution	Percentiles			
	1%	50%	95%	99%
Gamma	39 (40)	62 (68)	91 (104)	107 (124)
Weibull	37 (38)	63 (68)	89 (97)	100 (109)
Log-normal	41 (41)	65 (68)	99 (113)	121 (143)
	Mean	Mode	Max log likelihood	
Gamma	64.0 (69.9)	58 (62)	-3405.32 (-3419.53)	
Weibull	63.9 (69.0)	61 (66)	-3585.76 (-3482.38)	
Log-normal	67.5 (72.3)	60 (61)	-3378.39* (-3390.47)	

The maximum log-likelihood is the largest for the log-normal distribution. For this distribution, the mean a/o is around 6 years, an increase on all previous estimates, and the 99% value is around 10 (unadjusted) to 12 (adjusted) years.

11.5 Discussion.

11.5.1 Age at onset frequency distribution analysis findings.

11.5.1.1 Frequency distribution for complete groups.

Table 11.5.1.1(i) gives the a/o f/d analyses summaries for the three complete groups.

Table 11.5.1.1(i) BSE in cattle; summary of a/o f/d analyses.

Group	N	F/d compatible with:-		Skew for:-	
		Normality	Log-normality	Raw data	Transformed data
1	488	No	No	R	R
2	204	"	Yes	R	NS
3	5801	"	No	R	R

Table 11.5.1.1(ii) gives the f/d description for the 3 groups.

Table 11.5.1.1(ii). BSE in cattle; a/o f/d descriptions.

Group	Description
1	Right skew too elongated even for log-normality
2	Compatible with log-normality
3	Right skew too elongated even for log-normality

The two larger samples are too right skewed to be compatible with log-normality; the smaller set is not. Sample size has already been implicated as important in observed skewness but, here, all samples are relatively large.

11.5.1.2 Effect of censoring on frequency distribution.

When examined by birth cohort, groups 1 (not shown) and 3Y both demonstrate a trend for the a/o f/d skewness to move gradually from right to left over time. Table 11.5.1.2 describes f/d skewness, by cohort, for the larger set 3Y.

Table 11.5.1.2 BSE in cattle; group 3Y; f/d skewness description by cohort.

Cohort	Description
81	Right skew too elongated even for log-normality
82	"
83	"
84	"
85	Skew compatible with log-normality
86	"
87	Skew compatible with normality
88	Left skew too elongated for normality
89	"

Although for the cohorts 85-86 the a/o skewness is compatible with log-normality, for the remaining 7 of the 9 cohorts it is not.

When subdivided by feedflag status, (an indicator of probable exposure in the first year of life), both 'Y' and 'N' subsets from group 3/cohort 87 had skews which were compatible with normality, as did the complete 87 cohort. However, the smaller group ('N'; N = 99) was compatible with normality, whilst the larger group ('Y'; N=1281) was not. This may reflect the discriminant ability of the tests, when applied to different sample sizes.

11.5.2 Censoring: association of birth cohort with age at onset and coefficient of skewness.

During an epidemic of a disease with a long i/p, if the 'true' i/p f/d is not changing over time, then censoring results in the reduction of mean i/p in subsequent infection cohorts (Peterman, 1987). Clearly, if the 'true' i/p is changing with time, the pattern observed will become even more complex, however censoring will still be present, and constitute an important component of any observed i/p change over time. In addition, if a/o is used to estimate i/p, then if infection is not close to birth, apparent i/p will be affected.

Effective exposure to the BSE agent is thought to have commenced in the winter of 1981/82 (i.e. the 81 birth cohort) (Wilesmith et al, 1988). Therefore, in order to give at least the possibility of exposure from birth (and thus increase the probability of a/o as a valid estimate of i/p) this is the first cohort included in the in the BSE groups analysed here.

It is possible, for BSE, that the 'true' i/p changed over time, perhaps due to early low doses, or post-passage adaptation of agent in cattle after recycling. Any major first-passage change in 'true' i/p f/d due to these effects might be expected to cause a sudden, rather than a gradual change in the pattern of a/o f/d by cohort observed. However, the gradual pattern of a/o change observed fits very closely to the theoretical pattern which would be caused by censoring, and in particular right censoring. Although it is therefore unproven that censoring is the only influence on the observed a/o, it seems reasonable to conclude that it is a major influence, which even by itself renders the observed a/o f/d unreliable.

Using the a/o to estimate the i/p, considered valid since early exposure is likely, both groups 1 (not shown) and 3Y demonstrate an average reduction of 3.6 months in mean a/o per birth cohort through the epidemic.

If 'true' i/p is stable, once the theoretical maximum i/p for any infected cohort has passed, assuming good case ascertainment, right censoring will be eliminated for that cohort. If, due to good case ascertainment, left censoring is also eliminated and adequate case numbers exist, then observed i/p mean and range should reflect the underlying theoretically possible values. For subsequent cohorts where maximum i/p has also been passed, the observed mean, maximum and minimum values should stabilise. Right censoring will again be apparent in later cohorts where this maximum has not yet occurred. The modal i/p should be more stable than the mean, and be affected by censoring only when outside the observed i/p range. A gradual reduction of 'true' i/p would cause modal instability, whilst a sudden, sizeable reduction in 'true' i/p should, eventually be apparent as a second region of mean and modal a/o stability, given adequate later cohort sizes.

For both groups, the a/o for cohorts 85-86 appears relatively stable, in the region of 60-63 months (and have a skewness compatible with log-normality). Due to the rather flattened f/d for most cohorts (possibly, partly, a reflection of spread of infection age through the first year) a modal region may be more appropriate than a specific point. The mode for group 1 is relatively stable from 83 onwards, at around 60 months whilst for group 3Y it is less stable; it is possible that this is an indication of a 'true' i/p change superimposed upon the censoring effect.

The calculated c/s reduces with successive cohorts, as skewness changes from right to left; therefore as mean a/o reduces, so does c/s. Thus, in the presence of epidemic censoring, and for analysis by cohort, association between mean a/o and c/s is (artifactually) reversed from that suggested by other data, where a younger mean a/o is generally associated with larger c/s.

Group 3, the 'pre-order' herds, might be expected to suffer less left censoring than other herds. The data is not inconsistent with this, but impossible to confirm due to the small size of comparison groups.

11.5.3 Effect of age at infection on mean age at onset.

Using the a/p to estimate the maximum age of exposure (group 2), increasing a/p resulted in a (non-significant) reduction in a/o of 0.01 months per day (equivalent to 3.7 months per year). However, if first exposure is generally soon after birth, and if infection by the agent generally occurs at the first exposure, maximum age of exposure may be unimportant.

The increase in the mean a/o for animals recorded as not fed suspect foods in the first year of life (group 3N) is statistically significant, at 2.3 months. Even allowing for a potential age at infection effect of i/p reduction of 3.7 months per year (see section 11.2.1), a greater increase might be anticipated. Possible explanations include recording errors, incorrect assumptions about infection age, poor infection efficiency at first exposure (group 3Y), access to concentrates intended for other age groups, general farm and pasture contamination and lateral transfer, resulting in unrecorded earlier exposure and infection. In addition, i/p may be dependant upon additional unknown biological factors resulting in a tendency for overt disease to develop at a particular modal age (assuming infection had occurred prior to that age).

11.5.4 Culling.

Cattle are culled for many reasons, including ill health and poor productivity. Culling patterns may be different in BSE-negative herds, although earlier investigations into other diseases found a similar pattern (J Wilesmith, personal communication). In addition, herds with high BSE attack rates may suffer heavy culling at particular ages. The culling pattern derived will, therefore, be only approximate.

Although culling reduces cases of overt BSE, the modal and mean a/o and a/o f/d appear relatively unaffected. However, no adjustment can be applied for ages when no case was observed. Therefore the maximum a/o for an unculled sample can only be estimated by extrapolation, giving a right tail extension and increased skew. This would only be observed if, due to a different culling pattern, sufficient cattle lived longer. Since in practice, culling patterns are thought to have remained relatively stable over time, the observed f/d is more likely to be useful

in predicting overt disease. Culling adjustment will only affect the observed a/o f/d if dairy herd management alters to the extent that the culling pattern changes on many farms.

This culling adjustment gives an estimate of the numbers of adult dairy cows culled whilst incubating disease; here, approximately 17% of the total potential cases. Some may be slaughtered for food.

11.5.5 Back-calculation applied to BSE data.

Under the assumptions given (see appendix 25), log-normal distribution fitted the data better than either Weibull or gamma distributions. Using this distribution, comparison of group 3 data, unadjusted for culling, gives a mean a/o of 60.8 months for observed data (with or without the atypical 21 month case), and 67.5 months by the back-calculation method (which excluded that case). The minimum observed a/o (excluding the 21 month case) of 31 months is less than the 41 months for back-calculation 1st percentile, whilst the maximum observed a/o (133 months) is greater than the 121 months for the back-calculation 99th percentile. The mode, by back-calculation estimated as 60 months, is within the modal region of 49-65 for observed data.

If log-normality is an appropriate distribution to use, then this method gives no evidence for censoring of the a/o extremities in this dataset, although increase in the mean a/o suggests a shortfall in observed numbers with older a/o's, possibly due to right censoring in recent cohorts. Alternatively, this discrepancy may arise because log-normality is not the most appropriate distribution, and/or that the working assumptions were incorrect

11.5.6 Comparison with Sartwell's conclusions.

There is some indication, both from the 85/86 cohorts of group 3Y (over most of their range, probably the least distorted by censoring) and from the maximum log-likelihood results for distributions used in back-calculation, that log-normality may, for BSE, be a reasonably close approximation to a/o f/d over much of the a/o range.

Nevertheless, even if the a/o is a good estimate of the i/p, these results do not confirm Sartwell's hypothesis. The a/o f/d's for the largest two complete groups are too right skewed for log-normality, although the smallest one is compatible. Previous results indicate an association between sample size and skewness, most probably due to low probability, long i/p (a/o) cases.

CHAPTER 12

ASSOCIATION BETWEEN SAMPLE SIZE, STANDARD DEVIATION, AND COEFFICIENT OF SKEWNESS FOR A LOG-NORMAL FREQUENCY DISTRIBUTION.

12.1 Theoretical values for coefficient of skewness for a log-normal frequency distribution.

12.2 Observed values from random number generation.

12.3 Frequency distribution of observed coefficient of skewness values.

12.4 Discussion.

CHAPTER 12

ASSOCIATION BETWEEN SAMPLE SIZE, STANDARD DEVIATION, AND COEFFICIENT OF SKEWNESS FOR A LOG-NORMAL FREQUENCY DISTRIBUTION.

12.1 Theoretical values for coefficient of skewness for a log-normal frequency distribution.

Theoretical c/s were calculated for a range of SD values (method; appendix 26). Examples of results are as follows:-

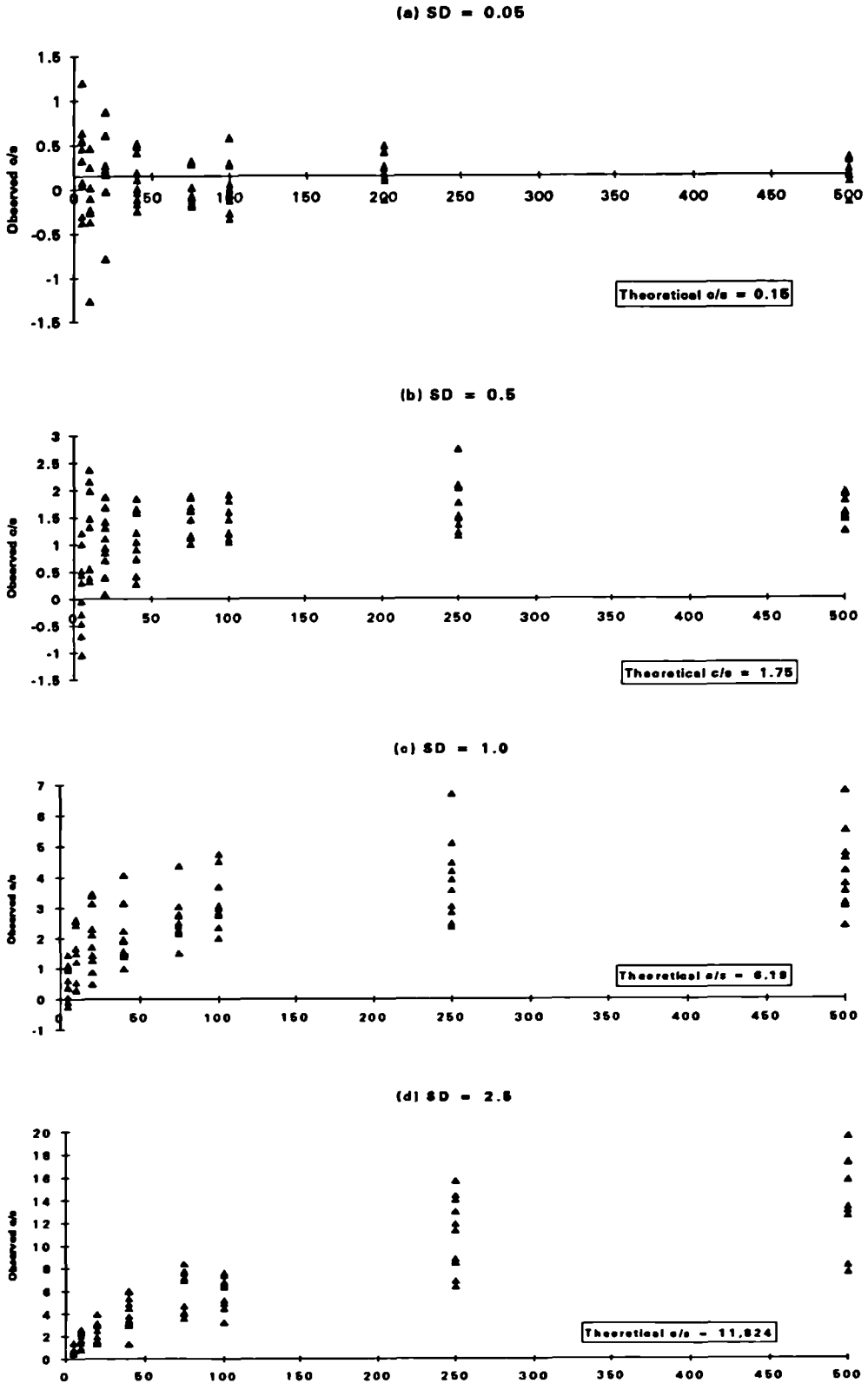
SD	C/s	SD	C/s
0.005	0.015	0.05	0.15
0.010	0.030	0.10	0.30
0.015	0.045	0.50	1.75
0.020	0.060	1.00	6.19
0.025	0.075	1.50	33.5
0.030	0.090	2.00	414.4
0.040	0.120	2.50	11,824

Results shown suggest that as SD decreases, c/s also decreases; intuitively this makes sense, since an SD of zero would mean that all values were identical, and no skew would exist. Further analyses will involve only SD values of 0.05 and larger.

12.2 Observed values from random number generation.

Initial investigation of the association between sample size and c/s when the underlying distribution is log-normal was undertaken by random number generation (method: appendix 27), for many combinations of SD and sample size, with 10 random-generation repeats for each. Figure 12.2(a)-(d) gives examples of results; each graph plots, for a particular SD, and sample sizes from 5 to 500, the 10 individual c/s values per combination .

Figure 12.2(a)-(d) Association between sample size and c/s for different SD values: log-normal distribution.



For SD = 0.05, for sample sizes below 50, c/s is unstable, with values both above and below the theoretical c/s, after which it converges around theoretical c/s value.

For larger SD values it is apparent that as sample size increases, observed c/s tends to increase until the theoretical c/s value is approached; however, as SD increases, so does the sample size needed in order to attain theoretical c/s.

Table 12.2(i) gives mean c/s (and SD of the c/s values) values for the ten observations, for sample sizes 5, 40, 100, and 500 for each SD of underlying distribution used to generate the random numbers.

Table 12.2(i) Comparison of observed and theoretical c/s by sample size and SD; 10 observations per combination

	Mean (and SD) for 10 observed c/s values for:-				Theoretical
SD	N =5	N=40	N=100	N=500	c/s
0.05	0.31 (0.47)	0.11 (0.28)	0.04 (0.27)	0.16 (0.14)	0.15
0.5	0.10 (0.73)	1.15 (0.57)	1.43 (0.32)	1.66 (0.24)	1.75
1.0	0.58 (0.56)	2.03 (0.93)	3.16 (0.89)	4.18 (1.30)	6.19
2.5	1.02 (0.40)	4.13 (1.49)	5.92 (1.39)	13.24 (4.32)	11,824

Increase in mean c/s with increasing sample size confirms graphical interpretation. Even for an underlying SD of only 0.5, mean observed c/s with sample size of 500 does not quite match theoretical c/s (95%), and for SD of 2.5, it is only 0.1%

For further analysis, numbers of repeats per SD/sample size combination was increased from 10 to 5000. Table 12.2(ii) gives examples of results.

Table 12.2(ii) Comparison of observed and theoretical c/s by sample size and SD; 5000 observations per combination.

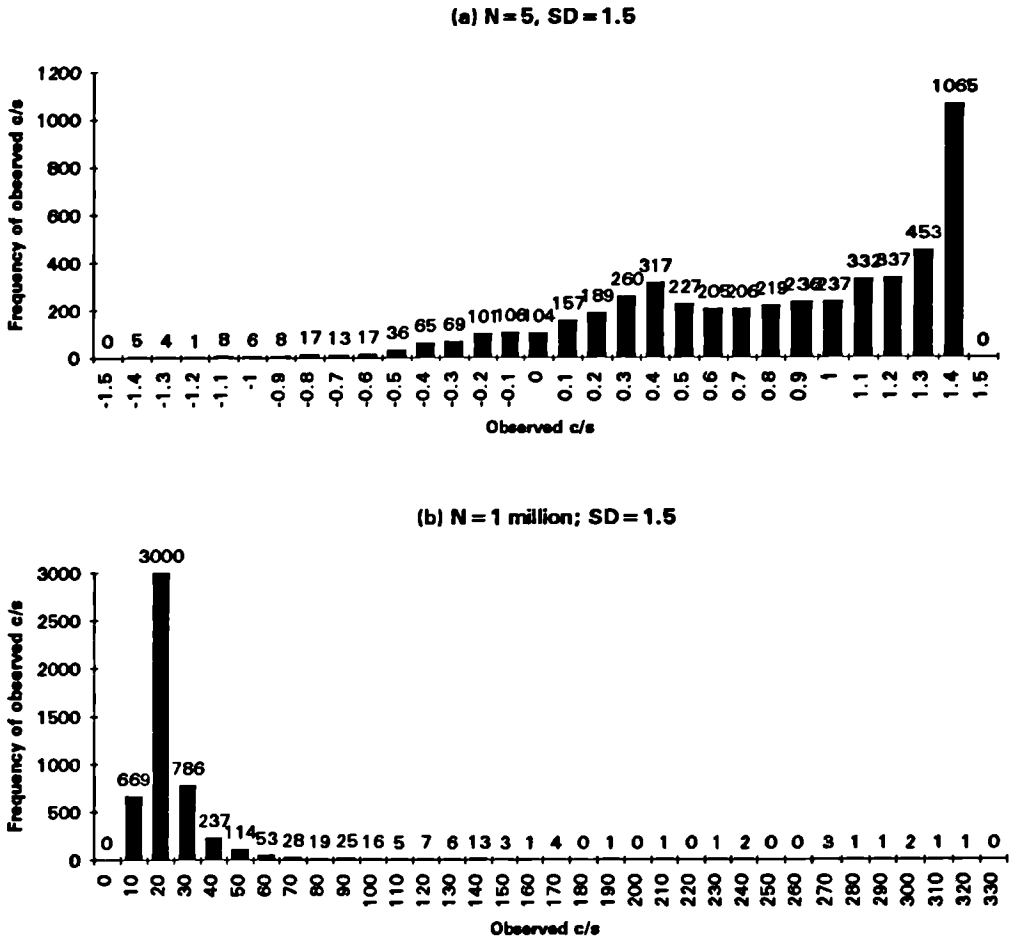
	Mean (and SD) for 5000 observed c/s values for:-			Theoretical c/s
SD	N=5	N=500	N=1 million	
0.05	0.04 (0.62)	0.15 (0.11)	0.15 (<0.00)	0.15
1.5	0.84 (0.57)	7.96 (3.49)	29.77 (20.20)	33.5

With 5000 random-generations per set of conditions, SD of 0.05 has a similar pattern to larger SD's. Even with a sample size of 1 million, with a large SD of 1.5, observed mean c/s only reaches 89% of theoretical.

12.3 Frequency distribution of observed coefficient of skewness values.

Figure 12.3 (a)-(b) shows c/s f/d for 5000 random-generations for SD=1.5, for sample sizes of 5 and 1 million.

Figure 12.3(a)-(b) Coefficient of skewness frequency distribution for 5000 random-generations; SD of theoretical distribution = 1.5.



Range and c/s distribution both vary greatly with sample size, even though samples were randomly generated from the same theoretical distribution.

C/s for the distribution of the individual c/s's are as follows:-

N= 5: C/s for c/s f/d = -0.736 (P<0.001)
N=1 million: C/s for c/s f/d = 7.009 (P<0.001)

C/s f/d for N=1million is also incompatible with both normality and log-normality (correlation coefficients: 0.665 and 0.901 respectively; critical value 0.999), and log-transformed data is still markedly right skewed (c/s = 2.257, P<0.001).

12.4 Discussion.

Results from previous chapters indicate that sample size and f/d c/s (and hence compatibility with any particular distribution) may be associated.

Results presented in this chapter demonstrate that for a theoretical log-normal distribution, when observed sample size is small, mean observed c/s tends to be much less than theoretical c/s; as sample size increases, mean observed c/s tends to increase towards theoretical c/s. Theoretical c/s increases with increasing SD of the theoretical distribution from which the random numbers are drawn. The sample size at which mean observed c/s reaches theoretical c/s depends upon the SD of this theoretical distribution; with a small SD (e.g. 0.05) , theoretical c/s is reached with a smaller sample size than with a large SD (e.g. 1.5).

There is also a dependency upon the numbers of separate sample observations used to calculate the mean c/s. The distribution of the c/s values, both for N=5 and N=1 million with SD=1.5 both show a wide range of values; although the mean observed c/s for N=1million is actually reasonably close to the theoretical value, at 89%, range of the 5000 separate c/s values is from just over 10 to over 320 units; 69 (1.4%) c/s estimations are over 100, and only 76% are within the range 20-<40. A singe observation even of a sample this large could well give a highly misleading result.

Although this analysis applies to log-normal distributions, and the distributions from which samples in earlier chapters are taken are unknown, it is apparent that the calculated c/s of a small observed sample cannot be equated to that of the theoretical distribution from which it is generated; and since the SD of the underlying distribution is unknown, minimum sample size required cannot reliably be estimated. In addition, one observation of any sample size is prone to wide errors.

CHAPTER 13

OVERVIEW OF MAJOR FINDINGS AND CONCLUSIONS.

13.1 Evaluation of Sartwell's hypothesis as applied to TSE/prion disease.

13.1.1 Datasets with incubation period measurements.

13.1.2 Datasets with age at onset measurements.

13.2 Association of modal age at onset with incubation period and age at onset frequency distribution and frequency distribution coefficient of skewness.

CHAPTER 13

OVERVIEW OF MAJOR FINDINGS AND CONCLUSIONS.

13.1 Evaluation of Sartwell's hypothesis as applied to TSE/prion disease.

The main aim of this study was to evaluate Sartwell's hypothesis in relation to TSE/prion disease, with the underlying objective of validating or otherwise the incorporation of a log-normal i/p (or a/o) term into BSE epidemic modelling.

The major finding from the project is that Sartwell's hypothesis is not confirmed in this group of diseases, for a number of reasons which will now be summarised. Datasets analysed comprise those where i/p is known, and those where a/o is used as a 'proxy' i/p. They will be reviewed separately.

13.1.1 Datasets with incubation period measurements.

Of 43 datasets tested for compatibility with log-normality, and summarised in earlier discussion sections, 14 (33%) had a right skew too extensive for compatibility with log-normality; 21 (49%) mostly small datasets were compatible with both normality and log-normality; 2 (5%) had left skews; 1 (2%) was compatible with normality only; 4 (9%) were incompatible with log-normality for other reasons, and only one (2%) was compatible with log-normality only.

However, there are many problems in the interpretation of these results. The first is the variation in sample size, from 4 to 447. Appendix 14 indicates the loss of discriminatory power, with small sample sizes, for the tests used to evaluate compatibility with normality and log-normality, and chapter 12 shows the association of c/s with sample size. Both of these effects would tend to mask an extreme right skew, the former by loss of discrimination, the latter by usually underestimating the c/s. The 14 datasets with a right skew too extensive for log-normality are therefore the most persuasive evidence of non-confirmation of Sartwell's hypothesis as the underlying theoretical f/d in TSE prion disease.

The second major problem is that of variability in the cases within each dataset, most of which were experimental results. Taking the 'full experimental set' datasets for each species generally means an amalgamation of results with differences in age, environment and genetic constitution for both donor and host, different infection routes, infecting materials, doses, agent 'strains' and passage histories, plus many other variables. However, Sartwell's hypothesis is based on observation of a 'mixed' situation, even if not quite so extreme. In his datasets, each case was

genetically different, each living in their own particular way and environment. Age, health, dosage, and many other factors for each case were also different. Armenian (1995) stresses the robustness of this model. The 'full' datasets for goats, several primates, BSE in mice, and CJD in guinea-pigs are very right skewed suggesting that, in a situation of great variability, this is not uncommon.

In addition, there is the problem of which variables to use to categorise experimental data. Information varies with data source, splitting reduces the sample size (e.g. the 4 case dataset), and many variables are by their nature, unknowable. In general, in those examples where subdivision into categories by known variables has been performed, the main effect is to reduce right skewness, and to reduce the tests discriminatory powers, without markedly changing the overall appearance of the f/d histogram. This suggests that observed extreme right skewness is associated with dataset size rather than exclusively with the mixed nature of the full set.

The 'full' sheep experimental set suggests that sample size is not always the skewness-determinant however, and the effect of a reduction of variables may be dependant on the proportions of the different categories present. One particularly marked exception is that of the highly inbred strains of mice inoculated with the highly stabilised and characterised 'strains' of scrapie agent. Here, reduction of variables to the practical minimum produces sets of cases which appear to be more clustered, with a narrower i/p range, and less 'tailing'. It is hypothesised that reducing variation even further (for example, mice at the same age at inoculation) would further tighten the i/p range around a particular modal i/p.

Thus, it may be that from a dataset with many variables (such as might be encountered in the natural situation for a TSE), and with an f/d with a right skew too extensive for log-normality, stepwise reduction of the number of variables reduces the right skewness through a stage of log-normality to result in a tightly clustered i/p.

13.1.2 Datasets with age at onset measurements.

Excluding cattle, complicated by the censoring involved in the epidemic, a total of 24 datasets with a/o measurements were tested. Of these, 3 (13%) had a right skew too extensive for compatibility with log-normality; 7 (29%) mostly small datasets were compatible with both normality and log-normality; 5 (21%) had left skews; 4 (17%) were compatible with normality only; 2 (8%) were incompatible with log-normality for other reasons, and 3 (13%) were compatible with log-normality only.

Excluding cattle, sample size varied from 15 to 1334. Those datasets with extensive right skew tended to be large, and all were suspected of having a mainly or entirely infectious aetiology (although this may be misleading; see section 13.2). In addition, they may have been associated with less culling or other artificial data truncation. Culling would tend to reduce right skewness. As well as the problems of sample size and extreme data variability (which, for natural data is far less easily controlled or categorised), there is here the additional problem of the validity of using a/o as a proxy for i/p.

For those cases believed to be mainly or wholly infectious, for example natural scrapie in sheep and goats and BSE in cattle, a/o is a valid proxy if infection occurs close to the time of birth. Actual infection time is unknown and thus no other measurement available, and in the presence of exposure to an infected flock or herd or contaminated food from birth, it is at least possible. It is therefore assumed (but not proven) that a/o is a close estimate of i/p for the majority of cases.

For those cases believed to be the result, either directly or indirectly, of a congenital abnormality (here, in the form of a genetic mutation), for example familial human cases and possibly some cases in sheep, then a/o can be seen as equivalent to the i/p of infectious disease. For those cases where aetiology is as yet unknown (mainly sporadic CJD), the assumption that an as yet undiscovered congenital abnormality (possibly an unidentified genetic mutation, not necessarily of the PrP gene, predisposing to susceptibility in some way) is involved, again may render a/o equivalent to i/p. In practice, it is the only relevant information that is generally available.

The use of the a/o as a proxy for the i/p is therefore considered to be a valid procedure but, for particular individual cases, it may not always be appropriate. Whilst this is perhaps unlikely to affect modal a/o values, it may be relevant to the extremes of the f/d range. In particular very large a/o values may be the result of later infection or disease initiation; it is possible that this is at least partly responsible for extensively right skewed f/d's.

In sheep, three datasets were compatible with only log-normality, one of which at least was subject to culling. Two other datasets classed as incompatible for other reasons were very close to log-normality. Although overall, Sartwell's hypothesis cannot be confirmed for a variety of reasons as the underlying theoretical f/d in TSE/prion disease, it may be that in sheep, in the natural situation, and particularly if culling removes very old animals, log-normality is a reasonable approximation to the observed situation.

Due to the effects of censoring in an epidemic, results taken directly from the cattle data analyses may be misleading, as discussed in chapter 11. Sartwell's hypothesis has not been confirmed as the underlying theoretical i/p in TSE/prion disease, and it is not certain that a/o is an exact

representation of i/p in BSE. In addition, results of analysis of AIDS data for compatibility with Sartwells's hypothesis as an example of another infectious disease with a long i/p were inconclusive, again due to the problems of epidemic censoring.

Nevertheless, of those distributions used in applying back-calculation methods adapted from AIDS modelling to BSE data, a log-normal distribution gave the best fit to the data, although it is possible that some other uncommonly used distribution might be even more appropriate. In addition, data from sheep suggests that log-normality might be a reasonably good distribution for practical use in that species. From these two findings, and if the disease in cattle behaves similarly to that in sheep, it appears that for practical purposes the use of log-normality to describe a/o f/d in the modelling of cattle with BSE is the most appropriate of the commonly used distributions to use in the first instance, as was done (Wilesmith et al, 1988), further evaluation being dependant on the continued closeness of fit with the observed data.

13.2 Association of modal age at onset with incubation period and age at onset frequency distribution and frequency distribution coefficient of skewness.

It became apparent from these analyses that the observed disease onset f/d, whether measured as an i/p or a/o f/d, may be dependant upon the modal a/o. Several of these correlations are based on evaluation of the observed c/s, and it has already been demonstrated that this is, in isolation, an unreliable estimator of the true c/s for a log-normal f/d. Extrapolating, it seems reasonable to expect that any type of skewed distribution may suffer from similar effects, particularly if close to a log-normal distribution. From this, it might be concluded that c/s estimation is of little use. However, it is suggested that if a group of otherwise comparable datasets show a trend in c/s values, then provided that the actual c/s values are used only as a guide, some inference might reasonably be drawn from them.

That such an association existed was first suggested by primate data analysis, when the i/p was evaluated with respect to life-history measures. It was observed that, although not statistically significant, the f/d right skewness decreased with increasing modal (and mean) i/p (see sections 8 4 and 8.5.2). With BSE in mice, increasing mean age at infection (and thus increasing modal a/o) was associated with a decrease in the c/s (see sections 9.1.2.2 and 9.4.1.2), and for scrapie in mice, the association between increasing modal i/p and decreasing c/s was clearly shown (sections 9.2.6 and 9.4.2.3).

However, the association becomes very apparent with the examination of human data. Looking first at all cases without a known infectious aetiological mechanism, examination of the a/o f/d

for cases with recorded genetic mutations is discussed at length (section 10.4.1.4). Mutations with an early modal a/o (from around 30 years of age for inserts) tend to be right skewed, whereas mutations with a late modal a/o (around 60 years for codon 200 mutations) tend to be left skewed. Not surprisingly therefore, analysis of all cases with a known familial history (an amalgam of all cases with genetic mutations) tends to have a mode (at around 45-50 years), roughly mid-way between the extremes of those for individual mutation types, and an f/d compatible with normality. For sporadic CJD cases, with a late modal a/o (in the region of 65 years), f/d is left skewed.

All human datasets with a known infectious aetiology are small, with a resultant lack of discriminatory power between different distributions and a probable underestimation of the true c/s. Nevertheless, for all three sets with a recorded a/o, the modal a/o is early (from 16 to around 30 years of age) and all appear right skewed. Clearly, this early modal a/o is in turn dependant upon a prior, and therefore also early, age at infection, which is documented for iatrogenic infection and assumed for kuru cases.

Examination of the i/p for the two relevant iatrogenic datasets gives a very short mean and modal i/p particularly for the 'CNS-inoculated' group. In addition, data values for 'maximum i/p' for the pituitary hormone cases described in section 10.3.2.1 are in general a few years less than the a/o data for the comparably sized set of kuru cases, perhaps indicating similarities in pathology during incubation. For both these datasets analysis indicates a tendency to right skewness for the i/p f/d, as for the a/o f/d. Comparison with experimentally infected datasets, for example BSE in mice, suggests that if a large enough group of patients for some reason suffered iatrogenic infection at a late age (resulting co-incidentally in a late modal a/o), they would exhibit a much reduced right 'tail' to the i/p f/d, possibly, but not necessarily, associated with death from other causes, as well as a less right skewed a/o f/d.

In summary, it is suggested that the i/p f/d c/s is determined by the difference between the modal a/o and either the mean or modal age at death, and that a right skewed i/p f/d (whether more or less skewed than log-normality) will only occur when this margin is wide enough to accommodate that skew. The modal a/o will depend upon time/s of infection, plus the modal i/p. Ultimately, therefore, this margin depends also upon modal i/p. For congenital disease, i/p equates to a/o. Skewness will then depend directly on the modal a/o.

It is therefore hypothesised that for TSE/prion disease in humans, the modal a/o, whether dependant upon a congenital abnormality or prior infection, is inversely associated with the observed a/o c/s. Examination of much larger datasets, particularly for specific mutations, and for known, recorded infection amongst different age groups (unlikely to be easily accomplished), is

required to fully evaluate this hypothesis, and especially to estimate whether the degree of any association is similar for both mutation group datasets and known infectious datasets. Any difference may be indicative of differing aetiologies.

If this hypothesis is correct, it may extend to other classes of disease. Huntingdon's chorea, an hereditary disease of humans, has a mid-life modal a/o of 38.3 years (SD 11.5) and an overall a/o f/d which is very close to normality, based on an examination of 3153 cases (Ridley et al, 1992). The differences in sample size makes comparison difficult, but the f/d does appear to be much more symmetrical than the marked left skew of the sporadic CJD datasets with their mean a/o's at around 60 years.

Returning to Sartwell's hypothesis, his data was based on diseases with short i/p's, the longest being for 'serum hepatitis' with a modal value of three months, the majority measured in days. With such short mean and modal i/p's, unless an outbreak of one of the diseases which he examined occurs in a close-to-death population, such as might be found in a hospice, or during active combat in war, mean duration of time left to death from other causes will generally be so large in comparison with the i/p of the disease in question that a right skewed f/d is probably inevitable. Its compatibility with log-normality may depend on for example sample size, case ascertainment, and duration of observation, as well as power of the statistical method being used to reject this (or any other) hypothesis. This would fit with his observations regarding experimental datasets, where observation and case ascertainment is generally more complete.

And finally, with respect to human TSE/prion disease, if this hypothesis is correct, the question now is not 'why is the a/o f/d as it is?', but rather 'what determines the modal a/o?'.

APPENDIX 1

EXAMPLES OF HISTOGRAMS OF INCUBATION PERIOD FREQUENCY DISTRIBUTIONS, PLUS OTHER DETAILS, FOR DISEASES WITH SHORT INCUBATION PERIODS AS DISCUSSED IN CHAPTER 1.

Figure A1.1 Frequency distribution: Serial interval for measles as an estimation of i/p (after Stillerman & Thalheimer, 1944).

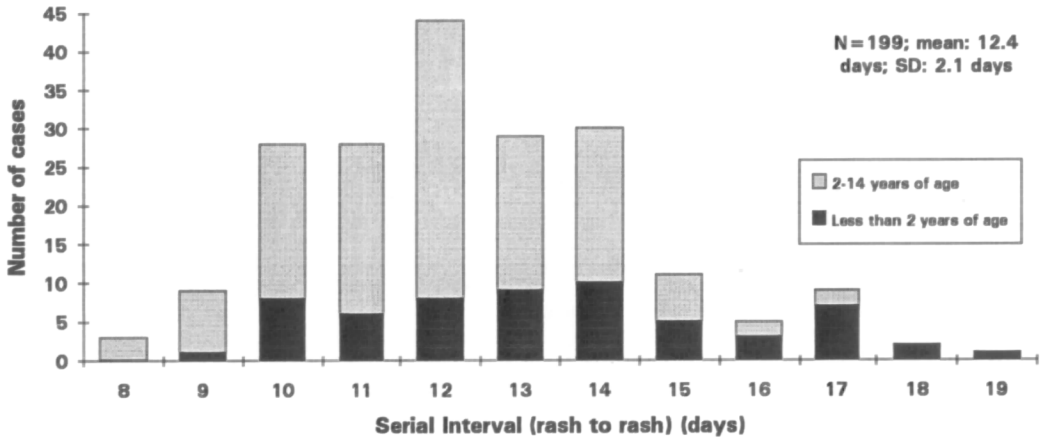


Figure A1.2 Frequency distributions for incubation periods for 5 outbreaks of 'Typhoid Fever' reviewed by Miner (1922).

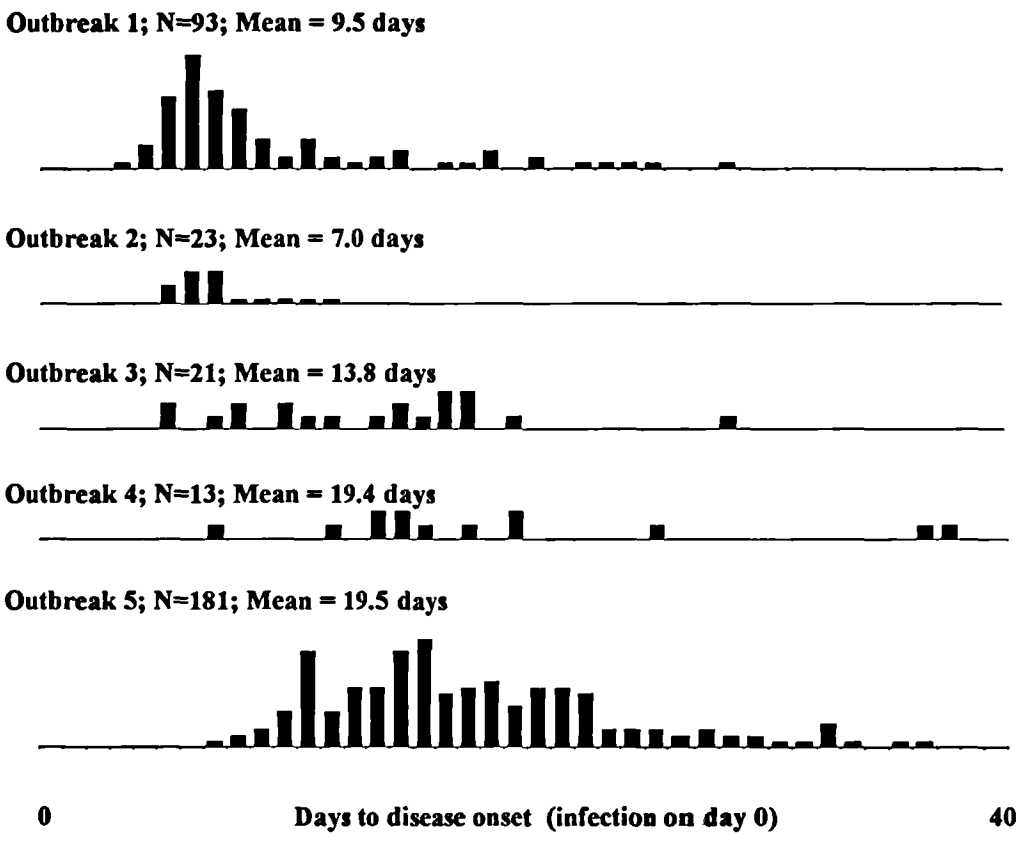
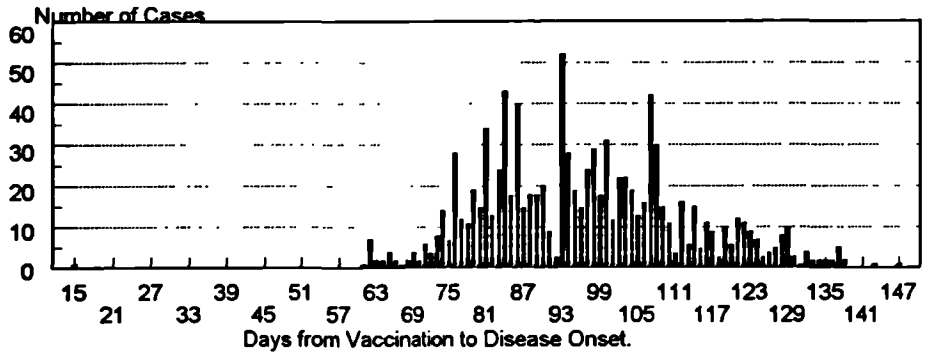


Figure A1.3 Details of i/p for 'Serum Jaundice'; Camp Polk series (after Parr, 1945)



Vaccinated at Camp Polk, Feb 27th 1942.

Remained at Camp Polk through the outbreak.

Vaccination lot: Number 369.

Mean Incubation Period: 96.4 days

Standard Deviation : 16.1 days

Mode: 93 days (52 cases)

Range: 15 to 147 days

(it was suggested that the case on day 15 may have been coincidenctal)

Number of Cases: 1004

Estimated Attack Rate: 18.4 cases per 100 soldiers vaccinated with this vaccine lot.

Figure A1.4 Frequency distributions: Reported cases of GBS by period between vaccination and onset in 1-week periods and 2-day periods (after Schonberger et al, 1979)

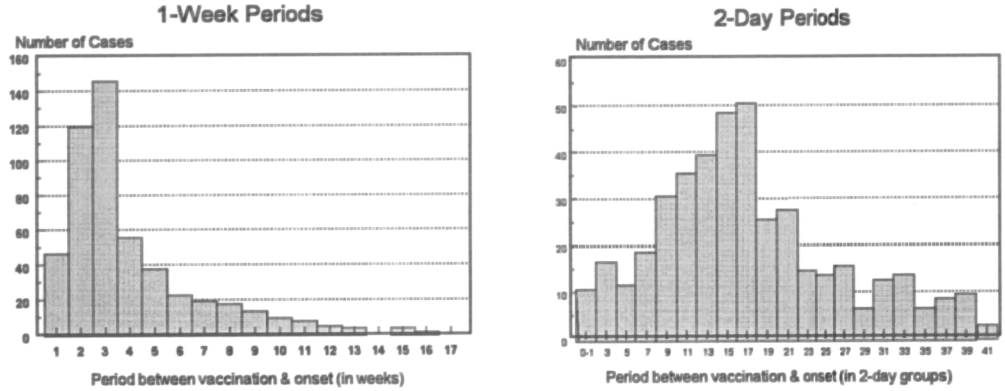
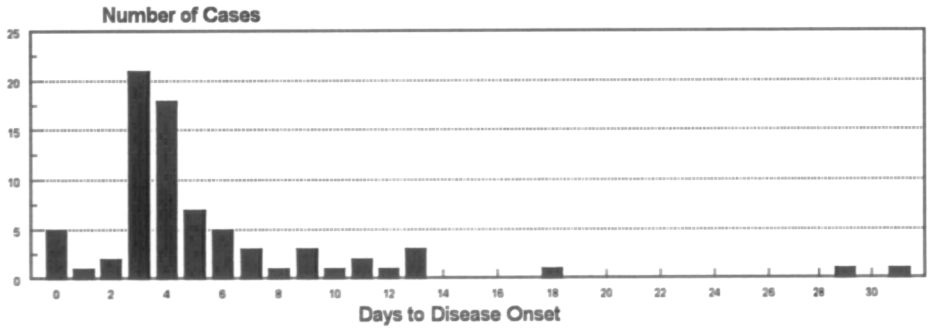


Figure A1.5 Frequency distribution: Experimental gonorrhoea using human volunteers; i/p f/d (after Mahoney et al, 1946)

Mean: 5.6 days; SD:5.3 days.



Males, age 21-45 years, free from intercurrent disease or debility
Day 0 is day of experimental infection (24 hours)

APPENDIX 2

DETAILS OF SARTWELL'S METHOD FOR THE EXAMINATION OF THE FREQUENCY DISTRIBUTION OF AN INCUBATION PERIOD.

Procedure Undertaken.

* The frequencies (number of cases with disease onset within a given time) are grouped in time intervals appropriate to the disease (hours, days or weeks).

* Cumulative frequencies, and corresponding cumulative percentages are calculated.

* The time intervals are transformed into logarithms.

* Cumulative percentages are plotted against log time on normal probability paper.

* If the points are nearly linear, a straight line is fitted by inspection.

* **"Estimated median":** the logarithm of the point at which this line intercepts the 50% frequency line (50 percentile) is read off. Its antilogarithm is the "estimated median" of the frequency distribution, in the original time units.

* **"Dispersion factor":** the logarithms corresponding to the 16 percentile and 84 percentile are read off (corresponding to +1 and -1 standard deviations). The difference between the 2 logarithms is found by subtraction, and this value divided by 2. The antilogarithm of the resultant figure is the "dispersion factor", in the original time units.

Notes:-

These 2 parameters correspond to the geometric mean, and geometric standard deviation, when calculated arithmetically.

The incubation period range given by "estimated median/dispersion factor" to "estimated median x dispersion factor" should encompass approximately 68% of observed incubation periods for that disease, for a log-normally distributed disease.

APPENDIX 3

EXAMPLES FROM SARTWELL'S TABLE OF FREQUENCY DISTRIBUTION RESULTS FOR OUTBREAKS IN WHICH HE FOUND A 'LOGARITHMIC NORMAL' DISTRIBUTION.

Disease	Number of cases	Estimated median	Dispersion factor
Streptococcal sore throat (1)	51	56 hours	1.53
Salmonellosis (2) *(S. typhimurium)*	227	2.4 days	1.47
Measles (3)	25	12.4 days	1.32
Common cold (4) (experimental)	92	2.4 days	1.50
Amoebic dysentery (5)	215	21.4 days	2.11
Induced malaria (6)	24	6.4 days	1.33

Sources of Data.

(1) Ingraham, personal communication to Sartwell.
(2) Mosher et al, 1941.
(3) Goodall, 1931.
(4) Andrewes, 1949.
(5) National Institute of Health, 1936.
(6) Greig, 1939.

APPENDIX 4

PRACTICAL DEMONSTRATION OF SARTWELL'S METHOD USING TWO SETS OF OUTBREAK DATA ALSO EXAMINED BY SARTWELL.

a) Measles Data from Stillerman & Thalheimer, 1944.

- serial interval used as best measure of incubation period
- grouped into 2 day intervals for plotting

Day	Log day	Cases	Cum. cases	Cum. %	Day	Log day	Cases	Cum. cases	Cum. %
8	.90	3	3	1.5	14	1.15	30	171	85.9
9	.95	9	12	6.0	15	1.18	11	182	91.5
10	1.00	28	40	30.2	16	1.20	5	187	94.0
11	1.04	28	68	34.2	17	1.23	9	196	98.5
12	1.08	44	112	56.3	18	1.26	2	198	98.5
13	1.11	29	141	70.9	19	1.28	1	199	100

Results from Graph (figure A4.1, next page).

	Sartwell	This Plot
Estimated median:	12.2 days	11.8 days
Dispersion factor:	1.18	1.19

Calculated results for "estimated median" and "dispersion factor" here differ slightly from Sartwell's results since this method depends upon the fitting of the line by eye, with associated observer bias. Alternative methods, such as fitting the line by regression, would remove this observer bias, but this was not done by Sartwell; neither were any 'goodness of fit' tests to see whether observed distribution differed from hypothesised distribution.

b) 'Typhoid Fever' Data from Sawyer, 1914.

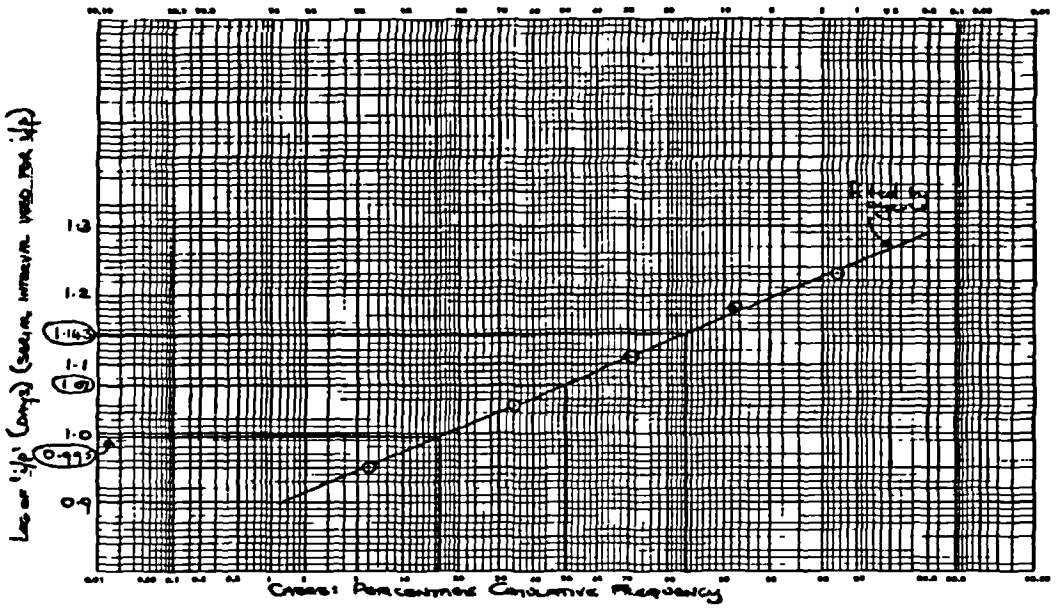
- grouped into 2 day intervals for plotting

Day	Log day	Cases	Cum. cases	Cum. %	Day	Log day	Cases	Cum. cases	Cum. %
3	.48	1	1	1.1	17	1.23	1	82	88.2
4	.60	4	5	5.9	18	1.26	1	83	89.2
5	.70	12	17	18.3	19	1.28	3	86	92.5
6	.78	19	36	38.7	20	1.30	0	86	92.5
7	.85	13	49	52.7	21	1.32	2	88	94.6
8	.90	10	59	63.4	22	1.34	0	88	94.6
9	.95	5	64	68.8	23	1.36	1	89	95.7
10	1.00	2	66	70.8	24	1.38	1	90	96.8
11	1.04	5	71	76.3	25	1.40	1	91	97.8
12	1.08	2	73	78.5	26	1.41	1	92	98.9
13	1.11	1	74	79.6	27	1.43	0	92	98.9
14	1.15	2	76	81.7	28	1.45	0	92	98.9
15	1.18	3	79	84.9	29	1.46	1	93	100
16	1.20	2	81	87.1					

Inspection of Graph (figure A4.2, next page): - points not on a straight line
- therefore not a log normal distribution
- Sartwell obtained similar result.

[Note: Cum. = cumulative]

Figure A4.1 Sartwell's method plot for measles; data from Stillerman & Thalheimer, 1944



Sartwell's method for calculations from graph:-

a) Estimated median = antilog of 1.07 = **11.75 days**

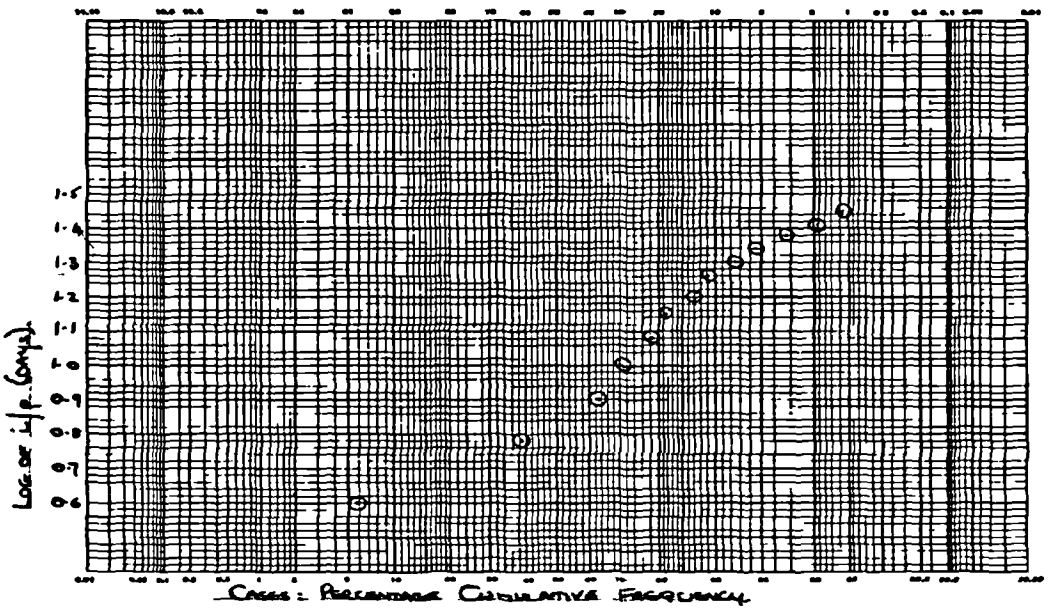
b) Dispersion factor: (log 84 percentile - log 16 percentile)/2

= (1.143 - 0.995)/2

= 0.148/2 = 0.074

Antilog 0.074 = **1.19**

Figure A4.2 Sartwell's method plot for 'typhoid fever'; data from Sawyer, 1914



APPENDIX 5

REVISED CASE DEFINITION OF AIDS (CDC, 1987).

A. With definitive diagnosis of indicator disease (with or without definitive diagnosis of HIV infection).

1) Candidiasis - oesophagus.
2) Cryptococcosis - extrapulmonary.
3) Cryptosporidiosis; diarrhoea > 1 month.
4) Cytomegalovirus infection (other than liver, spleen or lymph node) in patient > 1 month old.
5) Herpes simplex -> pneumonitis, oesophagitis, mucocutaneous ulceration for > 1 month.
6) Kaposi's sarcoma; patient < 60 years.
7) One cerebral lymphoma; patient < 60 years.
8) Lymphoid interstitial pneumonia; patient < 13 years.
9) *Mycobacterium avium-intracellulare* or *M. kansasii* (disseminated).
10) *Pneumocystis carinii* pneumonia.
11) Progressive multifocal leucoencephalopathy.
12) Toxoplasmosis of brain; patient > 1 month.

B. With definitive diagnosis of indicator disease PLUS definitive diagnosis of HIV infection.

1) Recurrent bacterial infection in a child.
2) Coccidioidomycosis - disseminated.
3) HIV encephalopathy.
4) Histoplasmosis - disseminated.
5) *Isospora* diarrhoea > 1 month.
6) Kaposi's sarcoma; any age.
7) One cerebral lymphoma; any age.
8) Non-Hodgkin's lymphoma.
9) Any disseminated mycobacterial disease (not M. tuberculosis).
10) *M. tuberculosis* - extrapulmonary.
11) Recurrent *Salmonella* septicaemia (not *S. typhi*)
12) HIV wasting syndrome.

C. With definitive diagnosis of HIV infection, and presumptive diagnosis of indicator disease.

1) Oesophageal candidiasis.
2) Cytomegalovirus retinitis with visual loss.
3) Kaposi's sarcoma.
4) Lymphoid interstitial pneumonia; patient < 13 years.
5) Disseminated mycobacterial disease (species not defined).
6) *Pneumocystis carinii* pneumonia.
7) Toxoplasmosis of brain.

APPENDIX 6

A BRIEF OUTLINE OF THE BACK-CALCULATION METHOD USED (AFTER BROOKMEYER AND GAIL, 1988) TO OVERCOME THE PROBLEM OF LEFT AND RIGHT CENSORING IN AIDS MODELLING.

A stable i/p over time is generally assumed, and the f/d assigned a function, usually from the Weibull or gamma families of curves; occasionally normal or log-normal functions are used. Changes in infection rate from time T (start of infection) are assigned a further function, frequently exponential. A function for the convolution of these two curves is then evaluated.

Thus, for an epidemic with infection commencing at time T, expected case numbers at time $(T+t^1)$, $(T+t^2)$ etc. can be calculated for any time X after T, as the product of the number infected at each time up to X and the proportion of all those infected at each time which would have disease onset at time X, estimated from the i/p f/d. Hence the total number of cases in a given period of time, from say $(T+X^1)$ to $(T+X^2)$, is described by the integral of the convoluted form, between those times.

If X^1 and X^2 are taken as the time-window of case observations during which case ascertainment was considered reliable, then numbers observed can be compared with numbers predicted. Since the number given by the equation depends upon certain parameters in the equation, those parameters are found (using maximum likelihood methods) which give best fit to the observed data, for X^1 to X^2. The complete integral can then be evaluated using those parameters, and the epidemic's total size estimated.

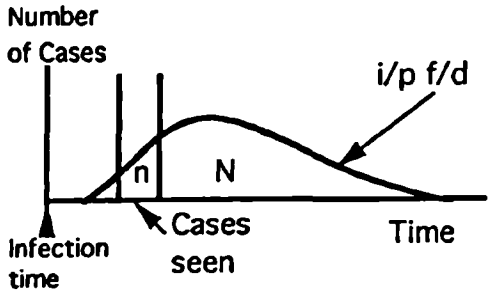
The reliability of observational data used from X^1 to X^2 depends on the dates chosen for X^1 and X^2. If X^1 is too early in the epidemic, ascertainment bias is again likely. If X^2 is too late, reporting-lag leads to under-representation of later cases. Assumptions are made about the time T (start of infection) and changes in infection rate. These will be based on incomplete information, possibly compounded by secrecy; possibly different for different risk-groups.

APPENDIX 7

DIAGRAMMATIC REPRESENTATION OF THE PRINCIPLES OF BACK-CALCULATION AND COMPARTMENTAL MODELS ILLUSTRATED FOR ONE INFECTION TIME POINT.

Back Calculation

- allows for censoring



Key:

n = incidence (cases seen) over time observed

N = total (unknown) number of cases

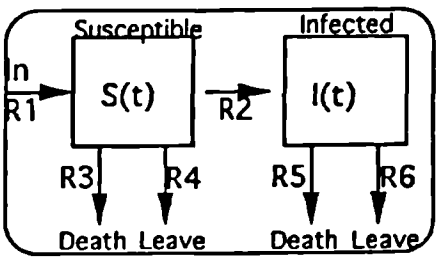
Ratio n/N = ratio of the integrals

Integral of N (i/p f/d) = 1

Therefore can calculate N

Compartmental Models

- simulates transmission pathways etc



With: rates *R1 - R6*, at time *t*

APPENDIX 8

EXAMPLES OF THE ESTIMATES OF THE INCUBATION PERIOD FOR AIDS

Reference	Group studied	Function used for i/p f/d	Resultant i/p (years)
Andreani et al, 1983	Tx in Haiti	Single case	4
Curran et al, 1984	TA-AIDS	Direct calculation	Med. 2.3 (1.3-4.8)
Curran et al, 1985b	TA-AIDS; paed.	Direct calculation	Med. 1.2 (0.3-3.8)
Lui et al, 1986	TA-AIDS; adult	Weibull i/p f/d	Mean 4.5 (90% CI; 2 6-14.2)
Rees, 1987	TA-AIDS	Normal i/p f/d	Mean 15 (SD; 5)
Rodgers et al, 1987	TA-AIDS; paed.	Direct calculation	Mean 2.0 (0.3-6.8)
"	Perinatal infection	"	Mean 1.4 (0.1-7.2)
Medley et al, 1987	TA-AIDS; >60 y	Weibull i/p f/d	Mean 5.5, med. 5.4
"	TA-AIDS; adult	"	Mean 8.2, med. 8.0
"	TA-AIDS; paed.	"	Mean 2.0, med. 1.9
Medley et al, 1988a	TA-AIDS; adult	Gamma i/p f/d	Mean 27.4 (SD 17.5)
"	TA-AIDS; paed.	"	Mean 2.4 (SD 1.5)
Medley et al, 1988b	TA-AIDS	Gamma i/p f/d	Mean 24.1, med. 20.8
"	"	Weibull i/p f/d	Mean 7.6, med. 7.3
Lifson et al, 1988	TA-AIDS; adults	Method unknown	Med. 3.8
"	TA-AIDS; paed.	"	Med. 2.8
Lui et al, 1988a	TA-AIDS; paed.	Weibull i/p f/d	Mean 2.4 (90% CI; 1.5-7.2)
Lui et al, 1988b	Gay men	Weibull i/p f/d	Mean 7.8 (90% CI; 4.2-15)
Lemp et al, 1988	Gay men	Non-parametric	Med. 10.8
Frosner, 1988	Haemophiliacs	Normal i/p f/d	Mean 9 (SD 3)
Kalbfleish & Lawless, 1988	TA-AIDS	Parametric	Med. 7.3 (95% CI; 4.6-inf)
Bacchetti & Moss, 1989	Gay men	Non-parametric	Med. 9.8 (95% CI; 9.4-inf)
Day et al, 1989	UK cases	Gamma i/p f/d	Med. > 12
"	"	Weibull i/p f/d	Med. 7-8
Longini & Clark, 1989	Mixed	Complex	Mean 9.8 (95% CI; 8.4-11.2)
Munoz et al, 1989	Gay men	Non-parametric	Med. 11
Blaxhult et al, 1990	TA-AIDS	Non-parametric	29% by 5
Giesecke et al, 1990	Gay men	Non-parametric	Mean 8
Taylor et al, 1990	Gay men	Complex	36% by 7 (95% CI; 26-47%)
Erice et al, 1991	Transplantees	Direct calculation	Mean 2.7
Hendriks et al, 1993	Gay men	Gamma i/p f/d	Mean 10.2, med 9.2
Robinson et al, 1993.	Ugandan cases	Complex	Mean 5-6
Alcabes et al, 1994	Injecting drug users	Non-parametric	Med. 10.2 (95% CI; 7.9-12.3)
Volkov et al, 1994	TA-AIDS; Mexico	Direct calculation	Med 2.4 (max. 4)
Ward, 1994	TA-AIDS	Direct calculation	Med 4.8 (0.1-12.8)

[Notes on table: Tx = transfusion; med = median; paed. = paediatric; inf = infinity]

APPENDIX 9

INFORMATION SUPPLIED WITH THE TA-AIDS DATASET DISCS.

Transfusion-Associated AIDS Dataset

AIDS Program
Center for Infectious Diseases
Centers for Disease Control

This dataset contains information on transfusion-associated AIDS (TA-AIDS) cases reported to the CDC through December 31, 1990. These data have been voluntarily reported to CDC by State and local health departments, and are protected under an Assurance of Confidentiality (Sections 306 and 308 (d) of the Public Health Service Act, 42 U.S.C. 242k and 242m) which prevents disclosure of any information that could be used to either directly or indirectly identify individual patients or establishments. The statistical data contained on this dataset are being released for scientific use solely for the purpose of estimating the incubation time for AIDS. These data should not be published or rereleased in any form which would either directly or indirectly identify the month and year of transfusion or diagnosis of individual patients. Persons with transfusion-associated AIDS who are later found to have other risks for HIV infection are reclassified into the appropriate transmission category. Thus, the number of transfusion-associated AIDS cases on the dataset may not be the same as that found on the December, 1989 Monthly Surveillance Report.

The HIV-antibody test was licensed in March 1985 and in the following months was implemented by blood banks to screen blood donations. TA-AIDS cases reported with dates of transfusion after this time period have a high likelihood of having other risks for HIV-infection; of the first 34 TA-AIDS cases investigated with transfusion dates after June 1985, approximately 90% had other risks for HIV-infection or had also received unscreened blood donations before the recorded dates of transfusion. The TA-AIDS cases found on reinvestigation to have been infected with HIV after receiving antibody negative blood components are noted on the dataset.

The Centers for Disease Control and local health departments cannot verify that all of the TA-AIDS cases reported to have received blood before June 1985 actually received a blood transfusion or if the reported transfusions were donated by persons infected with HIV. Although misclassified cases are believed to represent a small proportion of reported TA-AIDS cases, the users of this data set should consider that some person with other risks for HIV infection may be incorrectly classified as TA-AIDS cases. As persons with TA-AIDS tend to be older than persons infected by other routes, this possible misclassification may be most likely for men in the 20-40 year old age group.

Some TA-AIDS cases are reported with the transfusion(s) having occurred very shortly before or at the time of the diagnosis of AIDS. As the shortest verified incubation period after parenterally-acquired HIV infection is 2 months, reported TA-AIDS cases reported with very short incubation periods may have received other transfusions or may have had other risks for HIV infection. Thus, all cases with a reported interval of time from transfusion to diagnosis of AIDS of 2 months or less are investigated again and, if new information is found, the reports for these cases are revised.

Approximately 21% of adult TA-AIDS cases are reported to CDC to have died the same month as the month of diagnosis, compared to about 12% for all other reported adult AIDS cases. Much of this difference can be explained by the age distribution of TA-AIDS cases compared to other cases. Within most risk groups, death in the month of diagnosis is more common for patients age 50+ than for younger patients, and 64% of TA-AIDS patients are age 50+ at diagnosis vs. 9% for other adults. Possible explanations for this higher early mortality are a poorer prognosis for TA-AIDS cases because of host or viral factors unique to this group, and delayed recognition by physicians of opportunistic illnesses in persons with transfusion-associated HIV infections. Further information can be obtained from the CDC surveillance public use data set, which contains risk group, age at diagnosis, month of diagnosis, and calendar quarter of death for each AIDS case. Users of the data set should consider that the period between the date of transfusion and diagnosis may be prolonged because of this possible delay in diagnosis.

The diskette specifications are:

Density	Double sided/Double Density
Format	IBM PC DOS/MS DOS
Code	ASCII
File Name	TD90Q4.DAT
Record size	36 bytes (34 bytes of data plus carriage control and line feed)
Record Terminator	CR/LF
Total number of records	4010
Total bytes	250783

The diskette contains exactly one record for each reported transfusion-associated AIDS case, coded as follows:

Variable	Columns	Description
Age	01	Age at diagnosis of first AIDS-related opportunistic disease, coded as 1 = Less than one year 2 = 1-12 years 3 = 13-29 years 4 = 30-39 years 5 = 40-49 years 6 = 50-69 years 7 = 70 years or older 9 = Unknown
SEX	02	Sex of patient, coded as 1 = Male 2 = Female
REPDATE	03-06	Year and month (YYMM) the AIDS case report was received at the CDC. Cases reported during 1981 are coded as '8199'.

Page 3 - TA-AIDS Dataset

Variable	Columns	Description
DXDATE	07-10	Year and month (YYMM) during which the first AIDS associated opportunistic disease was diagnosed. Cases with opportunistic diseases diagnosed before 1982 are coded as '8199'.
TXDATE1*	11-14	Year and month (YYMM) of either the first transfusion or the one most likely to have resulted in infection if that can be determined (see SUSDONOR). A code of "8899" indicates that at least one transfusion was received sometime before that indicated on TXDATE2, but that the month and year are unknown.
TXDATE2*	15-18	Year and month (YYMM) of the second transfusion.
TXDATE3*	19-22	Year and month (YYMM) of the third transfusion.

*Special note regarding transfusion dates: Every effort has been made in the case of multiple transfusions to determine the month during which infection was most likely to have occurred. However, information on patients varies considerably so that this can not always be done. If SUSDONOR=1 then it is likely that the transfusion(s) on the first date listed resulted in infection. If not, the number of units transfused (if available) may be used to suggest which date is most likely. Also, U.S. blood banks began screening donated blood in March, 1985, so that transfusions received after this date are less likely to have resulted in infection than prior transfusions. Currently the CDC collects only the date of the first transfusion and last transfusion prior to an AIDS diagnosis as part of routine AIDS surveillance.

Page 4 - TA-AIDS Dataset

UNITS1	23-25	Total number of blood units received during first transfusion, if known. Blank otherwise.
UNITS2	26-28	Total number of blood units received during the second transfusion, if known. Blank otherwise.
UNITS3	29-31	Total number of blood units received during the third transfusion, if known. Blank otherwise.
SUSDONOR	32	Coded as '1' if a case investigation identified a donor with either diagnosed AIDS, a positive HIV test, an abnormal immunologic finding, generalized lymphadenopathy, or a known history of being at risk for AIDS. Blank otherwise.
AGETX1	33	Age at first transfusion, coded the same as AGE.
VERIFIED	34	Coded as '1' if all transfusion dates are after March, 1985, and case investigation identified HIV transmission from blood screened for HIV antibody. Blank otherwise.
BACT	35	Bacterial infections, multiple or recurrent (including Salmonella septicemia)(pediatric cases only)
BURKL	36	Burkitt's Lymphoma
CANDESOP	37	Candidiasis, esophageal
CANDLUNG	38	Candidiasis, bronchi, trachea, or lungs
CMV	39	Cytomegalovirus infection
CMVRET	40	Cytomegalovirus retinitis (with loss of vision)
COCCI	41	Coccidioidomycosis, desseminated or extrapulmonary
CRYPTOCO	42	Cryptococcosis, extrapulmonary
CRYPTOSP	43	Cryptosporidiosis, chronic intestinal

Page 5 - TA-AIDS Dataset

DEMENTIA	44	HIV encephalopathy
HISTO	45	Histoplasmosis, disseminated or extrapulmonary
HS	46	Chronic mucocutaneous herpes simplex infection
IBL	47	Lymphoma, immunoblastic
ISO duration)	48	Isosporiasis, chronic intestinal (>1 month
KS	49	Kaposi's sarcoma
LIP	50	Lymphoid interstitial pneumonia and/or pulmonary lymphoid hyperplasia (pediatric cases only)
MAVIUM	51	Mycobacterium avium complex
MYCO	52	Mycobacterium, of other species or unidentified species
PC	53	Pneumocystis carinii pneumonia
PLB	54	Primary lymphoma of the brain
PML	55	Progressive multifocal leukoencephalopathy
SALS	56	Salmonella septicemia (adult cases only)
TB	57	Mycobacterium tuberculosis
TP	58	Toxoplasmosis
WASTING	59	Wasting syndrome due to HIV

Page 6 - TA-AIDS Dataset

AIDS-indicator opportunistic diseases (columns 35 through 59)

Columns 35 through 59 contain information each of the AIDS-indicator diseases listed on the AIDS confidential case report form. Each of these variables is one character long and is coded as follows:

0 = AIDS-indicator opportunistic disease was not diagnosed
1 = AIDS-indicator opportunistic disease was diagnosed definitively.
2 = AIDS-indicator opportunistic disease was diagnosed presumptively.

Categ (column 60)

1 = Case meets the pre-1985 surveillance definition
2 = Case meets the 1985 surveillance definition
3 = Case meets the 1987 surveillance definition and was diagnosed definitively
4 = Case meets the 1987 surveillance definition and was diagnosed presumptively

Race (column 61)

1 = White (not Hispanic)
2 = Black (not Hispanic)
3 = Hispanic
4 = Asian/Pacific Islander, American Indian/Alaskan Native, or unknown

****

Updated versions of the Transfusion-associated AIDS case dataset will be available each quarter upon request. Send your request along with a blank double-sided/double-density diskette to:

Gloria Gavin
Statistics and Data Management Branch
Division of HIV/AIDS CID
Centers for Disease Control
Mailstop E48
1600 Clifton Road
Atlanta, Georgia 30333

[Note: Race (column 61); Asian/Pacific Islander, etc. are coded as group 4 on the information sheet but are actually coded as group 9 on the disc.]

APPENDIX 10

DETAILS OF DATA USED FOR TSE/PRION DISEASE EXAMINATION, INCLUDING SOURCE, VARIABLES, AND EXPERIMENTAL METHOD (WHERE APPLICABLE)

SHEEP.

Natural scrapie.

Two sources of data were used; a literature search (references; table A10.1, below), and VIDA. Stringent efforts were made to avoid case duplication. The database was checked visually, by comparison with source, for errors. In these analyses, natural scrapie in sheep includes cases of disease where the animals were kept in direct or indirect contact with experimentally induced cases of scrapie, either by chance or in a deliberate effort to investigate the possibility of contact as a natural transmission method.

Table A10.1 Data references: natural scrapie in sheep.

Greig (1940b) Parry (1962) Dickinson et al (1965)	Brotherston et al (1968) Gordon (1966)	Hourrigan et al (1979) Wooldridge (1991)

Since age at onset (a/o) was recorded in different ways in different papers, data from Parry (1962) and from Wooldridge (1991), papers detailing large numbers of cases, are analysed separately. VIDA data is also analysed separately. The remainder (108 cases) comprise the general literature search (GLS) described in the analyses. Additional details on these datasets are given below.

Data from Parry (1962).

Parry (1962) gives a/o details of 1008 cases by sex, in graphical form in 6-monthly increments, starting with 12-18 months. The increment mid-point is used in analyses, giving fifteen months as the first age point.

Parry collected data between 1952 and 1960 from approximately 50 pedigree sheep flocks of ten breeds, plus some commercial flocks, dispersed over 17 counties of Great Britain, giving variable local conditions. Most farms were in the 'Oxford flock health scheme'. They were visited regularly, and had good identification and recording systems. Diagnosis was based on 'unequivocal' clinical signs plus fatal outcome, and sometimes histopathology. Parry suggests that the number of diagnostic errors is likely to be small.

In these flocks, only one ram is kept for every 20-40 ewes, and most are culled before three years of age as normal management practice. For ewes, generally 15 to 20 % are between 4.5 and 11

years, with no marked difference between flocks; prior to this there is probably a large cull at around four years of age.

Data from Wooldridge (1991).

Wooldridge (1991) reported a postal questionnaire survey of scrapie in sheep spanning the years 1980 to 1990 inclusive, comprising data from 167 farms. The survey was not random and many biases were identified. The a/o was unrecorded for 591 (31%) cases; estimates were given in some cases. Cases comprise both those clinically diagnosed (with possible misclassification), and VIC diagnosed cases with histopathological confirmation. Some indication of the a/o was obtained for 1334 cases, 160 being VIC confirmed. For animals over five years old, questionnaire design allowed only one category ('five or more years') and the full extent of the right tail is thus lost. Therefore no means or ranges are calculated.

Data from VIDA.

The VIDA database records all scrapie in sheep confirmed histopathologically by the VIC's, from 1975 onwards. Data used here comprises 2361 such cases recorded by 7/10/1993. There are many biases in VIDA data (see Wooldridge, 1991), and no way of reliably estimating the situation regarding the culling of older animals in the population from which VIDA data comes.

Data from the GLS.

A total of 108 additional cases were found from the GLS. For 1 case, i/p but not a/o was given.

For 107 sheep a/o was given, either in months or years. For 55 cases, it was known that the dam was herself either a case, or had been exposed to a source of infection.

In order to analyse cases as one group, data given in years was converted to months (multiplication by 12), resulting in converted data being 'bunched' at twelve month intervals. Smoothing was therefore carried out by redistribution of those cases converted from 'years' to 'months' equally throughout the year, as two separate groups, as account was taken of the exposure status of the dam. It is recognised that this may introduce bias, particularly for any year which spans a steep part of the f/d curve.

Even if scrapie is always the result of infection (unproven), there is no way of knowing when infection occurred. If infection from the dam is possible, infection may be perinatal, or possibly

intrauterine. Two groups are identified; cases where the dam either was known to have contacted an infection source during pregnancy, or became a case herself (pregcode = 1; 55 cases), and cases where no such factors were recorded (pregcode = 0; 57 cases).

For ten cases, including nine with a known a/o, the sheep were positively recorded as grazing pasture previously grazed by scrapie infected sheep, a potential infection source which was actually being investigated using these sheep (Greig, 1940). The period of grazing was recorded, as was the time of disease onset. Two of these cases were born whilst their dams grazed this pasture and they were thus exposed from the time of birth. The remaining eight sheep were believed to have had no exposure prior to grazing this pasture, which was therefore assumed to be their source of infection. Truncation of the right tail by culling or flock dispersal is probable (Greig, 1940). For one additional case, both a/o and time from start of exposure to infection was given. Prior to recorded exposure, this animal was also believed to be from a scrapie-free source.

Thus there was a total of 11 cases with a documented exposure window. Actual infection time during this 'window' is unknown (if transmission was the aetiological mechanism). The i/p is therefore arbitrarily estimated as time from first recorded exposure until disease onset, the maximum possible value.

Experimental TSE in sheep.

Data was mainly obtained from a literature search (references; table A10.2). Experimental BSE data was obtained from M Dawson (personal communication). Stringent efforts were made to avoid case duplication; the database was checked visually by comparison with source, for errors. Details given for both experimental methods and results vary, and this is relevant when interpreting results.

Table A10.2 Data references: experimental TSE in sheep.

Cuille & Chelle (1936)	Pattison & Millson (1960)	Dickinson et al (1968)
Greig (1950)	Pattison & Millson (1961a)	Pattison et al (1972)
Wilson et al (1950)	Dickinson et al (1966)	Pattison et al (1974)
Pattison et al (1959)	Gordon (1966)	Gajdusek (1977)

A total of 447 cases of experimental sheep TSE was recorded in the database. The variables with some information recorded for most cases are passage status (either directly from a natural TSE case, or from a previous experimental TSE case, or from dam inoculation with TSE agent), immediately preceding donor species, original disease, infection route and infecting substance (i.e. tissue) used. Original disease segregated with immediately preceding donor. Sex, dose, agent strain or original source (e.g. SSBP1) and breed were, overall, poorly recorded. Two cases were

female, and for the remainder, sex was unspecified; this factor could not therefore be examined. Dose was recorded in different ways making comparisons difficult or impossible.

Subgroup 1.

Definition. Sheep with experimental *scrapie*, where the immediately preceding *donor was a sheep*, inoculation was *into the CNS*, and the infecting substance used was *CNS material only*. There were 77 such cases.

For some of these cases, additional information was available on passage status or number, breed, age at infection, or dosage. Thirty of the 77 sheep were infected from a natural case of scrapie (i.e. passage 1), and the remaining 47 from a previous experimental case. One experimental group had passage number given for seventeen cases, from passage 2 up to passage 9, but numbers are too small to investigate in any formal way. For 38 cases, age at infection was given. Of these, 21 were direct from natural cases and infected at 7 months

For all 77 cases, the breed of case was given, and for 54 cases, the breed of donor was given. Case breeds are as follows: Cheviot; 61; Greyface, 5; Welsh Mountain, 5: Halfbred, 2; Blackface, 2 Figures for donors are: Cheviot, 19; Greyface, 7; Welsh Mountain, 18; Halfbred, 3; Swaledale, 7 Since numbers were small for many donor/case breed combinations, breeds were regrouped as either 'donor and case = the same breed', or 'donor and case = different or unknown breeds' for analysis.

Seventy-five of the 77 cases had information on dosage. For all this was 1 ml of a 10% dilution of brain preparation. However, donor brains may have contained widely varying amounts of agent.

Subgroup 2.

Definition. Sheep with experimental *scrapie*, where the immediately preceding donor was a *sheep with experimentally derived scrapie*, infection was by *s/c inoculation*, and the infecting substance was *CNS material only*. There were 97 such cases.

Sex was unknown for any case. All were infected using a 10% dilution of brain material, all bar two being from the same preparation, the 23rd passage of a particular agent preparation; the two other cases were the 24th passage. Dose was given as volume of brain extract, 89 receiving the same 2 ml volume.

GOATS.

Natural scrapie.

Two sources of data were used; a literature search (references; table A10.3, below), and VIDA. Stringent efforts were made to avoid case duplication. However, it was impossible to avoid all overlap between cases from the literature search and VIDA, therefore they are analysed separately. The database was checked visually by comparison with source, for errors. In these analyses, natural scrapie in goats includes cases of disease where animals were kept in direct or indirect contact with experimentally induced cases of scrapie, either by chance, or in a deliberate effort to investigate the possibility of contact infection as a natural transmission method.

Table A10.3 Data references: natural scrapie in goats.

Mackay & Smith (1961)	Harcourt & Anderson (1974)	Toumazos & Alley (1989)
Brotherston et al (1968)	Stemshorn (1975)	Wood et al (1992)
Hourrigan et al (1969)	Hourrigan et al (1979)	

Culling of goats occurs, particularly for males, but it is difficult to assess quantitatively due to lack of data, as is the 'natural' lifespan of a goat.

Data from the GLS.

A total of 66 cases was found. For 37, sex was recorded, 32 being female, 5 male. The a/o was given in months for 47 cases, the remainder were given in years or six-month intervals. In order to analyse cases as one group, data was converted to months, the method depending on information available. For example two cases reported together as eight years, and four and a half years respectively, were converted as (8x12=96) and (4x12 + 6 = 54) months, whereas 'between three and four years' was converted to (36 + 6 months = 42 months). Conversion is thus a possible source of error. Small numbers and mixed recording methods meant that 'smoothing' was not feasible.

Three contact groups were identified. Fifteen cases where the dam also became a case, 36 cases where flock or herd contact with scrapie was recorded, and 15 cases with no contact recorded. However, for the group with no contact recorded, unrecorded contact may have occurred. In only one case was it stated categorically that there had been no contact with either sheep or goat, directly or via pasture. Similarly, where flock/herd contact is recorded, this does not exclude absolutely the possibility of the dam being scrapie positive.

For 9 cases, the latest possible age by which removal from known exposure had occurred was recorded; all were exposed from birth. It is difficult to be certain that such removal was effective. Some of these goats may have been kept together subsequently, with possible lateral spread.

Data from VIDA.

The VIDA database records all scrapie cases in goats confirmed histopathologically by VIC's, from 1989 onwards. Data comprises 28 such cases recorded by 7/10/1993. Bias and other aspects of VIDA data are as for sheep.

Experimental TSE in goats.

Data was mainly obtained from a literature search (references; table A10.4 below). Experimental BSE data was obtained from M Dawson (personal communication). Stringent efforts were made to avoid case duplication; the database was checked visually by comparison with source, for errors. Although approximately 75% of the cases originate from the same group of investigators, at the same research establishment, details given both for experimental methods and results vary, as they do for sheep. This is relevant when interpreting results.

Table A10.4 Data references; experimental TSE in goats.

Gordon et al (1957)	Pattison & Millson (1961c)	Marsh et al (1969)
Pattison et al (1959)	Pattison & Millson (1962)	Pattison et al (1972)
Gordon (1960)	Pattison (1964)	Hadlow et al (1974)
Pattison & Millson (1960)	Pattison (1965a)	Gajdusek (1977)
Pattison & Millson (1961a)	Pattison (1965b)	Gibbs et al (1979)
Pattison & Millson (1961b)		

A total of 400 cases of experimental TSE's in goats was recorded in the database. As for sheep, and for the same reasons, the first i/p analysis utilises the complete dataset. Again, as for sheep, variables with some information recorded for most cases are passage status (either directly from a natural TSE case, or from a previous experimental TSE case), immediately preceding donor species, original disease, infection route and infecting substance used. Again, original disease segregated with immediately preceding donor. Dose was poorly recorded, and a large number of preparation methods and pre-treatments were used. Forty cases were female, 126 were male. For the remainder, sex was unspecified.

Subgroup 1

Definition. Goats with experimental *scrapie*, where the immediately preceding *donor was a goat*, inoculation was *into the CNS*, and infecting substance used was *CNS material only*. There were 201 such cases, over 50% of the total.

For some cases additional information was available on passage number, age at infection, dose, and other manipulations. For 60 cases, the number of goat-to-goat passages was given. Passage 2 indicates the donor was a goat infected directly from a natural sheep case, that is, the donor was itself a passage 1 case, and so on. All cases had similar experimental conditions and protocol. For 71 cases age at infection was given, ranging from 1 to 9 months. Numbers infected at any age were small, the maximum being 23 cases at 6 months, most age categories having ten or fewer cases infected.

To examine the effect of dose, Pattison & Millson (1961b) describe a complete dilution experiment starting with a single scrapie agent homogenate. Dilutions are described as 1/5 (i.e. 20%), 1/10, 1/20 and so on, down to 1/81920, in fourteen doubling steps; i.e. 15 strengths in all. Either one, two or three goats were given 1 ml at each dilution (23 in all), and all but one goat (at the most dilute level, observed for 19 months) developed disease; this goat is excluded from analysis.

Ten other agent manipulations were identified; the majority had very few examples. 'Enzyme treatment', freeze/thaw cycling, boiling for up to 3 hours, autoclaving, and ether extraction all had 8 or fewer cases. There were 19 cases of formalin treatment, and 15 cases of host pre-treatment by CNS inoculation with 100 mls of 'dense suspension of normal goat brain' carried out 10 weeks (2.5 months) before agent inoculation.

Subgroup 2

Definition. All cases with a plausible 'natural' infection route; *s/c (plus i/d) inoculation* (27 cases) and *oral dosing* (7 cases).

There were 34 cases, of which 28 were s/c (or i/d) inoculated, and 6 oral infections

Subgroup 3

Definition. Passaged directly *from a 'natural' case of scrapie in sheep,* the generally assumed infection source in natural goat scrapie.

There were 14 cases. For 6, breed of donor sheep was known: 4 Welsh Mountain and 2 Swaledale; all six were CNS transmitted using CNS material, but different agent doses or goat breeds may have been used.

PRIMATES.

Two sources of data were used, a literature search and personal communications (references; table A10.5). These sources were amalgamated. Age at inoculation was generally unknown; other details available varied from case to case. For data originating from the National Institutes of Health (NIH) (over 90% of the 635 cases), animals were always observed until disease onset or death from other causes (P Brown, personal communication). The proportion of inoculated primates that die from TSE varies both with the primate species, and to some extent with the type of inoculated material, i.e. sporadic, familial, iatrogenic, kuru etc. (P Brown, personal communication). Therefore estimating the extent of right censoring is complex (see Brown et al, 1994 for available information).

Table A10.5 Data references and personal communications: experimental TSE in primates

Gajdusek, 1977 Peterson et al, 1978	Masters et al, 1979 Gibbs et al, 1979	Baker et al, 1990	M Dawson, CVL (p.c) P Brown, NIH (p.c.)

Primate classification used in analyses is as given in source data, since no further taxonomic information is available; sometimes this refers to a species, sometimes a genus. As for goats and sheep, experimental conditions varied. Again as for goats and sheep, and for the same reasons, the first i/p analysis utilises each complete dataset. Many experimental variables are present, and information available varies. Table A10.6 gives information on some variables.

Table A10.6 Experimental TSE in primates; some major variables by primate group.

Group	Passage		Original disease					Info on 'dose' given
	Pg=1	Pg>1	CJD (sp)	CJD(f) +GSS	Kuru	Scrapie	BSE	
Marmoset	20	30	12	5	31	0	2	15
Capuchin	67	4	48	15	6	2	0	63
Squirrel m.	306	16	266	40	9	7	0	296
Spider m.	54	3	36	9	11	1	0	44
African green	18	2	19	1	0	0	0	20
Cynomolgus	8	4	9	0	0	3	0	9
Rhesus	25	5	25	1	2	2	0	26
Chimp	63	10	49	7	17	0	0	54

[Notes on table:- Pg=1: passage direct from natural case: Pg>1: passage from experimental case
Information given on 'dose' varied in report method;
sp = sporadic, f = familial.]

Data on additional variables, including infection route, infecting substance, formalinisation of agent etc., is given for some cases. Relevant taxonomic details are given in appendix 21.

Datasets with greater homogeneity; Subgroups 1 to 6.

To better assess whether any consistent i/p f/d pattern exists for similarly treated sets of experimental animals, more homogenous sets of experimental conditions were selected and

applied to each primate group. In this process numbers within subgroups necessarily decreased. The following subgroups were defined:-

Subgroup	Original disease	Passage number
1	Kuru	1
2	CJD (all types) + GSS	1
3	"	>1
4	CJD (sporadic only)	1
5	CJD (familial) + GSS	1
6	CJD (sporadic only) **	1

** Additional conditions for subgroup 6:- all from single laboratory (for a given primate); dose given as '10% dilution of brain extract'; no 'manipulation' of extract by formalin, paraffin wax , or fractionation recorded; inoculation route includes an intracerebral component, either alone or as part of a mixed route infection. Subgroup 6 was therefore the most homogeneous.

Data for squirrel monkeys.

Numbers of cases for squirrel monkeys, the largest group, allowed examination in more detail. However, they are a genus, not a species and it is not certain that all cases were from the same species. Nevertheless, the majority were from one laboratory, therefore if different, possibly they would have been categorised separately.

Many sources of variation are recorded. For most cases passage status (whether directly from a natural TSE case, or from a previous experimental case), original disease, preceding donor, infection route and substance were recorded. For some cases, a genetic mutation of PrP gene in the originating case, or manipulation of the inoculum was recorded. Dose was recorded in a variety of ways. Sex was unknown in all cases. For most variables, the majority of cases had the same value, and i/p f/d for that category would therefore have resembled the full set; others comprised very small numbers of each of several variables.

For the 129 squirrel monkeys in subgroup 6, it was possible to further stratify by inoculation route: 53 cases had been CNS inoculated only, 76 had undergone multiple site inoculation (but which included a component of CNS inoculation in all cases).

RODENTS.

BSE in mice.

Data was kindly supplied by H Fraser, NPU, Edinburgh, and comprised BSE i/p in 196 RIII (genotype *'sinc' s7s7*) mice, sex (females, 100; males, 320), infection age and date, and originating BSE-case identity number.

All cases represent primary transmission (passage 1). Inocula originated from seven field cases of cattle BSE. Other than this, and inoculation date, experimental protocol was unchanged. Mice were injected intracerebrally (0.02 mls) and intraperitoneally (0 1 ml) with a 10% brain homogenate. All inoculations from a single case of BSE were performed on the same date. All mice underwent daily examination for signs of BSE, and histopathological examination after death or killing. Initially, 213 mice were inoculated; 17 (8%) died from other causes.

Inoculation age and BSE group tend to segregate together. Compared with overall inoculation age range, within each BSE group range is tight, presumably due to age of mice available when required, which is likely to be clustered by litter membership.

To attempt to reduce effects of BSE group on c/s, in order to further investigate any age effect, data was regrouped as follows:-

Groups A,B,C,D: Dataset divided into four age groups with *equal age-range-spans* of 26 days for inoculation age.
Groups E,F,G,H: Dataset divided into four age groups as *equally sized* as possible without splitting mice inoculated at the same age.

Scrapie in mice.

Data was extracted by measurement, from histograms (Dickinson & Miekle, 1971) where it was grouped by 10 day intervals. Data comprised an i/p for 407 cases, plus scrapie agent 'strain'; for some cases, mouse 'sinc' genotype was given.

Experimental design was complex, in order to investigate effects of host genotype and agent strain interactions. All mice were inoculated intracerebrally with 0 02 mls of fluid, an estimated agent dose of 10^4LD_{50}. Inoculation age range was 22-152 days; a decrease in i/p of 0.05 +/- 0.01 days per increased day of age had previously been found.

Two fully inbred strains of mouse were used: C57BL (homozygous for the s7 'sinc' * allele) and VM (homozygous for the p7 'sinc' allele); these were also crossed to produce s7p7 offspring. Further crossed generations were also used; these comprise the 'unrecorded genotype' in the data. All inoculated mice developed scrapie, except for 'a few' which died early of unrelated causes. Two strains of scrapie agent were used, ME7 and 22A. Both had previously been passaged several times through mice, most recently twice through VM mice.

[* The 'sinc' gene, a contraction of 'scrapie incubation period gene', was originally characterised by the response of various strains of mice to infection with certain scrapie agents, including ME7; s7 originally signified the 'short i/p' allele and p7 the 'prolonged i/p' allele; the correlation of the ME7 results with this designation is therefore expected; that strain 22A did not follow this pattern was a major focus of interest of the authors.]

CJD in guinea-pigs.

Seventy-nine Hartley strain guinea-pigs were inoculated intracerebrally (Manuelidis & Manuelidis, 1979), using 0.1 ml of a 10% dilution of brain material. All developed CJD, thus there was no truncation. Passage number varied from 1 (directly from two human cases) to 10; two first passage guinea-pig cases are described; it seems probable that one derived from each human case.

HUMANS.

For analysis, data is categorised initially into sporadic disease, non-sporadic disease, and documented iatrogenic disease. Sporadic disease is defined here as non-recording of any known risk factor for TSE/prion disease. Non-sporadic disease is defined as any case (other than iatrogenic) with a recorded risk factor, and includes all cases with a recorded genetic mutation, as well as those with a documented family history.

Data supplied by Dr R Will, National Creutzfeldt-Jacob Disease Surveillance Unit, Edinburgh.

Data used comprised information on 376 cases of sporadic CJD with onset from 1969 to 1990, supplied on disc as a database. Sex and diagnostic method (histopathology, or clinical plus characteristic EEG) was recorded, plus age at death. For the period 1980 to 1990, where disease duration was over one year (6%), a correction was made to give age at onset rather than age at death; comparable information was unavailable for earlier cases.

Additionally a list of i/p data for iatrogenic cases (pituitary hormone injections, and dura mata graft) of international origin was supplied.

For pituitary hormone injection aetiology, data comprises 23 cases (8 females, 15 males), from a variety of countries. Treatment (tx) generally consisted of multiple injections, over widely differing durations and occasionally continuing until disease onset. Treatment (tx) duration is summarised as follows:-

	Mean (years)	Range (years)
Age at first treatment:	9.2	1-30
Duration of treatment:	7.6	1-17
Age at onset:	25.0	10-44

Precise infection time during treatment is unknown, therefore the i/p was estimated in two ways: from first treatment to disease onset (maximum possible), and from treatment mid-point to disease onset. Since conclusions on the i/p f/d were similar, only the results from the first (maximum possible i/p) estimate are shown in the thesis.

Eight of the 9 dura mata grafts cases were also from this source; presumed infection time is well documented.

Data supplied by Dr Paul Brown, National Institutes of Health, Bethesda Washington, USA.

I was allowed access to all records of TSE's/prion disease (i.e.; sporadic CJD, familial CJD, GSS, FFI) recorded at NIH. There were several categories:-

Experimentally transmitted cases as of November, 1993, from USA, abroad, iatrogenic cases, and kuru.
Cases in France from 1968 to 1982.
Additional observations by, and reports and referrals to NIH, often having particular features including genetic mutations, familial history, or ethnicity of interest.

There was category overlap, but cases were individually identified and duplication eliminated. Data was abstracted from records, and after database entry, checked visually for errors.

Of 785 cases, 271 were experimentally transmitted; this is considered diagnostic. Transmission criteria have altered with time. Early in investigations, transmission was routinely attempted. Later it was undertaken only for particularly interesting cases, leading to selection bias. Most cases without positive transmission were histopathologically confirmed, the remainder diagnosed by clinical signs. However, for the French survey approximately 40% were diagnosed clinically.

There is no notification requirement to NIH; it occurs because of a clinicians interest, or a request for material of a particular type, introducing a selection bias. However, the French survey (the largest single component), a countrywide search for all clinical cases of CJD is less likely to suffer this bias, although other clinical manifestations, for example GSS, were not included.

Eighteen kuru cases with recorded a/o (the accuracy of which is unknown) were obtained from NIH data; the time of infection is unknown. All transmitted to laboratory animals.

Additionally a list of French pituitary hormone injection cases was supplied, comprising 21 cases. There is evidence to suggest that use of contaminated inoculum in France may have occurred only in 1985 (P Brown, personal communication), therefore i/p was estimated from this time.

Data from a general literature search (GLS).

The literature search for all clinical manifestations of TSE/prion disease was time delineated, rather than exhaustive. Diagnostic method depends on the paper, and includes histopathology, genetic mutation, and family history. There is a probable reporting bias; cases with particular

features, whilst reported, may not be typical. Extensive cross-referencing within the GLS, to eliminate duplication was believed to be successful. Checking of the data within the database was done as for the previous dataset. After database entry, data was error-checked visually.

Table A10.7 Data references; TSE in humans; GLS.

Alema (1973)	Nochlin et al (1989)	Nakazato et al (1991)
Gajdusek (1977)	Blisard et al (1990)	Tranchant et al (1991)
Brown et al (1979)	Collinge et al (1990)	Amano et al (1992)
Mayer et al (1979)	Nakazato et al (1990)	Collinge et al (1992)
Neugut et al (1979)	Nisipeanu et al (1990)	Genthon et al (1992)
Galvez et al (1980)	Sadeh et al (1990)	Hsiao et al (1992)
Foncin et al (1982)	Tateishi et al (1990)	Kretzschmar et al (1992)
Manuelidis et al (1985)	Brown et al (1991)	Terao et al (1992)
Vinters et al (1986)	Hsiao et al (1991a)	Collinge et al (1993)
Pearlman et al (1988)	Hsiao et al (1991b)	Kitamoto et al (1993)
Farlow et al (1989)	Kretzschmar et al (1991)	Martinezlage et al (1993)
Hsiao et al (1989)		

GLS data includes sporadic, familial, and iatrogenic TSE/prion disease cases, separated for analysis.

Data duplication.

It proved impossible to cross-reference a considerable number of GLS cases with other data, therefore not all case duplication could be eliminated. However, where identified, the case was deleted from the GLS to reduce duplication as far as possible. The NIH dataset is known to contains some duplicates of UK data. The exact extent of duplication is difficult to assess, therefore individual analysis of the datasets, rather than combined analysis, was undertaken.

Genetic mutations recorded in data.

Mutations of PrP gene recorded in this data comprise codon substitutions, and base pair inserts. Table A10.8 details the codon mutations. Insert mutation details recorded vary, but comprised so few of each that they are analysed together.

Table A10.8 TSE in humans; codon substitutions recorded in data used and resultant amino acid change.

Codon substituted	Amino-acid change in protein.
217	Glutamine -> Arginine
200	Glutamic acid -> Lysine
198	Phenylalanine -> Serine
178	Aspartic acid -> Asparagine
145	Tyrosine -> 'Stop'
117	Alanine -> Valine
102	Proline -> Leucine

For NIH data, 214 cases were recorded (189 with known family history); for GLS data, there were 123 cases (122 with family history). Numbers from each data source are as follows:-

Mutation	NIH data	GLS data
217	-	1
200	69	19
198	16	19
178	73	-
145	-	1
117	17	25
102	20	29
Insert	19	29

For analyses, the larger set from each source was used (see appendix 23 for method of estimation of overlap).

CATTLE.

Using programmes specially written by A Mitchell, on 5/8/94 data was abstracted as files of fixed field-length, from the BSE database maintained on the Sequent computer at CVL Epidemiology Department, Addlestone, Weybridge, Surrey, KT15 3NB.

By 29/7/94, 132,538 histologically confirmed cases were recorded in this database, from 31,138 different farms, plus suspect cases undergoing investigation. Many records lack sufficiently precise information on age, date of onset, and feeding history (recorded by the year) for these analyses. The following criteria are used for all cases included in analyses:-

(a) Month and year of birth recorded; month and year of disease onset recorded.
(b) BSE confirmed by histopathology.

For all cases, date of infection is unknown; a/o is therefore used as an estimate of i/p, which necessitates assuming infection close to birth. Therefore, cases born before 1/7/81 are excluded, as exposure from birth was considered unlikely (Wilesmith et al, 1988). There is often a reporting lag of several months duration, therefore cases with onset within six months prior to data abstraction are excluded.

Cohorts are described by 'calving season of birth', from the 1st July until the subsequent 30th June inclusive, to fit the usual pattern of cattle management and farm recording. For example, cohort 81 indicates birth from 1/7/81 to 31/6/82 inclusive.

Three final cattle groups were obtained as follows:-

Cattle group 1: 488 cases.
In order to closely estimate i/p by a/o, cattle were all known or strongly suspected to have been offered protein concentrates (potentially containing ruminant derived protein) from the first few weeks of life, and none after their first year; achieved by selecting cattle recorded as being fed suspect feeds only in their first year of life. In addition, to reduce potential recording errors, all cases were homebred, in dairy herds, from a farm where all homebred cases were recorded as having a similar feeding regime.

Cattle group 2: 204 cases.
In order to closely estimate i/p by a/o, cattle were all strongly suspected to have been offered protein concentrates (potentially containing ruminant derived protein) from the first few weeks of life, and none after their first year; achieved by selecting cases occurring on beef suckler farms,

which were purchased onto those farms at 365 days old or less, with no record of being fed suspect feeds subsequently.

Such purchased replacement calves are generally bred on dairy farms and surplus to requirements; they will generally have had concentrates from weaning, around 3 weeks of age, as for other dairy-farm calves; they are assumed infected by such feeding before purchase.

Cattle group 3: 5,810 cases.

Censoring (as for AIDS; see chapter 6 and Peterman, 1987) was anticipated in the BSE epidemic; left censoring results in non-observation of early, short i/p cases. 'Pre-order' farms were identifying and reporting cases before BSE became notifiable. Farmers on these farms may therefore be more observant, and less likely to have missed early cases, thus possibly reducing left censoring. Even on these farms, it was unlikely that left censoring was eliminated, and it is known that histological samples were unobtainable for a number of cases with typical clinical signs (J Wilesmith, personal communication). All 'pre-order' farms were all selected. All cases on these farms meeting the original criteria were then selected from these farms. In addition, each case was assigned a 'feedflag' (F) as follows:-

F	Feeding status in first year of life:-	N	%:-
Y	Recorded as fed suspect feeds in first year.	4678	80.5
N	Recorded as not fed suspect feeds in first year.	325	5.6
U	Feeding in first year unknown/unrecorded.	807	13.9

For examination of the effect of censoring, in order that all cattle in the analysis were exposed to a possible infection source in their first year of life, and were therefore comparable to group 1, only feedflag status 'Y' is used.

Typical cattle 'concentrate' feeding regimes.

Until the 'ban' on feeding ruminant protein to ruminants (18th July 1988) all concentrates potentially contained ruminant derived meat and bone meal (MBM), and thus, possibly, the BSE agent. Dairy-herd calves are generally weaned at a few days of age onto milk substitutes. In addition, they are generally introduced to some form of 'concentrate', for example calf pellets, either at the same time, or very soon after. The majority of dairy calves have had access to such calf concentrates before one month of age, and are gradually weaned off the milk substitute. Soon, roughage is added, and 'rearing' concentrate is generally substituted for calf concentrate. During their second year, calves intended as replacement dairy heifers often receive no concentrate, their diet comprising only grazing and conserved fodder. Once they have borne their first calf, at around two years, they generally receive, in addition to grazing, some form of adult dairy ration concentrates. Beef-herd calves and cows may well receive no concentrates at any stage.

APPENDIX 11

FILIBEN'S METHOD FOR FREQUENCY DISTRIBUTION ANALYSIS FOR COMPATIBILITY WITH NORMALITY AND LOG-NORMALITY USING MINITAB STATISTICAL SOFTWARE.

Theoretical basis of method (Filiben, 1975).

Filiben's test statistic, the normal probability plot correlation coefficient *r*, is defined as the product moment correlation coefficient between the ordered observations X_i and the order statistic medians M_i from a normal (0,1) distribution.

$$r = \mathrm{Corr}(X, M)$$

$$= \frac{\sum(X_i - \overline{X})(M_i - \overline{M})}{\sqrt{\sum(X_i - \overline{X})^2 \sum(M_i - \overline{M})^2}}$$

A normal frequency distribution will give near-linear normal probability plots, which will in turn give near-unity values for the probability plot correlation coefficient *r*.

Filiben's paper provides a table of the percentage points of the normal probability plot correlation coefficient r, for sample size values up to N=100, which the author states in a corrigendum to be nearly equivalent to the Wilk-Shapiro statistic (Shapiro & Wilk, 1965). Additional values for sample sizes up to N=2000 were kindly calculated by R Sayers.

Method as used in Minitab statistical software, release 7.2 (1989).

The method used is as described in Minitab reference manual (1989), release 7, section 4-8; normal probability plots and correlation.

For estimation of compatibility with normality, the data used (X_i in Filiben's method; column C1 in the Minitab manual example) is the incubation period for each case. For estimation of compatibility with log-normality, the data used is the logarithm (to base 10) of the incubation period for each case. The command NSCORES produces, for each value of X, the corresponding value for M.

For each sample, a normal probability plot (X_i plotted against M_i), and a correlation coefficient (X_i correlated with M_i) designated CC in this thesis (*r* in Filiben's method) is then produced.

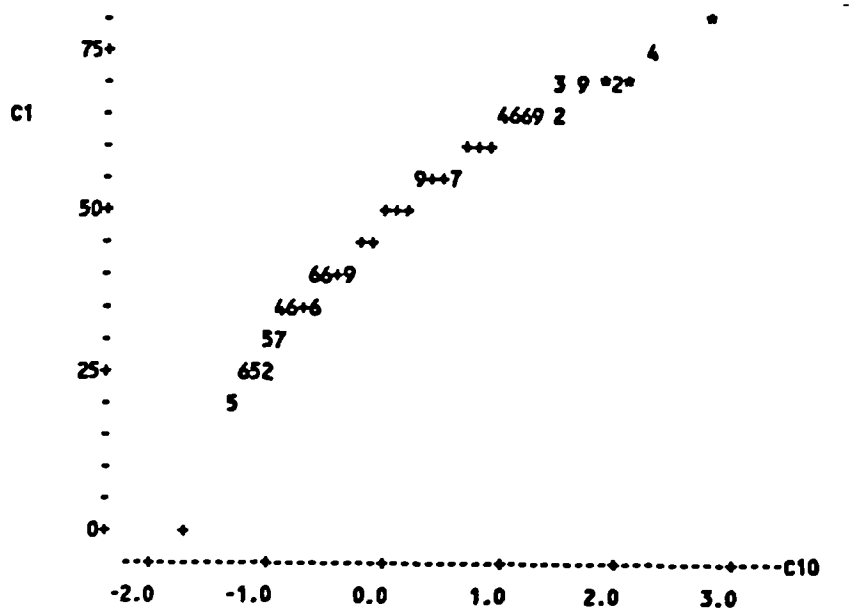
If the dataset f/d is close to normal, then the plot approaches linearity, and the CC approaches 1.

The Minitab manual also provides a table of 'critical values' (C/v's), which corresponds to, but is much less extensive than, the table of the percentage points of the normal probability plot correlation coefficient *r* given in Filiben's paper. The values are also not quite identical, with the third decimal place generally being one integer different; they are actually the Wilk-Shapiro statistic referred to by Filiben.

If the calculated CC falls below the C/v for a given percentage point then the hypothesis of normality (or log-normality, when using the log-transformed data) is rejected with that level of probability; the f/d is incompatible with that distribution. Due to the lack of percentage points given in the Minitab C/v table, Filiben's table was used for samples with size greater than 20.

Unless otherwise stated, the percentage point level (P value) used for estimating compatibility with normality (or log-normality) for all datasets is the 0.05 level.

Example of the normal probability plot produced by Minitab.



Age at onset.

Estimation of compatibility with normality (or log-normality) of f/d for age at onset was calculated by the same method.

APPENDIX 12

THE METHOD DESCRIBED BY SNEDECOR AND COCHRAN (1967) FOR ESTIMATION OF FREQUENCY DISTRIBUTION COEFFICIENT OF SKEWNESS USING MINITAB STATISTICAL SOFTWARE.

Theoretical basis of method (Snedecor and Cochran, 1967).

Skewness of the frequency distribution (f/d) of a population can be measured by the average value of $(X - \mu)^3$, a quantity called the 'third moment about the mean'. An extended right tail gives a positive value, and an extended left tail a negative value for the skewness. In order to make this measurement independent of scale, it is divided by σ^3. The resulting measurement is called the coefficient of skewness, denoted by $\sqrt{\beta_1}$.

The coefficient of skewness (c/s) of a sample from that population is denoted by $\sqrt{b_1}$, and calculated as follows:-

First, m_3 and m_2 are calculated.

$$m_3 = \sum (X - \overline{X})^3 / n$$

$$m_2 = \sum (X - \overline{X})^2 / n$$

The coefficient of skewness $\sqrt{b_1}$ is then calculated.

$$\sqrt{b_1} = m_3 / (m_2 \sqrt{m_2})$$

In calculating m_2, the sample variance, division is by n rather than the more usual $(n - 1)$, making subsequent calculations slightly easier.

Method as used in Minitab statistical software, release 7.2 (1989).

The data, or X values, used for calculation of the coefficient of skewness ($\sqrt{b_1}$ in the above method; designated c/s in this thesis) are, for the f/d of the raw data, the incubation period values. For the f/d of the log-transformed data, the logarithm (to base 10) of the incubation period values are used.

A macro was used to compute the c/s values, as follows:-

```
mean c1 k1
let c5=(c1-k1)**2
let c6=(c1-k1)**3
mean c5 k2
mean c6 k3
let k4=k3/(k2*sqrt(k2))
print k1 k4
end
```

Where c1 = X values and k4 = c/s.

The skewness probability (P) value was calculated by the method of Royston (1985), using a FORTRAN programme (kindly written by R Sayers) on a Prime computer.

Age at onset.

Skewness of f/d for age at onset was calculated by the same method.

APPENDIX 13

SHEPPARD'S METHOD FOR THE CORRECTION OF THE ESTIMATE OF THE STANDARD DEVIATION FOR FREQUENCY DISTRIBUTIONS WITH GROUP INTERVALS.

Theoretical basis of the method (Sheppard, 1897).

An error in the estimation of the variance (and thus the standard deviation) is introduced, when grouping data, by the assumption that each observation takes the mid-value of the group. For frequency distributions with equal group intervals, an allowance can be made for this error.

Method.

If C is the length of the group interval, then Sheppard's correction is made by reducing the variance by $(1/12)C^2$. Where observations $x_1, x_2 \ldots x_n$ have frequencies of $f_1, f_2, \ldots f_n$, the formula for the standard deviation S (designated SD in this thesis) therefore becomes:-

$$S = \sqrt{\left\{ \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f} \right)^2 - \frac{1}{12} C^2 \right\}}.$$

Age at onset.

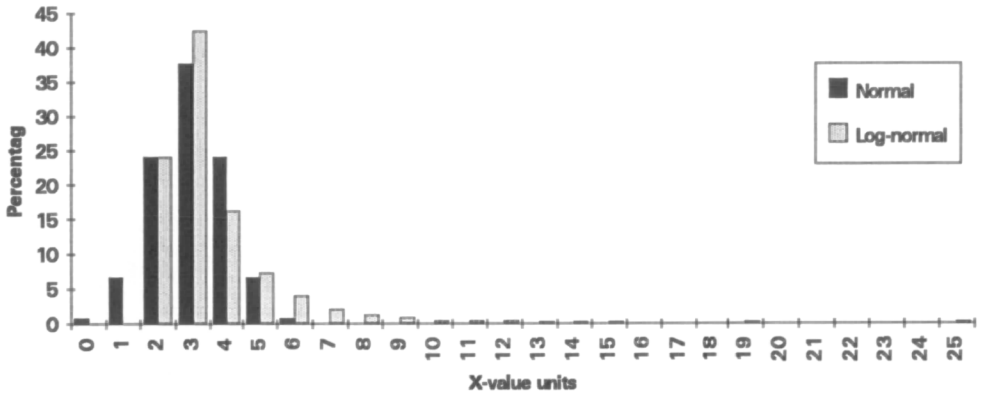
Where appropriate, Sheppard's correction for SD for age at onset was calculated by the same method.

APPENDIX 14

DISCRIMINATORY POWER OF THE TESTS USED TO ESTIMATE COMPATIBILITY WITH NORMALITY AND LOG-NORMALITY.

Figure A14.1 compares normal and log-normal distributions. These were produced by averaging 100 computerised random simulations (using Minitab statistical software, release 7.2, 1989). In each case a sample size of 500 and an SD of 1.0 were used. Modes have been aligned for ease of comparison of the 'tails' of the distributions and the minimum x-value set to zero, thus altering the means from those originally used for sampling (0 for both distributions).

Figure A14.1 Composition of Normal and Log-normal Distributions



Only 93.8% of the log-normal distribution is shown, in order to use the same X-axis scale, as the maximum value was close to 500 units.

Figures A14.2 and A14.3 show the effect of sample size on the appearance, firstly for a normal, and then for a log-normal distribution. The distributions were produced as above (except that only 10 simulations were performed for each sample), altering only the sample size, for comparison of N=500, 50 and 5.

Figure A14.2 The Normal Distribution: Comparison by Sample Size

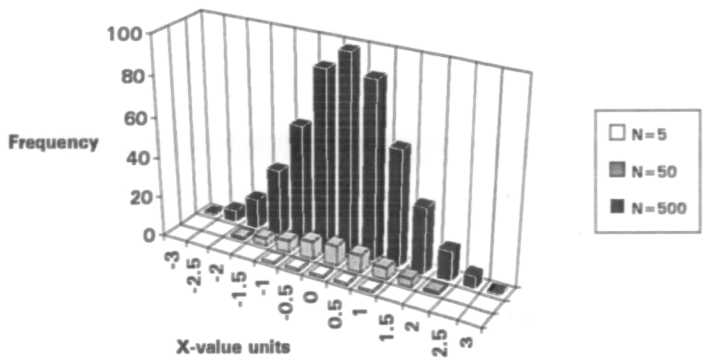
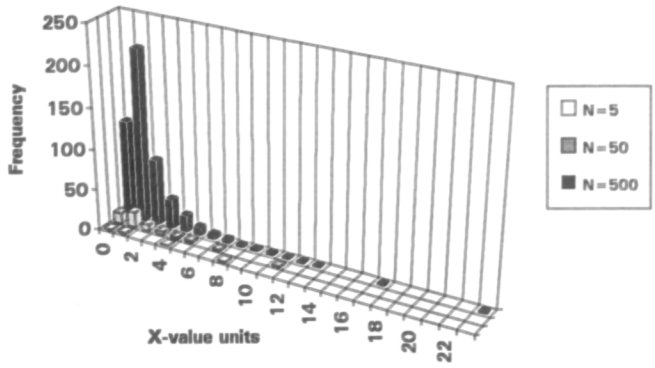


Figure A14.3 The Log-normal Distribution: Comparison by Sample Size



In this thesis, tests for compatibility with normality and log-normality are performed at the 5% (P=0.05) level. If the calculated correlation coefficient (CC) falls below the critical value (C/v), the null hypothesis of normality (or log-normality) is rejected. However, at this level of significance, it will be rejected on one occasion in 20 when it is, in fact, true. If the calculated CC is at or above the C/v, the null hypothesis will not be rejected and the observed distribution is described as compatible with normality (or log-normality) at that level.

Clearly, a variety of non-identical datasets may be classified as compatible with normality (or log-normality) by this test. The question arises as to what extremes of shape will be accepted as normal (or log-normal). The following information attempts to give some pictorial representation of the ranges for three different sample sizes (as used above) broadly representative of the size of datasets examined in the thesis. Firstly, a sample size of N=50 is examined, the most closely representative of these sizes for many of the datasets used. The situation is then compared for N = 5 and 500. C/v's for the CC's (at P=0.05) for samples of these sizes are given in table A14.1.

Table A14.1. Correlation coefficient critical value (at P = 0.05) by sample size.

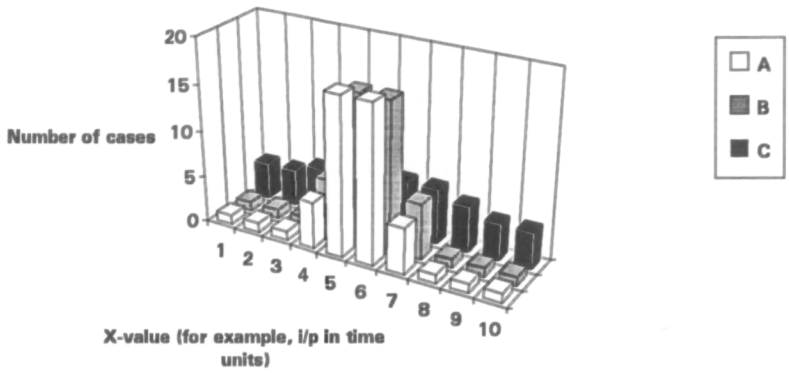
Sample size N	Correlation coefficient critical value
500	0.997
50	0.976 (Filiben) : 0.977 (Wilk-Shapiro)
5	0.880 (Filiben) : 0.879 (Wilk-Shapiro)

Sample size = 50; C/v=0.976.

Normality. Figure A14.4 gives an indication of the range of different distribution shapes which will be accepted by this test as compatible with normality, and shows distributions A, B and C, which are all symmetrical. All were produced by stepwise symmetrical adjustments of a normal

distribution (A and B by piling up 'cases' into the centre of the distribution, C by moving them away from the centre) in an attempt to define the distribution shapes at which non-compatibility with normality (by this test, at this level) was reached. There is little difference in the distribution of A and B, B having two cases (4% of the sample) with X-values different to A and therefore being only slightly flatter than A. C is clearly much flatter, with a very different appearance.

Figure A14.4 N = 50; f/d; normality and non-normality

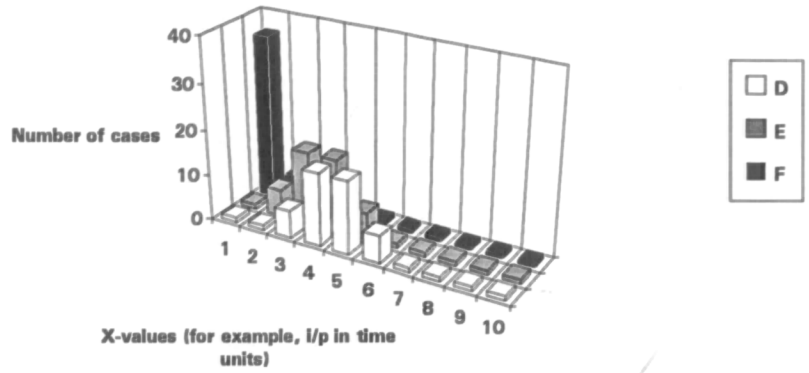


Despite the very small difference in the distribution of A and B, by the test used distribution A is not compatible with normality (CC = 0.975) whereas the slight change puts distribution B at the limit of compatibility (CC = 0.980). Perhaps unexpectedly, distribution C is also compatible with normality (CC = 0.996). In fact, even flattening the distribution completely, with 5 cases at each X-value gives compatibility with normality (CC = 0.992; not shown). A very wide variety of shapes between these two extremes is thus compatible with normality, by this test.

As might be expected, none of these distributions are compatible with log-normality (CC = 0.885 for A, CC = 0.893 for B, and CC=0.954 for C).

Log-normality. Figure A14.5 attempts to give an idea of the skewness compatible with log-normality. Distributions D, E and F are all right skewed. D and E are similar in shape except that sample E has been derived from D by being shifted by one unit leftwards along the X-axis, with a resultant small adjustment (as can be seen) to the lengths of the tails. F appears visually to be very different, and much more right skewed.

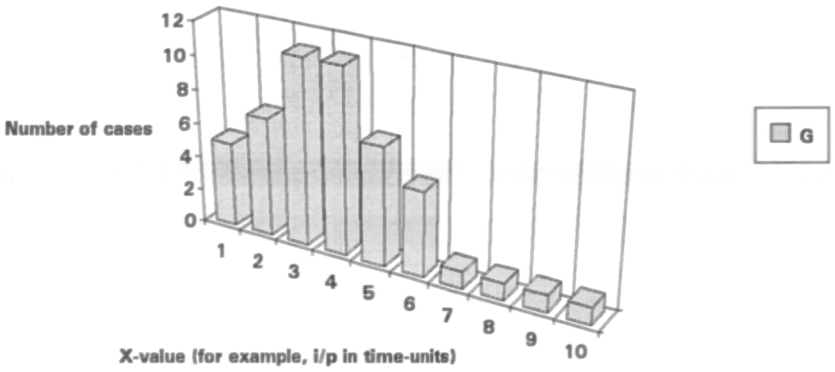
Figure A14.5 N = 50; f/d; log-normality and non-log-normality



Distribution D has a statistically significant right skew (C/s = 0.988; P=0.005), and is not compatible with normality (CC = 0.966); however, neither is it compatible with log-normality (CC = 0.951). The apparently similar distribution E is also significantly right skewed (C/s = 1.622; P<0.001) and also incompatible with normality (CC = 0.935) but the very small increase of right skewness has rendered it compatible with log-normality (CC = 0.987). Despite the very different shape of F (C/s = 2.397; P<0.001) compared with E, conclusions are similar; that is, it is incompatible with normality (CC = 0.914) but compatible with log-normality (CC=0.979). A very wide variety of shapes between these two extremes is thus compatible with log-normality by this test. Neither E nor F look particularly similar to the expected shape of a log-normal curve illustrated in figure A14.1.

Compatibility with both normality and log-normality. Figure A14.6 illustrates distribution G (visually somewhat similar to E and F) which, although statistically significantly right skewed (C/s = 0.863, P=0.012), is compatible with both normality (CC =0.982) and log-normality (CC = 0.980) by this test. Other distributions with similar characteristics, and the resulting inability to discriminate by this test, were found (not shown).

Figure A14.6 N = 50; f/d; compatibility with both normality and log-normality



Reducing the level of significance and hence increasing the C/v may be of some help in further

assessing such samples. For example, at P=0.10, the C/v for N=50 is 0.981; distribution G at this level is compatible with normality, but not log-normality. Hence the hypothesis of log-normality would be rejected at the 10% level of significance, but not at the 5% level, giving weak evidence for incompatibility with log-normality.

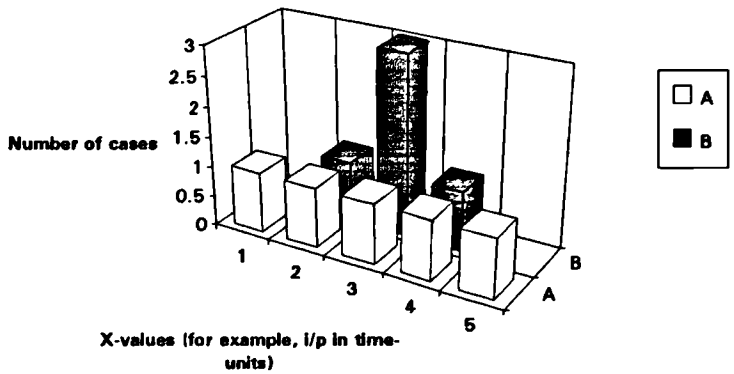
Conclusion. Discrimination between normality and log-normality by this test, for a sample size of 50, whilst often possible is not always so. In addition, there is a wide variety in the distributions which are compatible with normality, and with log-normality. Compatibility with either of the two distributions does not therefore mean a sample appears visually to be, or is, the same as other compatible distributions, nor that it is visually similar to the appropriate theoretical distribution.

<u>Sample size = 5; C/v=0.880.</u>

As might be inferred from figures A14.2 and A14.3, with only 5 cases in a sample, it is very difficult to discriminate between distributions of different shapes.

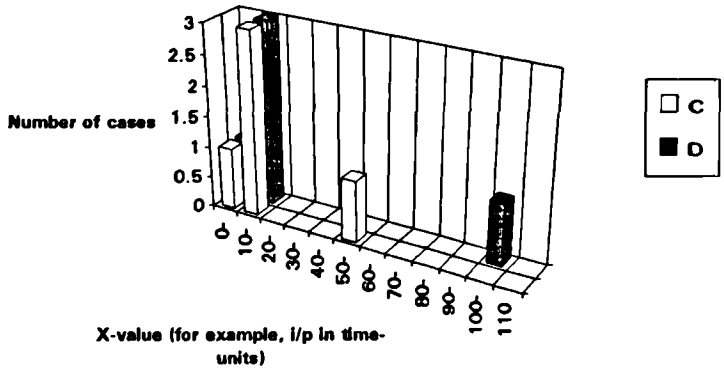
Symmetrical examples. Figure A14.7 shows two symmetrical distributions, A (completely flat) and B (with a marked mode), each of which is compatible both with normality (CC= 0.998 for A; CC=1.000 for B) and log-normality (CC = 0.975 for A; CC=0 980 for B).

Figure A14.7 N = 5; f/d; symmetry, normality and log-normality



Right skewed examples. Figure A14.8 shows two highly right skewed distributions, C (C/s = 1.339) and D (C/s = 1.466), both of which are compatible with log-normality (CC =0.991 for C; CC=1.000 for D). Despite such marked skewness, only D is incompatible with normality (CC=0.898 for C; CC = 0.845 for D).

Figure A14.8 N = 5; f/d; asymmetry, normality and log-normality



For N=5, even at a level of P=0.10, the C/v is only 0.903; the hypothesis that distribution C is compatible with normality would even then only just be rejected .

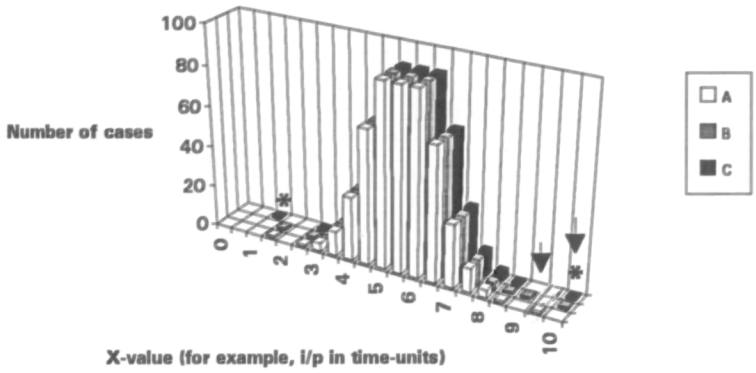
Sample size = 500; C/v=0.997.

For a sample of size 500, results (not all shown) indicate that a wide variety of distributions are compatible with normality, and similarly for log-normality, but that overlap, i.e. compatibility with both, is uncommon. Therefore discrimination by this test, at this level, between the two distributions for N=500 is usually possible.

The C/v for P=0.05, at 0.997, is very high. Nevertheless, there is broad agreement with the findings for a sample of size 50 with regard to the varieties of shapes compatible with normality, and with log-normality. However, since one case represents a much smaller percentage of the total sample, changes from compatibility to non-compatibility (and vice-versa) may be very sensitive to small percentage changes in the distribution.

Example: an X-value change in 0.4% of cases. Figure A14.9 shows three sample distributions, A, B and C which appear to be very similar. However, they are not identical. Distribution A is symmetrical. B has 2 cases with slightly larger X values than A (arrowheads indicate changes), and is thus asymmetrical. C has 1 case with a larger, and one with a smaller X value than A (asterisks indicate changes), and remains symmetrical.

Figure A14.9 N=500; f/d; normality and non-normality

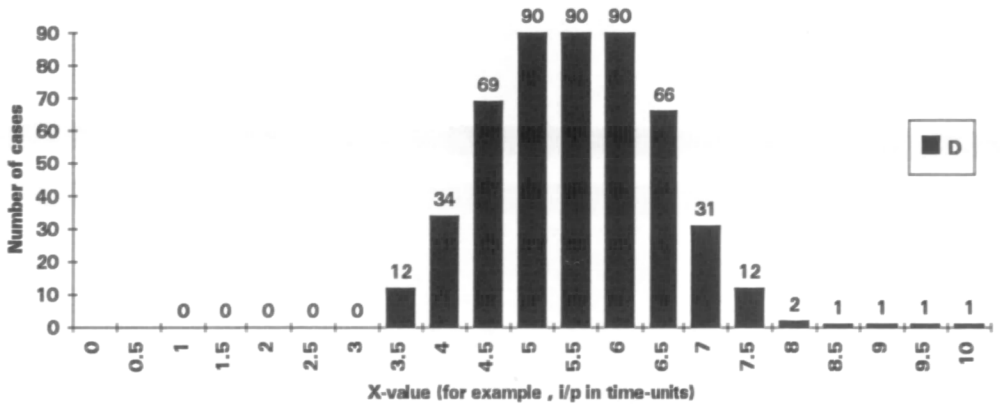


Distribution A, by this test is compatible with normality (CC = 0.998), but not log-normality (CC = 0.979). Distribution B (asymmetrical with a slight right skew; C/s = 0.093; P=0.390) is not compatible with normality (CC = 0.996), nor with log-normality (CC = 0.970). Distribution C (symmetrical, the original range having been symmetrically extended), is no longer compatible with normality (CC = 0.996), nor is it compatible with log-normality (CC = 0.962).

In both B and C, a small change in the X-value of only 2 of the 500 cases (0.4%) has rendered the distribution incompatible with normality, even though, in distribution C, it is still symmetrical. Distribution compatibility with normality is much less sensitive to changes in the X-values of small numbers of cases with x-values near the centre of the distribution (not shown).

Range of distributions compatible with log-normality. Figure A14.10 shows distribution D, made by starting with distribution A and altered by stepwise changes to X-values (i.e. one 'case' moved one unit along, at a time, and recalculated) until it is just compatible with log-normality by this test.

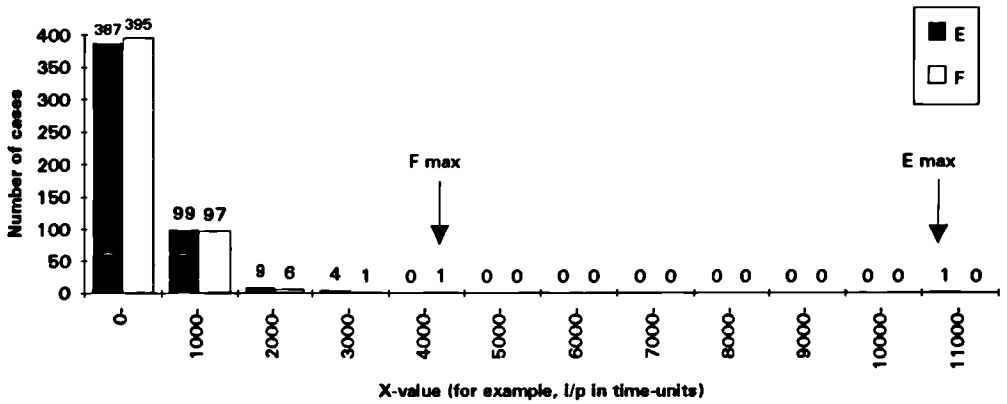
Figure A14.10 N=500; f/d; log-normality (i)



The numbers on the graph show the extent of the asymmetry (C/s = 0.423; P < 0.000). D is, not surprisingly, incompatible with normality (CC= 0.994), but has now reached compatibility with log-normality (CC = 0.997).

Figure A14.11 shows other distributions which are also compatible with log-normality by this test (generated as single samples from computerised random sampling of a log-normal distribution), to give some idea of the extent of variety possible with such a large sample. For the illustration, the X-values have been grouped in thousands, thus the left 'tail' is not apparent.

Figure A14.11 N=500; f/d; log-normality (ii)



Maximum values and skewness for these two distributions E and F are clearly very different, and both are different from D. However, conclusions for all 3, by this test, are similar. As for D, distribution E (C/s 9.703; P<0.000) is incompatible with normality (CC = 0.660) but compatible with log-normality (CC 0.998); similarly for distribution F (C/s 3.442; P<0.001; CC = 0.831 for normality and 0.999 for log-normality respectively).

Summary.

For the test as used in this thesis, the possibility of discrimination is seen to be more likely as sample size increases. The 5% significance level was chosen in order to specifically categorise, with a high level of probability, each dataset as either compatible or not with each of the two distributions in order to look for any recurrent pattern; it is recognised as an artificial and arbitrary cut-off level. For samples of around 50, it will often, but not always, be possible to discriminate at this level. It is not possible, on the basis of the work presented, to give a precise minimum sample size for useful analysis, and further work on this issue is outside the scope of this thesis, but it is suggested that where a distribution is found to be compatible with only one distribution at the 5% level then a useful statement can be made about that distribution. Conversely, when no discrimination is found at this level, evidence for the compatibility of the observed distribution with either theoretical distribution is weakened.

APPENDIX 15

TESTS USED IN THIS THESIS TO ESTIMATE COMPATIBILITY WITH NORMALITY AND LOG-NORMALITY, AND TO CALCULATE COEFFICIENT OF SKEWNESS, APPLIED TO A SELECTION OF THE DISEASES WITH SHORT INCUBATION PERIODS SIMILAR TO THOSE EXAMINED BY SARTWELL, PLUS ONE 'PROXY' INCUBATION PERIOD AS USED BY ARMENIAN AND LILIENFELD (1974).

1) Five outbreaks of 'typhoid fever' reviewed by Miner (1922) and examined by Sartwell (1952).

Appendix A1.2 shows the histograms for the i/p f/d for these five datasets.

a) Outbreak 1, taken from Sawyer (1914); due to contaminated Spaghetti.

Sample size: 93 I/p mean: 9.5 days I/p SD: 5.7 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.904	0.986	1.550	<0.001
Log-normality	0.975	0.986	0.687	0.008

Conclusion from these tests: Compatible with neither normality nor log-normality; very marked right skew. (This was one of the few samples which Sartwell also found to be not log-normal).

b) Outbreak 2; due to contaminated ice-cream.

Sample size: 23 I/p mean: 7.0 days I/p SD: 1.9 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.947	0.958	1.295	0.008
Log-normality	0 980	0.958	0.816	0.070

Conclusion from these tests: Compatible with log-normality, but not normality.

c) Outbreak 3; water-borne.

Sample size: 21 I/p mean: 13.8 days I/p SD: 5.8 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.975	0.952	0.492	0.272
Log-normality	0.972	0.952	-0 546	0.225

Conclusion from these tests: Compatible with both normality and log-normality.

d) Outbreak 4; water-borne.

Sample size: 13 I/p mean: 19.4 days I/p SD: 9.2 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.929	0.930	1.040	0.058
Log-normality	0.971	0.930	0.050	0.923

Conclusion from these tests: Compatible with log-normality but just incompatible with normality.

e) Outbreak 5; water-borne (using mid-point of the possible 4-day water-contamination interval as i/p start).

Sample size: 181 I/p mean: 19.5 days I/p SD: 6.1 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.978	0.991	0.809	<0.001
Log-normality	0.998	0.991	0.068	0.699

Conclusion from these tests: Compatible with log-normality but not normality.

2) Serum hepatitis (viral hepatitis type B); Camp Polk series reported by Parr (1945).

Appendix A1.3 shows the histograms for the i/p f/d for this dataset; the case at 15 days is omitted from the calculation, as it was thought, by Parr, likely to be co-incidental.

Sample size: 1003 I/p mean: 96.6 days I/p SD: 15.9 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.993	0.998	0.356	<0.001
Log-normality	0.998	0.998	-0.031	0.687

Conclusion from these tests: Compatible with log-normality but not normality.

3) Measles; data from Stillerman and Thalheimer (1944).

Appendix A1.1 shows the histograms for the i/p f/d for this dataset.

Sample size: 199 I/p mean: 12.4 days I/p SD: 2.14 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.993	0.993	0.488	0.006
Log-normality	0.999	0.993	0.033	0.844

Conclusion from these tests: Compatible with log-normality; may also just be compatible with normality but results equivocal as right skew present.

4) Measles; data from Goodall (1931), and examined by Sartwell (1952).

Sample size: 115 I/p mean: 12.2 days I/p SD: 3.3 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.983	0.988	0.617	0.008
Log-normality	0.998	0.988	0.010	0.962

Conclusion from these tests: Compatible with log-normality but not normality.

5) Experimental gonorrhoea using human volunteers; data from Mahoney et al (1946), and examined by Sartwell (1966).

Appendix A1.5 shows the histograms for the i/p f/d for this dataset. Isolations reported on day 1 are excluded, as they are considered likely to be the result of residual inoculation material.

Sample size: 71 I/p mean: 5.9 days I/p SD: 5.2 days

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.807	0.982	3.091	<0.001
Log-normality	0.971	0.982	0.898	0.003

Conclusion from these tests: Compatible with neither normality nor log-normality; very marked right skew. (Again, Sartwell also found this to be not log-normal). Similar conclusions result from inclusion of the day 1 isolation cases (not shown).

6) Leukaemia after irradiation for ankylosing spondylitis; data from Cobb, Millar and Wald (1959).

Example of the method used for 'proxy' i/p; here the 'proxy' i/p used is from the mid-point of the radiation exposure period.

Sample size: 51 I/p mean: 51.8 months I/p SD: 30.0 months

F/d under analysis	Correlation Coefficient	Critical value at P= 0.05	Coefficient of skewness (c/s)	P value for c/s
Normality	0.960	0.977	1.155	0.001
Log-normality	0.934	0.977	-1.557	<0.001

Conclusion from these tests: Compatible with neither normality nor log-normality. Using the 'proxy' i/p as from the start of the exposure gave similar results and conclusions (not shown).

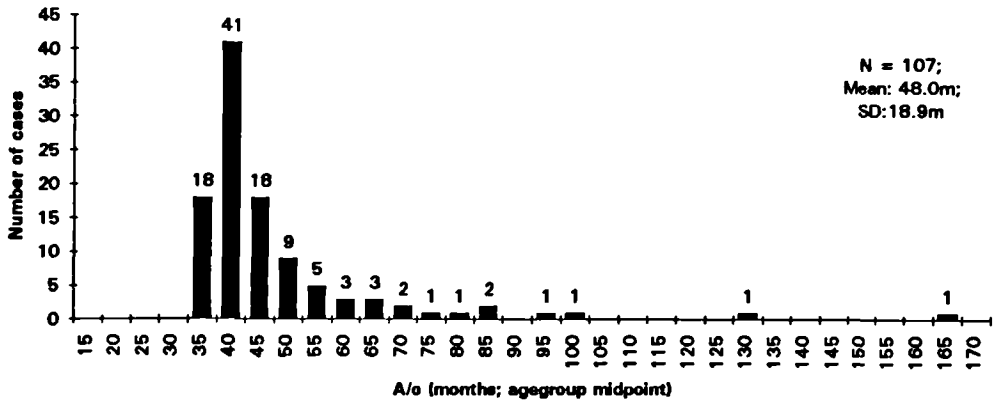
APPENDIX 16

CONSTRUCTED LOG-NORMAL DISTRIBUTION FOR SELECTED DATASETS FOR COMPARISON WITH ACTUAL DATA

The following graphs give an idea of the visual appearance of a 'typical' log-normal distribution for selected TSE/prion disease datasets, for comparison with the actual distribution. The distribution is constructed from the average of 10 repeats of computerised random sampling of a log-normal distribution, using Minitab statistical software (release 7), matching the mean and sample size with that of the original relevant dataset. Not surprisingly, all constructs are broadly similar in shape; the exact SD of each constructed distribution is a result of the random sampling process, dependant upon the actual random numbers selected.

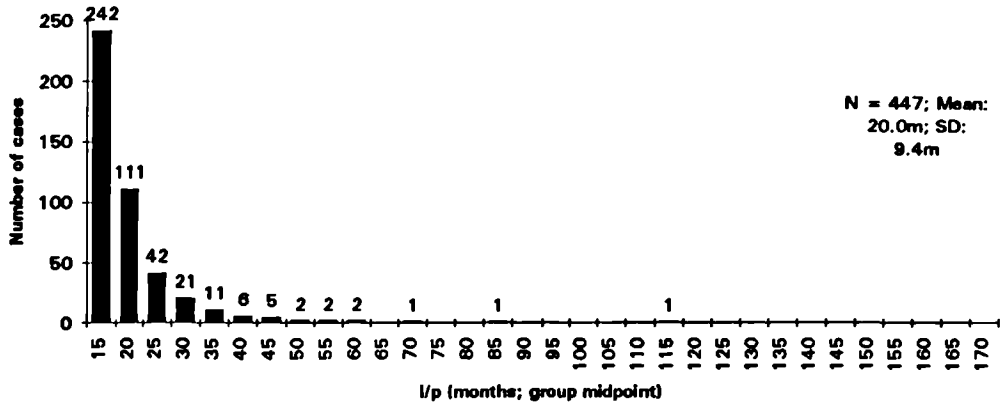
Sheep. Two examples are shown; the first for comparison with figure 6.1.4.1, page 55, and the second with figure 6.2.1.1, page 58.

Figure A16.1 Log-normal construct for natural scrapie in sheep; GLS full dataset



The real dataset on which this construct is based was itself compatible with log-normality, and although the two shapes are not identical, they are broadly similar with respect to overall shape and SD; it is unlikely that sheep would remain alive to demonstrate the largest a/o values in this construct.

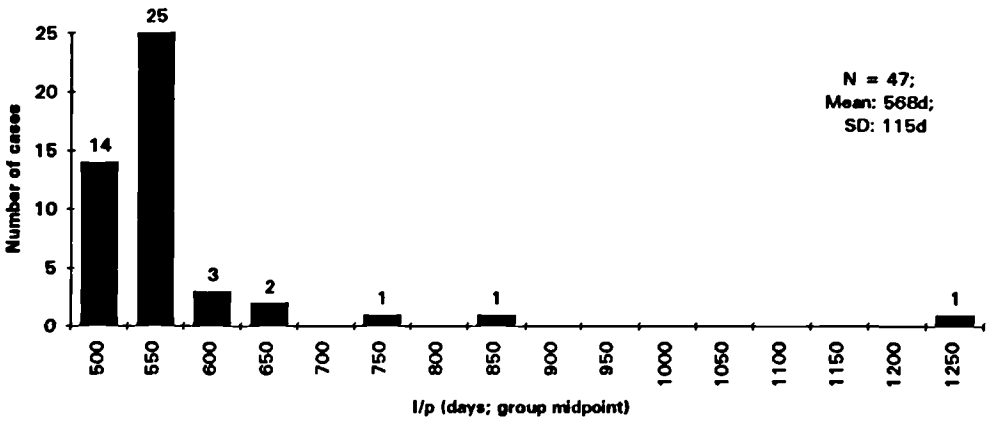
Figure A16.2 Log-normal construct for experimental TSE in sheep; full dataset



The real dataset on which this construct is based was incompatible with log-normality, but not because of the skewness; the comparison with the 'thick' right tail in that dataset is clear.

Rodents. The figure shown is for comparison with figure 9.2.3(a), page 106, the s7p7 genotype.

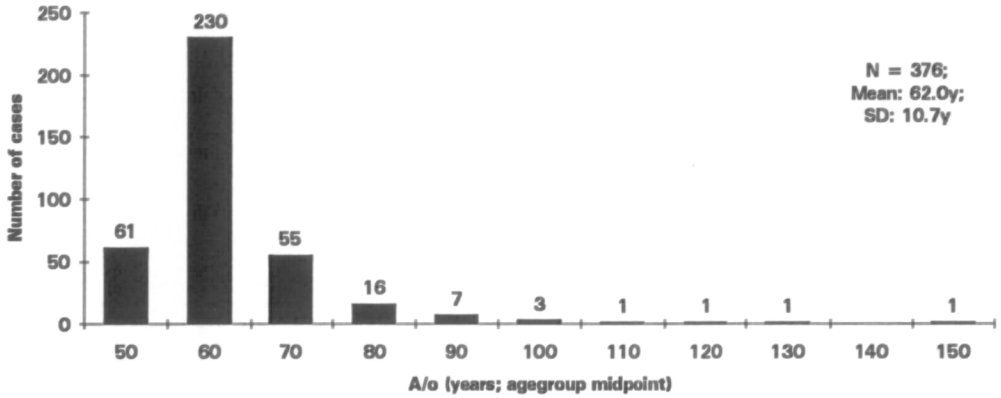
Figure A16.3 Log-normal construct for scrapie in mice; strain 22A in 'sinc' s7p7 mice



The real, small, dataset on which this construct is based, although compatible with both normality and log-normality, had a slight left skew. Clearly a 'typical' log-normal dataset is very different, and, with the same mean, takes much of the range outside the usual lifespan of a mouse, the mean for C57BL being around 2 years, and that for VM possibly less (A Weller, personal communication).

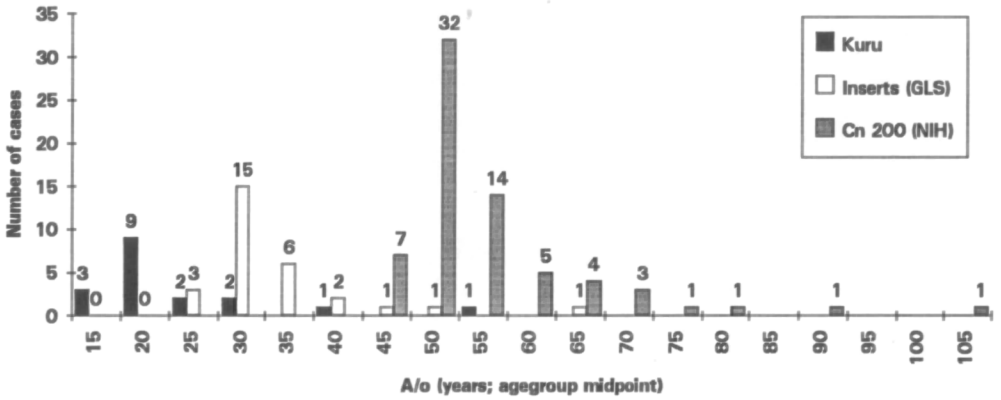
Humans. The first figure is for comparison with figure 10.1, page 120, and the second, a composite, with figures 10.2.2.2 (page 123) and 10.3.1 (page 126).

Figure A16.4 Log-normal construct for TSE disease in humans (ii); sporadic CJD, UK dataset



The real dataset on which this is based has a significant, extended left skew; this construct reverses that skew, giving, for matched means, an age-of-onset range extending considerably outside a typical human lifespan.

Figure A16.5 Log-normal construct for TSE disease in humans (ii); comparison of 3 datasets.



Data for the above constructs.

Dataset	N	Mean a/o	SD
Kuru	18	24.1	10.1
Inserts (GLS)	29	33.7	8.4
Cn 200 (NIH)	69	55.1	10.5

Similarity of these three constructs to the original dataset decreases as the mean a/o increases; for kuru (an infectious manifestation), and for the group with insert mutations (with an uncertain, but possibly purely hereditary aetiology), both groups with a young mean a/o, shape and age-range are broadly similar to the real, right-skewed datasets. For the group with codon 200 mutations however (again with an uncertain, but possibly purely hereditary aetiology), with a much older mean a/o, the original left skew is completely reversed, and the range, even for this small dataset, extends to the maximum of human lifespan.

APPENDIX 17

SELECTION OF DATA FOR ESTIMATING THE INCUBATION PERIOD AND ITS FREQUENCY DISTRIBUTION FOR TA-AIDS.

The data-set as supplied contained many incomplete records, lacking information necessary to calculate an incubation period (i/p) in months. A few records were clearly erroneous, with for example, report dates before diagnosis. The maximum number of transfusions for any case recorded was three. Although some transfusions may appear to be less likely than others to be infectious, with the exception of those transfusions recorded as after the date of diagnosis, there is no theoretical justification for elimination of any of them on grounds of date of transfusion.

Cases with a single transfusion recorded.

Producing a 'clean', complete data-set of cases with a history of only one (presumed infectious) transfusion, reduces numbers from 4010 to 911 (22.7%). These were used for an initial estimate of i/p.

Cases with a traced suspect donor.

Fifty-three cases had a traced suspect donor, 34 of which were single transfusion cases. Where a suspect transfusion was recorded, it was treated as a single transfusion case in all further analyses.

Inclusion of cases with multiple transfusions and incomplete data in overall i/p estimate.

Data was 'cleaned' for all transfusions recorded, and mean time to onset calculated for each transfusion.

With incomplete date recording, one way to maximise utilisable data is to assign missing information in a reasonable way. Table A17.1compares mean time to onset estimated by two methods for each of the three transfusions. The first method uses only the 'cleaned' data, where complete records for all of month and year of transfusion, diagnosis, and report are given. The second estimate for each transfusion includes additional records where year of transfusion, diagnosis and report are known, but the month of any of these may be unknown, and is then assigned as 'mid-year', i.e. July 1st (code 07).

Table A17.1 Time to disease onset: Comparison of 'Cleaned' data for 1st, 2nd, and 3rd transfusion with estimates where unknown month (coded 99) has been replaced by month 07 (i.e. mid-year), for transfusion, diagnosis, and report dates.

	Number N	Mean time (months)	Range (months)	SD (months)
1st (or only) transfusion:				
'Clean' data	2322	57.3	1-264	30.2
'07' substitution data	2433	57.9	1-415	31.4
2nd transfusion:				
'Clean' data	1419	51.9	1-177	27.6
'07' substitution data	1497	52.2	1-415	29.3
3rd transfusion (no substitution needed):				
'Clean' data	8	19.9	8-44	12.0

[For explanation, see text.]

Substitution results in an insignificant increase in mean time to onset for Transfusions 1 and 2 (P=0.677 and 0.986, non-parametric tests, respectively), and a slight increase in SD, due to range extension, as the small number of early transfusions are now included; however, these early transfusions cannot be eliminated with complete certainty on theoretical grounds and the means are effectively unaltered. The 'substituted' data is therefore used in further analyses.

APPENDIX 18

METHOD OF TRANSFUSION-WEIGHTING FOR ESTIMATING INCUBATION PERIOD IN TA-AIDS IN MULTIPLE-TRANSFUSION CASES.

For those cases which have received more than one possibly infectious transfusion, there are several ways of estimating i/p.

Methods considered.

For some transfusions, units of blood are known; it may be that the transfusion with the most units is most likely to be the infecting dose (T Peterman, personal communication); however, complete blood-unit data was available for only eight cases.

Other methods include taking the mid-point between two transfusion dates, or the central transfusion date if there are three; or weighting years of transfusions to account for probable rate of increase and subsequent decrease of numbers of infected donors.

The method using the mid-(or central) point was then examined. Some cases have two transfusions recorded, but a date for only one. Comparing i/p estimate by taking the mid (or central) point for recorded transfusion (Tx) dates for cases with all dates given with that for cases having up to one date missing gives:-

	Cases with all Tx dates given:	Cases with up to one Tx date missing:
Mean i/p (months) :	55.3	55.2
SD:	29.1	29.0
Range:	1-415	1-415
N:	2433	2570

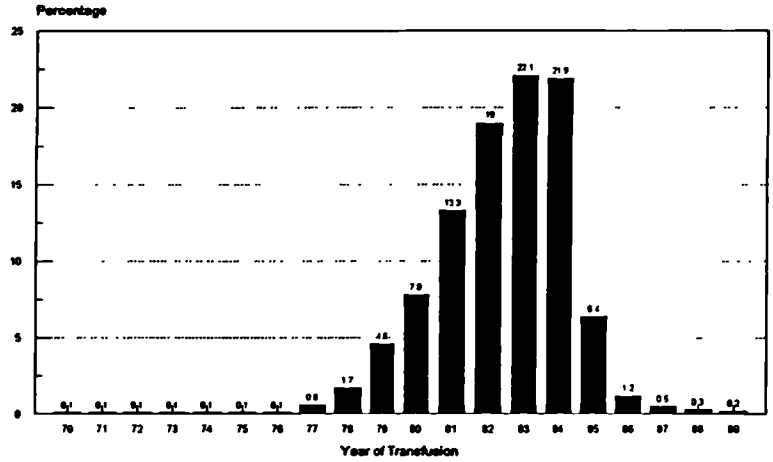
Estimates are almost identical; having up to one transfusion date completely missing (and therefore ignoring that transfusion in calculations) does not alter mean i/p estimate which, not surprisingly, lies between time to onset values for first and second transfusions; the eight cases with three transfusions affects this very little. On the basis of this result, plus the assumption that multiple transfusions are often associated with one medical problem, and therefore often clustered in time, cases with up to one transfusion with year unrecorded were included in further analyses; they were assumed infected in the known year of transfusion.

However, it was considered that weighting the transfusion dates would be likely to give the best approximation to the actual i/p. Many ways of weighting dates of transfusion data are possible; most require incidence data from outside the database. One method of time-weighting from within the database, to assign weighted 'infection date' to multiple transfusion cases is described.

The method finally used.

The method used was to first take only single (or suspect) transfusion cases, and calculate the proportion of transfusions by year; shown in figure A18.1.

Figure A18.1 TA-AIDS; Single transfusion cases by proportion of transfusions per year.



For weighting purposes only, transfusions prior to 1970 and after 1989 were assumed to carry zero risk as their percentage in the single transfusions was zero (unless all transfusions for a particular case were outside those limits; and in such cases all transfusions for each case had occurred in the same year). The decrease in 1985 was assumed to be due mainly to blood screening.

It was assumed that in multiple transfusion cases, the probability of infection from any transfusions occurring in the same year was equal, and dates averaged. However, for multiple transfusions in different years, the probability of infectious transfusion in any particular year was assumed to vary in the same way as that for the single transfusions, and the ratio of these probabilities used as the basis for weighting the i/p between them, appropriate additional weighting being made for those three-transfusion cases with two transfusions in one year.

The final, weighted transfusion date was that used as the infection date for multiple transfusion cases in all further analyses. The result (chapter 5. 1.1) is almost identical to that found by averaging dates; less than 400 cases required weighting out of the total of 2570, and for this set of data, averaging the two (or three) different transfusion dates appears to have been adequate. Since for a given problem, transfusions required are likely to be time-clustered, this may generally be the case.

APPENDIX 19

BACK-CALCULATION METHOD TO ESTIMATE THE MEAN INCUBATION PERIOD OF TA-AIDS.

Calculations performed by R Sayers.

The method used to estimate the mean incubation period was based on that developed by Medley et al (1988a).

Suppose that infection at transfusion occurs as a Poisson process with rate $r(t)$, where t is the time in months measured from 1/1/70, the (unproven) assumption being made that this was the earliest possible infectious transfusion date. Let $f(x)$ be the probability density function of the disease incubation period.

Since the distribution function for the incidence of infectious transfusion decreased sharply from 1/8/85 (most probably due to donor education/blood screening) only cases with infectious transfusion estimated to be before this date (i.e. up to month 180) are used. A further restriction is that case ascertainment was considered to be most reliable between 1/1/85 (month 181) and 31/12/89 (month 240) and therefore only the 724 cases reported between these dates are used.

Then if the observed incubation period and the corresponding times of infection are x_i and t_i respectively, the log likelihood for n cases is given by:-

$$l = \Sigma \log(r(t_i)) + \Sigma \log(f(x_i)) - \int_0^{180} r(t) \int_{180-t}^{240-t} f(x)dxdt$$

The following functions were used for r and f:-

$$\mathrm{r}(t) = \quad \exp(a + bt + ct^2) \text{ for } 0 \le t \le 180$$

$$f(x) = \frac{\alpha^{\beta} x^{\beta-1} \exp(-\beta x)}{\Gamma(\beta)} \quad \text{(Gamma)} \quad or$$

$$\frac{\alpha x^{\alpha-1} \exp(-(x/\beta)^{\alpha})}{\beta^{\alpha}} \quad \text{(Weibull)} \quad or$$

$$\frac{\exp(-(\log x - \alpha)^2 / (2\beta^2))}{\beta x \sqrt{2\Pi}} \quad \text{(Lognormal)}$$

A FORTRAN program then calculated the maximum likelihood estimates of the parameters *a*, *b*, *c*, *m*, α, β using subroutines from the NAG Fortran 77 Library, Mark 14 (NAG Ltd, Oxford, UK).

E04CCF - minimisation by the Simplex method.
D01DAF - double integration.

From these results the mean incubation period was calculated as follows:-

Gamma - $m + \beta/\alpha$

Weibull - $m + \beta\Gamma(1 + 1/\alpha)$

Lognormal - $m + \exp(\alpha + 0.5\beta^2)$

Attempts to estimate confidence limits for the parameters using the profile likelihood were unsuccessful due to problems with convergence in the NAG subroutines.

APPENDIX 20
METHOD FOR ESTIMATING POTENTIAL SCRAPIE CASE NUMBERS ADJUSTED FOR CULLING OF SHEEP IN SCRAPIE INFECTED FLOCKS FROM DATA COLLECTED BY WOOLDRIDGE (1991).

Information abstracted from a scrapie survey (Wooldridge, 1991) of 123 farms, flock sizes ranging from 2 to 2,150 sheep, allows an average culling pattern to be estimated for the intended breeding flock (i.e. sheep over 1 year of age), up to five years of age.

It is assumed that an equal proportion of culled and unculled animals would have developed disease at each age, and that there is no association between culling and incubating scrapie. Calculations are based only on sheep from one year of age; that is sheep intended for the adult flock.

Table A20.1 shows the steps in calculation, and gives the resulting estimate of adjusted case numbers.

Table A20.1 Natural sheep scrapie; Culling pattern estimate (Wooldridge, 1991)

	Age group (years)					
	1-<2	2-<3	3-<4	4-<5	5+	Total
Total 'flock' from survey	10,978	11,612	11,423	8,460	5,852	48,325
% of flock of that age	22.7%	24.0%	23.6%	17.5%	12.1%	100%
Adjusted total for culling	11,338	11,338	11,338	8,460	5,852	48,326
estimate (and % of flock)	23.5%	23.5%	23.5%	17.5%	12.1%	100%
% flock still alive	100%	100%	100%	74.6%	51.6%	-
% flock culled	0%	0%	0%	25.4%	48.4%	-
Actual case numbers	137	662	343	123	68	-
Adjusted case numbers	137	662	343	165	132	-

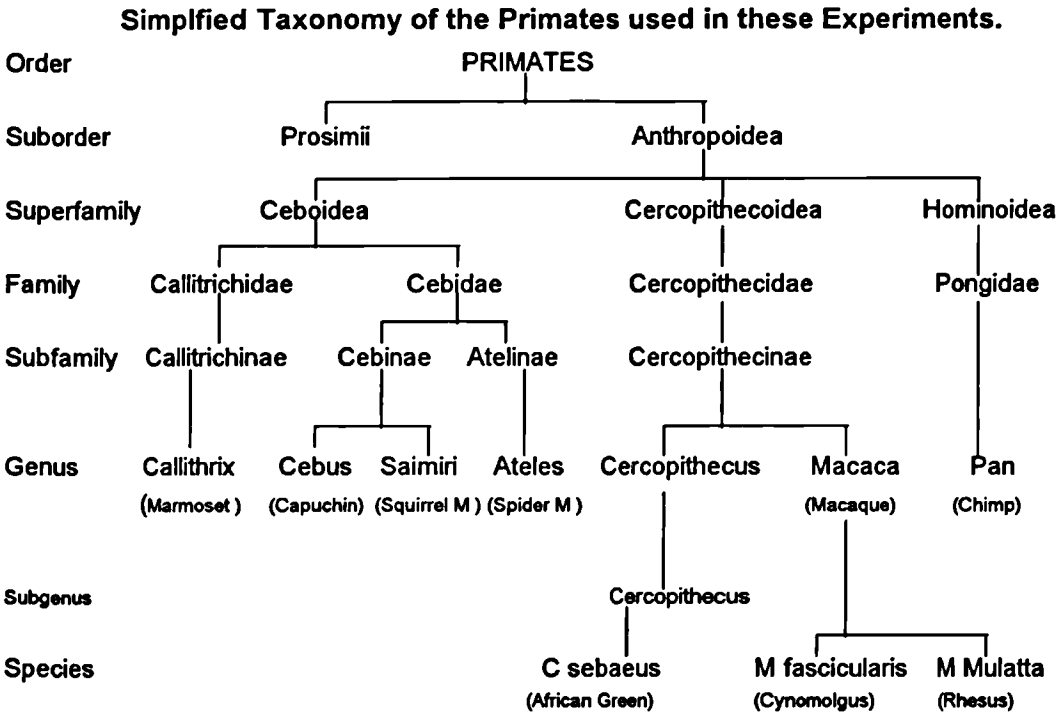
There is little change in the proportion of cases from 4-<5 years (9.2 to 11.5%); it is not possible to estimate the distribution or maximum a/o for cases from 5 onwards, although overall the proportion increases from 5.1% to 9.2%.

APPENDIX 21

PRIMATE TAXONOMY RELEVANT TO DATA CLASSIFICATION USED IN THIS THESIS.

Figure A21.1 shows a simplified scheme of taxonomic classification; names used in the analyses are given in brackets (derived from Napier & Napier, 1967).

Figure A21.1



As shown, marmosets, capuchin, squirrel monkeys, and spider monkeys are all genera within the superfamily ceboidea, whereas the chimpanzee is a genus in the superfamily hominoidea. African green, cynomolgus, and rhesus monkeys are all species, within the superfamily cercopithecoidea.

Ceboidea are the new-world monkeys; cercopithecoidea, the old-world monkeys; hominoidea, the apes and humans. It is generally considered that this order corresponds to levels of primate evolution, i.e. new-world monkeys being lower in the evolutionary scale than old-world monkeys, which are lower than apes. This is based on the assumption that *Homo sapiens* is 'highest' on the evolutionary scale; this point of view may well be species-dependent. Taxonomy does however give some indication of the closeness of the relationships between the different groupings, likely to be reflected in genetic make-up, and therefore possibly also in disease susceptibility and response to infection with TSE's/prion diseases.

APPENDIX 22

PRIMATES: SELECTED NATURAL LIFE HISTORY MEASURES BY GENUS, AND COMPARISON WITH DATA USED IN ANALYSES.

Published longevity data is given by genus. Table A22.1 compares maximum recorded lifespan (Napier & Napier, 1967) for each genus represented in the analysis data with the mean and maximum recorded experimental i/p for the primate group representing that genus in the analysis data.

Table A22.1 Primates; comparison (in months) of recorded maximum longevity with mean and maximum i/p (and as a percentage) for full dataset by genera.

Genus	Maximum recorded longevity	Represented in data by species:	Mean i/p in data	Max i/p in data	Max i/p as % of max longevity
Callithrix (marmosets)	144	u/k	20.7	94	63
Cebus (capuchin)	480	"	43.8	168	35
Saimiri (squirrel m.)	252	"	28.3	189	75
Ateles (spider m.)	240	"	31.4	86	36
Cercopithecus (sub-genus cercopithecus)	372	African green monkey	47.5	58	15
Macaca (macaques)	352	Cynomolgus	55.8	74	21
		Rhesus	62.9	102	29
Pan (chimps)	534	u/k	23.8	86	16

In no genus did maximum observed i/p equal maximum recorded longevity, although this recorded longevity may be much longer than natural mean lifespan, and vary considerably by species within genera.

Table A22.2 gives weight ranges (Napier & Napier, 1967) by genera, and the mid-point of these ranges. There is much variation in weight both for species within genera, and in some species by sex; since data availability varied considerably by genus, these variations are not taken into account in this table, therefore any comparisons or conclusions using these values must be considered a very broad guide.

Table A22.2 Primates; comparison of adult weight with certain i/p measures, for full dataset by genera.

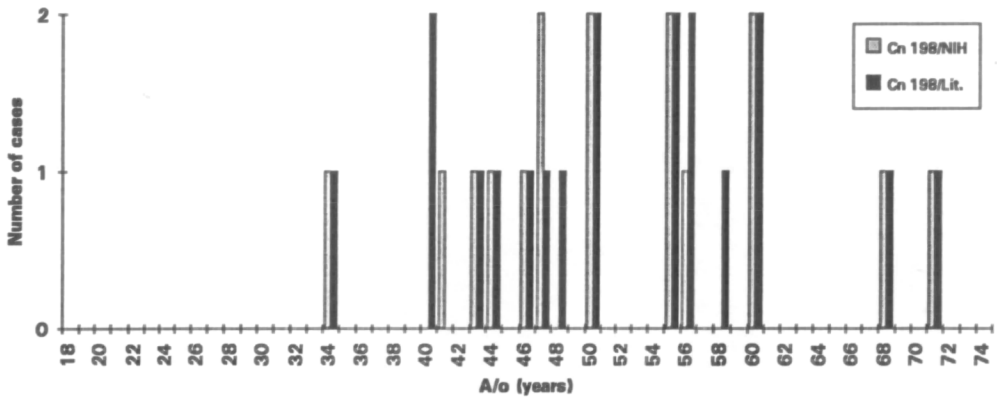
Genus	Approx adult weight Range (kg)	Mid-range (kg)	Represented in data by species:
Callithrix (marmosets)	0.1-0.4	0.25	u/k
Cebus (capuchin)	1.2-3.3	2.75	"
Saimiri (squirrel m.)	0.4-1.1	0.75	"
Ateles (spider m)	5.4-6.9	6.15	"
Cercopithecus (subgenus cercopithecus)	1.8-6.4	4.10	African green
Macaca (macaques)	2.5-18	10.25	Cynomolgus Rhesus
Pan (chimps)	40-50	45	u/k

APPENDIX 23

METHOD OF ESTIMATION OF DATA OVERLAP FOR HUMAN TSE/PRION DISEASE CASES WITH A GENETIC MUTATION, FROM THE TWO DATA SOURCES NIH AND GLS.

Because it is known that many of these cases, being particularly interesting, have been published, there is high probably of considerable dataset overlap. Inspection of the histograms paired by mutation (figure A23.1, example using codon 198 mutation; remainder not shown), shows that age-at-onset (a/o) for many cases duplicates exactly; this may indicate possible degree of overlap.

Figure 23.1 TSE in humand; Codon 198 mutation; comparison of f/d for NIH and GLS data.



Unfortunately, cases with the same a/o cannot be assumed to be the same cases. However, maximum possible overlap would occur if duplicate a/o cases were in fact the same cases. Duplicated a/o numbers and percentages are as follows:-

Dataset	N	Number with duplicated a/o	Percentage with duplicated a/o
NIH	214	60	28.0
GLS	123	60	48.8

Therefore range of possible overlap is 0-28% for NIH data, and 0-49% for GLS data. If all duplicates are genuinely the same cases, amalgamation of the datasets would markedly distort the a/o f/d; conversely if none are the same, then removal of one case from each duplicate pair would again distort results.

Because mean a/o's and the f/d's for each pair of sets are similar, and because they cannot be amalgamated, individual analyses for each mutation employs whichever dataset of each pair is largest.

APPENDIX 24

METHOD OF ADJUSTING DATASET FOR THE EFFECT OF CULLING FOR BSE IN CATTLE AND AS APPLIED TO CATTLE GROUP 3.

An estimate of culling was made, using lactation structure (l/s) recorded for dairy herds on the BSE database.

Data used:

The first recorded l/s for each herd supplying cases in these analyses, recorded as cow numbers per lactation period.

Working assumptions:

1) First lactation starts at two years of age (variation is actually from about 18-30 months); a lactation period is one year. Therefore, a cow recorded in 'lactation 1' is 2-2.99 years old. Reduced cattle numbers between sequential age groups indicates culling.
2) Culling rates are unaffected by subclinical BSE infection (unproven); therefore numbers of animals culled and unculled at any age contain the same proportion of infected animals, of which the proportion which would have developed clinical disease within the next year is the same.
3) All culling per age-group, other than for BSE, occurs prior to any animal developing overt BSE.
4) Numbers culled with BSE are small compared with total culling, therefore the percentage reaching the next age-group is unchanged by cases of BSE. (The validity of this is based on an average, at 5/8/94, of 5.3 total recorded cases per dairy herd since the epidemic began, and an average herd size of 90 cows).

Method:

1) Cattle numbers per age group are totalled over all herds.
2) Total 2 (actually 2-2.99) year old heifers are taken as 100% (the dairy herd comprises only those replacement heifers which actually enter the adult herd).
3) Total numbers at the start of each age group are converted to a percentage of original 2-year olds (equals percentage of original 2-year olds unculled during the previous year-of-age group).
3) Numbers culled during each lactation are converted to a percentage of original 2-year olds.
4) Culling ratio (percentage of original 2-year olds culled during year-of-age/percentage of original 2-year olds unculled by end of year-of-age) is calculated for each age group.

5) Total observed numbers of BSE for each age are multiplied by the culling ratio, and rounded to complete numbers, to give numbers in culled fraction which would have developed BSE at that age.
6) Observed and calculated BSE cases are totalled to give 'unculled' BSE case numbers per age-group; that is the number of cases expected if no culling had occured.

Table A24.1 gives culling pattern, calculated culling ratios, and case number adjustment, applied to group 3 (excluding the one case < 2 years old in group 3).

Table A24.1 BSE in cattle; culling correction, and method as applied to group 3.

Stage in method	Age group 2-	3-	4-	5-	6-	7-	8-	9-	10-	11-	12-	13-
Culling ratio:												
3) % alive: start	100	95.7	84.9	73.6	62.0	47.2	32.8	21.6	13.1	7.4	4.2	1.5
4) % culled in year	4.3	10.8	11.3	11.6	14.8	14.4	11.2	8.5	5.7	3.2	2.7	1.5
3)-4) % alive: end	95.7	84.9	73.6	62.0	47.2	32.8	21.6	13.1	7.4	4.2	1.5	0
5) Ratio culled/unculled	0.05	0.13	0.15	0.19	0.31	0.44	0.52	0.65	0.77	0.76	1.80	-
Group 3 adjustment:												
Observed cases	13	675	2288	1793	719	223	67	18	10	3	0	0
5) Culled 'cases'	1	86	351	335	225	98	35	12	8	2	?	?
6) Total cases	14	761	2639	2128	944	321	102	30	18	5	?	?

APPENDIX 25

BACK-CALCULATION METHOD TO ESTIMATE THE MEAN INCUBATION PERIOD OF BSE.

Calculations performed by R Sayers.

The method used to estimate the mean incubation period was based on that developed by Medley at al (1988a).

Suppose that infection occurs at birth as a Poisson process with rate $r(t)$, where t is the time in months measured from 1/07/81. Also assume that from 31/07/88 (month 85) infection tails off linearly to zero over the 21 months to 30/04/90. Let $f(x)$ be the probability density function of the disease incubation period which equals the age of the animal at onset. The final date of observation is taken as 28/02/94 (month 152) in view of the reporting lag for cases. Animals are assumed to be subject to culling from the age of 24 months with the surviving proportion represented as a function of age by $k(x)$.

Then if the observed ages of onset and the corresponding times of birth are x_i and t_i respectively, the log likelihood for n cases is given by:-

$$l = \Sigma \log(r(t_i)) + \Sigma \log(f(x_i)) + \Sigma \log(k(x_i)) - \int_0^{106} r(t) \int_0^{152-t} f(x)k(x)dxdt$$

The following functions were used for r, k and f:-

$$r(t) = \begin{cases} \exp(a + bt + ct^2) \text{ for } 0 \le t \le 85 \\ \exp(a + 85b + 7225c)\left(\frac{106 - t}{21}\right) \text{ for } 85 < t \le 106 \end{cases}$$

$$k(x) = \begin{cases} 0 \text{ for } 0 \le x \le 24 \\ \exp(-0.000243(24 - x)^2) \text{ for } x > 24 \end{cases}$$

$$f(x) = \quad 0 \text{ for } 0 \le x \le m, \text{ the minimum incubation period.}$$

$$= \left.\begin{array}{ll} \dfrac{\alpha^{\beta} x^{\beta-1} \exp(-\beta x)}{\Gamma(\beta)} & \text{(Gamma)} \quad or \\ \dfrac{\alpha x^{\alpha-1} \exp(-(x/\beta)^{\alpha})}{\beta^{\alpha}} & \text{(Weibull)} \quad or \\ \dfrac{\exp(-(\log x - \alpha)^2 / (2\beta^2))}{\beta x \sqrt{2\Pi}} & \text{(Lognormal)} \end{array}\right\} \quad for \quad x > m$$

A FORTRAN program then calculated the maximum likelihood estimates of the parameters *a*, *b*, *c*, *m*, α, β using subroutines from the NAG Fortran 77 Library, Mark 14 (NAG Ltd, Oxford, UK).

E04CCF - minimisation by the Simplex method.
D01DAF - double integration.

From these results the mean incubation period was calculated as follows:-

Gamma - $m + \beta/\alpha$

Weibull - $m + \beta\Gamma(1 + 1/\alpha)$

Lognormal - $m + \exp(\alpha + 0.5\beta^2)$

Attempts to estimate confidence limits for the parameters using the profile likelihood were unsuccessful due to problems with convergence in the NAG subroutines.

APPENDIX 26

CALCULATION OF THE THEORETICAL COEFFICIENT OF SKEWNESS OF A LOG-NORMAL FREQUENCY DISTRIBUTION USING EXCEL STATISTICAL SOFTWARE.

Theoretical basis; derivation of the expression for calculation of coefficient of skewness.

Derivation of expression for coefficient of skewness (c/s) undertaken by S Cousens.

Log-normal distribution function:-

$$f(x) = \frac{1}{\sigma x\sqrt{2\Pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right)$$

where μ, σ^2 are mean and variance of the normal distribution.

Taking:-

$$E(X^k) = \exp\left(k\mu + \frac{1}{2}\sigma^2 k^2\right)$$

$$E(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

$$E(X^2) = \exp\left(2\mu + 2\sigma^2\right)$$

$$E(X^3) = \exp\left(3\mu + \frac{9\sigma^2}{2}\right)$$

Calculating m_2 and m_3:-

For m_2 :-

$$E(X - E(X))^2 = E(X^2) - [E(X)]^2$$

$$= \exp\left(2\mu + 2\sigma^2\right) - \exp(2\mu + \sigma^2)$$

$$= \exp(2\mu + \sigma^2)[e^{\sigma^2} - 1]$$

For m_3:-

$$E(X-E(X))^3 = E(X^3) - 3E(X^2)E(X) + 3[E(X)]^3 - [E(X)]^3$$

$$= \exp\left(3\mu + \frac{9\sigma^2}{2}\right) - 3\exp\left(2\mu + 2\sigma^2 + \mu + \frac{\sigma^2}{2}\right) + 2\exp\left(3\mu + \frac{3\sigma^2}{2}\right)$$

$$= \exp\left[3\mu + \frac{3\sigma^2}{2}\right]\left[\exp(3\sigma^2) - 3\exp(\sigma^2) + 2\right]$$

Then coefficient of skewness is calculated as follows:-

$$c/s = \frac{m_3}{m_2\sqrt{m_2}}$$

$$= \frac{e^{3\sigma^2} - 3e^{\sigma^2} + 2}{\left[e^{\sigma^2} - 1\right]^{3/2}}$$

Example:-

if SD = 0.05

then σ^2 = $(0.05)^2$

$$c/s = \frac{\left[e^{3(0.05)^2} - 3e^{(0.05)^2} + 2\right]}{\left[e^{(0.05)^2} - 1\right]^{3/2}}$$

which gives c/s = 0.150.

Method using Excel statistical software, version 4.0a (1992).

The calculations were performed in this spreadsheet (methods as described in the Excel 4 0 Users Guide Volumes 1 and 2, 1992) using the following functions:-

SD was entered (cell A1)

Cell A2 = A1^2

Cell A3 = (exp(3*A2)) - (3*(exp(A2))) +2

Cell A4 = (exp(A2) - 1)^(3/2)

Cell A5 = A3/A4 = theoretical c/s

This was used to calculate theoretical c/s for a range of SD values.

APPENDIX 27

METHOD FOR SIMULATION OF SAMPLES FOR ESTIMATION OF COEFFICIENT OF SKEWNESS.

Method for 10-observation samples.

The following method was used to generate the samples used in the first part of section 12.2. Each sample was generated using Minitab statistical software, release 7.2 (1989) on a Prime computer. Ten samples were randomly generated using each set of conditions (i.e. sample size and SD), using the random log-normal command as described in the Minitab reference manual (release 7; 1989). C/s was calculated, as previously described (appendix 12), for each of the ten samples (and plotted, figure 12.2), and the mean and SD of the ten samples per set of conditions calculated (results, table 12.2(i)).

Method for 5000-observation samples.

The following method was used to generate the samples used in the second part of section 12.2, and section 12.3. Essentially, the method was as for 10-observation samples. However, in order to increase the speed of computation, for use with larger numbers of repeat-observations, but in particular with the increase of sample size to 1 million, each sample was generated using a Sun Workstation, utilising a programme specifically written in FORTRAN, by A Mitchell. Random numbers were generated by NAG routines (NAG Fortran Library Manual, Mark 14, 1990), using algorithm GO5DEF. Even on the more powerful Sun, these simulations still took just over one week to run.

Once samples were generated, c/s was calculated, using the calculation method previously described, translated into FORTRAN By A Mitchell, and performed on the Sun, for each of the 5000 samples, and the mean and SD of the 5000 samples per set of conditions calculated (results, table 12.2(ii)). In addition, for each of the two sample sizes (5, and 1 million) the f/d of the 5000 individual c/s values was plotted (figure 12.3), examined for compatibility with normality and log-normality (method, appendix 11) and c/s calculated.

REFERENCES

Aaby P (1991). Determinants of measles mortality: Host or transmission factors. Medical Virology, 10, 83-107.

Aaby P Coovadia H Bukh et al (1985). Severe measles: a reappraisal of the role of nutrition, overcrowding and virus dose. Medical Hypotheses 18, 93-12.

Abbey H (1952). An examination of the Reed-Frost theory of epidemics. Human Biol, 24, 201-233.

Achard Ch Bensaude (1896). Infections paratyphoidiques. Bulletin et Mem de la Soc Med de Hop de Paris 3, Serie 13, 820-833.

Aiken JM Williamson JL Marsh RF (1989). Evidence of mitochondrial involvement in scrapie infection. J Virol 63, 1686-1694.

Aiken JM Marsh RF (1990). The search for scrapie agent nucleic acid. Microbiol Rev 54, 242-246.

Aiken JM Williamson JL Borchardt LM Marsh RF (1990). Presence of mitochondrial D-loop DNA in scrapie-infected brain preparations enriched for the prion protein. J Virol 64, 3265-3268.

Akowitz A Sklaviadis T Manuelidis EE Manuelidis L (1990). Nuclease-resistant polyadenylated RNA's of significant size are detected by PCR in highly purified Creutzfeldt-Jakob disease preparations. Microb Pathog 9, 33-45.

Alcabes P Munoz A Vlahof D Friedland G (1994). Maturity of human immunodeficiency virus infection and incubation period of acquired immunodeficiency syndrome in injecting drug users. Ann Epidemiol 4, 17-26.

Alema G (1973). Aspectos epidemiologicos de la enfermedad de Jakob-Creutzfeldt (consideraciones sobre los casos italianos). Excerpta Medica Int Congress Sciences 274, 1221-1227.

Alper T (1993). The scrapie enigma: Insights from radiation experiments. Rad Res 135, 283-292.

Alper T Haig DA Clarke MC (1966). The exceptionally small size of the scrapie agent. Biochem Biophys Res Comm 22, 278-284.

Alper T Cramp WA Clarke MC (1967). Does the agent of scrapie replicate without nucleic acid? Nature 214, 764-766.

Amano N Yagashita S Yokoi S et al (1992). Gerstmann-Straussler syndrome - A variant type: Amylod plaques and Alzheimer's neurofibrillary tangles in cerebral cortex. Acta Neuropathol 84, 15-23.

Ammann AJ Cowan MJ Wara DW et al (1983). Acquired immunodeficiency in an infant: possible transmission by means of blood products. Lancet, 1, 956-8.

Anderson RM Medley GF May RM Johnson AM (1986). A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV) , the causative agent of AIDS. IMA J of Math Appl Med Biol 3, 229-263.

Anderson RM (1988). The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS. J of Acqu Immune Defic Syndr 1, 241-256.

Anderson RM May RM (1991). In: Infectious diseases of humans; dynamics and control. Oxford University Press.

Andrewes CH (1949). The natural history of the common cold. Lancet 1, 71-75 (Cited in Sartwell, 1950).

Andreani T Modigliani R Le Charpentier Y et al (1983). Acquired immunodeficiency with intestinal cryptosporidiosis: possible transmission by Haitian whole blood. Lancet 1, 1187-1191.

Anon (1990a). BSE case found on the continent. Vet Rec 127, 462.

Anon (1990b). Survey of encephalopathy incidence set up by Health ministry. Vet Rec 126, 447.

Anon (1991). BSE case confirmed in Brittany. Vet Rec 128, 218.

Anon (1994). Government extends ban on specified bovine offals. Vet Rec 135, 26.

Armenian HK (1995). Invited commentary on "The distribution of incubation periods of infectious disease". Am J Epidemiol 141,385.

Armenian HK Khoury MJ (1981). Age at onset of genetic diseases. An application for Sartwell's model of the distribution of incubation periods. Am J Epidemiol, 113, 596-05.

Armenian HK, Lilienfeld AM (1974). The distribution of incubation periods of neoplastic disease. Am J Epidem, 99, 92-100.

Armitage P, Doll R (1961). Stochastic models for carcinogenesis. In: Proceedings of the fourth Berkeley symposium on mathematical statistics and probability. Berkeley and Los Angeles, University of California Press. Vol 4, 19-34. (Cited in Guess & Hoel, 1977)

Auerbach DM Darrow WW Jaffe HW et al (1984). Cluster of cases of the acquired immunodeficiency syndrome: parients linked by sexual contact. Am J Med, 76, 487-92.

Austin AR (1993). Clinical signs of the transmissible spongiform encephalopathies in ruminants. Cattle Practice 1, 63-67.

Bacchetti P Moss AR (1989). Incubation period of AIDS in San Francisco. Nature 338, 251-352.

Bailey NTJ (1975). The mathematical theory of infectious diseases and its applications. 2nd edition. London, Charles Griffin & Co Ltd.

Baker H (1990). Human spongiform encephalopathy. SVJ 44, 19-30.

Baker HF Duchen LW Jacobs JM Ridley RM (1990). Spongiform encephalopathy transmitted experimentally from Creutzfeld-Jakob and familial Gerstmann-Straussler-Scheinker diseases. Brain 113, 1891-1909.

Baker HF Ridley RM Wells GAH (1993). Experimental transmission of BSE and scrapie to the common marmoset. Vet Rec, 132, 403-406.

Baker HF Poulter M Crow et al (1991). Aminoacid polymorphism in human prion protein and age at death in inherited prion disease. Lancet, 337, 1286.

Barlow RM Middleton DJ (1990). Dietary transmission of bovine spongiform encephalopathy to mice. Vet Rec 126, 111-112.

Barre-Sinoussi F Chermann JC Rey F et al (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for aquired immune deficiency syndrome (AIDS). Science, 220, 868-71.

Barton DE (1987). Striking the balance on AIDS. Letter. Nature, 326, 734.

Basler K Oesch B Scott M et al (1996). Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene. Cell 46, 417-428.

Beach SA Dolphin GW (1962). A study of the relationship between X-ray dose delivered to the thyroids of children and the subsequent development of malignant tumours. Phys Med Biol, 6, 583-598.

Beck E Daniel PM Parry HB (1964). Degeneration of the cerebellar and hypothalamo-neurohypophysial systems in sheep with scrapie and its relationship to human system degenerations. Brain, 87, 153-176.

Becker NG (1989). In: Analysis of infectious disease data. Chapman & Hall, London & New York.

Benenson AS (1985) In: Control of Communicable Disease in Man, 14th Edition. Ed: AS Benenson. Washington DC, American Public Health Association.

Bennett AD Birkett CR Bostock CJ (1992). Molecular biology of scrapie-like agents. Rev sci tech Off int Epiz 11(2), 569-603.

Berg LJ (1994). Insights into the role of the immune system in prion disease. Proc Natl Acad Sci USA 91, 429-432.

Bernoulli D (1760). Essai d'une nouvelle analyse de la mortalite causee par la petite verole et des advantages de l'inoculation pour la prevenir. Mem Math Phys Acad Roy Sci., Paris, 1-45.

Besnoit Morel C (1898). Notes sur les lesions neurveuses de la tramblante du mouton. CR Soc Biol, Paris, 5, 536-538.

Blakemore WF (1989). Bovine spongiform encephalopathy and scrapie: Potential human hazards. Outlook on Ariculture 18, 165-168.

Blattner WA (1991). HIV epidemiology: Past, present, and future. FASEB J 5, 2340-2348.

Blaxhult A Granath F Lidman K Giesecke J (1990) . The influence of age on the latency period to AIDS in people infected by HIV through blood transfusions. AIDS 4, 125-129.

Blisard KS Davis LE Harrington MG et al (1990). Pre-mortem diagnosis of Creutzfeldt-Jakob disease by detection of abnormal cerebrospinal fluid proteins. J Neurol Sci 99, 75-81.

Blythe SP Anderson RM (1988). Distributed incubation and infectios periods in models of the transmission dynamics of the human immunodeficiency virus (HIV). IMA J of Math Appl Med Biol 5, 1-19.

Board of Agriculture (1812). Reports. Cited by Stockman (1913).

Bosanquet FD Daniel PM Parry HB (1956). Myopathy in sheep: Its relationship to scrapie and to dermatomyositis and muscular dystrophy. Lancet 2, 737-746.

Bradley R Lowson RC (1992). Bovine spongiform encephalopathy: the history, scientific, political and social issues. In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London. 285-299.

Brookmeyer R (1991). Reconstruction and future trends of the AIDS epidemic in the United States. Science 253, 37-42.

Brookmeyer R Gail MH (1988). A method for obtaining short-term predictions and lower bounds on the size of the AIDS epidemic. J Am Stat Soc 83, 301-308.

Brotherston JG Renwick CC Stamp JT et al (1968). Spread of scrapie by contact to goats and sheep. J Comp Path, 78, 9-17.

Brown F (1993a). Editor. Transmissible spongiform encephalopathies - Impact on animal and human health. Dev Biol Stand, vol 80. Karger, Basel.

Brown P (1993b). Infectious cerebral amyloidoses: Clinical spectrum, risks and remedies. In: Transmissible spongiform encephalopathies - Impact on animal and human health. Ed: F Brown. Dev Biol Stand, vol 80. Karger, Basel. 91-101.

Brown P Cathala F Sadowsky D Gajdusek DC (1979). Creutzfeldt-Jakob disease in France: II. Clinical characteristics of 124 consecutive verified cases during the decade 1968-1977. Ann Neurol 6, 430-437.

Brown P Liberski PP Wolff A Gajdusek DC (1990). Resistance of scrapie infectivity to steam autoclaving after formaldehyde fixation and limited survival after ashing at 360°C: Practical and theoretical implications. J Inf Dis 161, 467-472.

Brown P Goldfarb LG Brown WT et al (1991). Clinical and molecular genetic study of a large German kindred with Gerstmann-Straussler-Scheinker syndrome. Neurology 41, 375-379.

Brown P Gibbs CR Rodgers-Johnson P et al (1994). Human spongiform encephalopathy: the National Institutes of Health series of 300 cases of experimetally transmitted disease. Ann Neurol 35, 513-29.

Bruce ME Dickinson AG (1979). Biological stability of different classes of scrapie agent. In: Slow Transmissible Diseases of the Nervous System: Volume 2. Eds: SB Prusiner & WJ Hadlow. New York, Academic Press. 71-86.

Bruce M (1994). Bovine spongiform encephalopathy: Experimental studies in the United Kingdom. In: Report of the meeting of the OIE ad hoc group on bovine spongiform encephalopathy; Paris, 1-2 September, 1994. Pub: Office International des Epizooties. 29.

Bruce M Chree A McConnell I et al (1994). Transmission of bovine spongiform encephalopathy and scrapie to mice: strain variation and the species barrier. Phil Trans R Soc Lond B 343, 405-411.

Bueler H Aguzzi A Sailer A et al (1993). Mice devoid of PrP are resistant to scrapie. Cell 73, 1339-1347.

Carolan DJP Wells GAH Wilesmith JW (1990). BSE in Oman. Vet Rec, 126, 92.

Centers for Disease Control (1981a). Pneumocystis pneumonia - Los Angeles. Morbidity and Mortality Weekly Report 30, 250.

Centers for Disease Control (1981b). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men - New York City and California. Morbidity and Mortality Weekly Report 30, 305-8.

Centers for Disease Control (1982a). Update on acquired immunodeficiency syndrome (AIDS) - United States. Morbidity and Mortality Weekly Report 31, 507-8, 513-14.

Centers for Disease Control (1982b). Pneumocystis carinii pneumonia among persons with haemophilia A. Morbidity and Mortality Weekly Report 31, 365-7.

Centers for Disease Control (1982c). Opportunistic infections and Kaposi's sarcoma among Haitians in the United States. Morbidity and Mortality Weekly Report 31, 353-4, 360-1.

Centers for Disease Control (1983). Immunodeficiency among female sexual partners of males with acquired immunodeficiency syndrome (AIDS) - New York. Morbidity and Mortality Weekly Report 32; 697-8.

Centers for Disease Control (1984). Update: aquired immunodeficiency syndrome (AIDS) - United States. Morbidity and Mortality Weekly Report 33, 661-4.

Centers for Disease Control (1985a). Update: aquired immunodeficiency syndrome (AIDS) - Europe. Morbidity and Mortality Weekly Report 34, 21-22, 28-31.

Centers for Disease Control (1985b). Update:Prospective evaluation of health-care workers exposed via parenteral or mucous-membrane route to blood or body fluids from patients with acquired immunodeficiency syndrome. Morbidity and Mortality Weekly Report 34, 101.

Centers for Disease Control (1987). Revision of the CDC surveillance definition for acquired immunodeficiency syndrome. Morbidity and Mortality Weekly Report 36 (Supplement 1) 1S-15S.

Chandler RL (1961). Encephalopathy in mice produced by inoculation with scrapie brain material. Lancet 1, 1378-1379.

Chelle PL (1942). Un cas trembante chez la chevre. Bull Acad Vet Fr 15, 249-295.

CJD SU (1994). Creutzfeldt-Jakob disease surveillance in the United Kingdom. Third Annual report, September 1994. The National CJD Surveillance Unit, Western General Hospital, Edinburgh.

Clark AM (1991). Diagnosis of scrapie. Vet Rec, 128, 214.

Clumeck N Sonnet J Tallman H et al.(1984). Acquired immunodeficiency syndrome in African patients. N Engl J Med, 310, 492-7.

Cobb S, Miller M, Wald N (1959). On the estimation of the incubation period in malignant disease: the brief exposure case, leukaemia J Chronic Dis 9, 385-393.

Collee JG (1990). Bovine spongiform encephalopathy. Lancet 336, 1300-1303.

Collinge J Owen F Poulter M et al (1990). Prion dementia without characteristic pathology. Lancet 336, 7-9.

Collinge Palmer MS Dryden AJ (1991) Genetic predisposition to iatrogenic Creutzfeldt-Jakob disease. Lancet 337, 1441-1442.

Collinge J Brown J Hardy J et al (1992). Inherited prion disease with 144 base pair gene insertion. 2. Clinical and pathological features. Brain 115, 687-710.

Collinge J Prusiner SB (1992). Terminology of prion diseases. In: Prion diseases of humans and animals, 5-12. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London.

Collinge J Palmer MS Campbell T et al (1993). Inherited prion disease (PrP lysine 200) in Britain: two case reports. BMJ 306, 301-302.

Collinge J Weissmann C (1994). Editors. Molecular Biology of Prion Diseases. Phil Trans R Soc Lond B. Pub: The Royal Society, London.

Comber T (1772). Real improvements in agriculture (on the principles of A Young Esq.) In: Letters to Reade Peacock Esq. and to Dr Hunter, physician in York, concerning the rickets in sheep. Nicoll, London.

Come JH Fraser PE Lansbury PT (1993). A kinetic model for amyloid formation in the prion diseases: The importance of seeding. Proc Natl Acad Sci USA 90, 5959-5963.

Costagliola D Downs AM (1987). Incubation time for AIDS. Nature, 328, 582.

Court Brown WM Doll R (1965). Mortality from cancer and other causes after radiotherapy for ankylosing spondylitis. BMJ, 2, 1327-1332

Creutzfeldt HG (1920). Z Ges Neurol Psychiatr 57, 1-18.

Crow TJ Poulter M Baker HF et al(1992). Familial dementia in relation to the 144 bp insert and its implications. In: In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London 200-214.

Cuille J Chelle P-L (1936). Pathologie animale: La maladie dite tremblante du mouton est-elle inoculable? CR Acad Sci 203, 1552-1554.

Curran JW (1985). The epidemiology and prevention of the acquired immunodeficiency syndrome. Ann Int Med 103, 657-662.

Curran JW Lawrence DN Jaffe H et al (1984). Acquired immunodeficiency syndrome (AIDS) associated with transfusions. N Engl J Med, 310, 69-75.

Curran JW Morgan WM Hardy AM et al (1985a). The epidemiology of AIDS: current status and future prospects. Science 229, 1352-1357.

Curran JW Jaffe HW Peterman TA Allen JR (1985b). Epidemiologic aspects of acquired immunodeficiency syndrome (AIDS) in the United States. Cases associated with transfusions. Progress Clin Biol Res 182, 259-269.

Dawson M Wells GAH Parker BNJ Scott AC (1990). Primary parenteral transmission of bovine spongiform encephalopathy to the pig. Vet Rec 127, 338.

Day NE Gore SM McGee MA South M (1989). Predictions of the AIDS epidemic in the U.K.: The use of the back projection method. Phil Trans R Soc Lond B 325, 123-134.

DeArmond SJ McKinley MP Barry RA et al (1985). Identification of prion amyloid filaments in scrapie infected brain. Cell 41, 221-235.

DeArmond SJ Prusiner SB (1993). The neurochemistry of prion diseases. J Neurochem 61, 1589-1601.

Denny GO, Wilesmith JW Clements RA Heuston WD (1992). Bovine spongiform encephalopathy in Northern Ireland: epidemiological observations 1988-1990. Vet Rec 130, 113-116.

Deslys J-P Marce D Dormont D (1994). Similar genetic suceptibility in iatrogenic and sporadic Creutzfeldt-Jakob disease. J Gen Viro 75, 23-27.

Detwiler LA (1992). Scrapie. Rev sci tech. Off int epiz. 11(2), 491-537.

Dickinson AG Young GB Stamp JT Renwick CC (1965). An analysis of natural scrapie in Suffolk sheep. Heredity, 20, 485-403.

Dickinson GB Stamp JT Renwick CC (1966). Scrapie: Experiments involving maternal transmission in sheep. Report of Scrapie Seminar; US Dept Agric. ARS 91-53, 244-248.

Dickinson AG Stamp JT Renwick CC Rennie JC(1968). Some factors controlling the incidence of scrapie in Cheviot sheep injected with a Cheviot-passaged scrapie agent. J Comp Path 78, 313-321.

Dickinson AG Mieckle VHM Fraser H (1969). Genetical control of the concentration of ME7 scrapie agent in the brain of mice. J Comp Path 79, 15-22.

Dickinson AG Meickle VMH (1971). Host genotype and agent effects in scrapie incubation: change in allelic interaction with different strains of agent. Molecular and General Genetics 112, 73-79.

Diringer H Oberdieck U Malchow M Beekes M (1993). Transmissible spongiform encephalopathies: Virally transmitted amyloidoses. In: Transmissible spongiform encephalopathies - Impact on animal and human health. Ed: F Brown. Dev Biol Stand, vol 80. Karger, Basel. 25-25.

DoH (1988). Short-term prediction of HIV infection and AIDS in England and Wales. Report of a working group. Dept of Health and Welsh Office. London; HMSO.

Downs AM Ancelle RA Jager C et al (1987). AIDS in Europe: Current trends and short term predictions estimated from surveillance data, January 1981-June 1986. AIDS 1, 53-57.

Doyle JO (1953). Case of leprosy seen in a venereal disease clinic in Britain. BMJ 2, 261-262.

Druckey H (1967). Quantitative aspects of chemical carcinogenesis. In: Potential carcinogenic hazards from drugs (evaluation of risks). Ed: R Truhaut. New York, Springer-Verlag. UICC Monograph Series, Vol 7, 60-78. (Cited in Guess & Hoel, 1977).

Eklund CM Hadlow WJ (1969). Pathogenesis of slow viral diseases. JAVMA 155, 2094-2099.

Eklund CM Kennedy RC Hadlow WJ (1976). Pathogenesis of scrapie virus infection in the mouse. J Inf Dis 117, 15-22.

Englert G Milbradt H (1977). Bovine-type tuberculosis of the kidneys in a farmer: A contribution to the relationship of human and bovine tuberculosis. Dtsch Tierarztl Wschr 84, 51.

Epi Info epidemiological software, release 5.01a; Epi Info reference manual, release 5 (1991). Epidemiology Program Office, Centers for Disease Control, Atlanta, GA 30333, USA.

Erice A Rhame FS Heussner RC et al (1991). Human immunodeficiency virus infection in patients with solid-organ transplants: Report of five cases and review. Rev Inf Dis 13, 537-47.

Evans AS (1982). The clinical illness promotion factor: A third ingredient. Yale J Biol & Med 55, 193-199.

Excel spreadsheet software, version 4.0a, and Excel User's Guides Volumes 1 and 2 (1992). Microsoft Corporation Inc.

Eyster ME Gail MH Ballard JO et al (1987). Natural history of human immunodeficiency virus infectionss in haemophiliacs: Effects of T-cell subsets, platelet counts, and age. Ann Int Med 107, 1-6.

Fanning A Edwards S (1991). *Mycobacterium bovis* infection in human beings in contact with elk *(Cervus elaphus)* in Alberta, Canada. Lancet 338, 1253-1255.

Farlow MR Yee RD Dlouhy SR et al (1989). Gerstmann-Straussler-Scheinker disease. 1. Extending the clinical spectrum. Neurology 39, 1446-1452.

Fenner F (1948a). The epizootic behaviour of mouse-pox (infectious ectromelia). Brit J Exp Path 29, 69-91.

Fenner F (1948b). The pathogenesis of the acute exanthems. Lancet 2, 915-920.

Filliben JJ (1975). The probability plot correlation coefficient test for normality. Technometrics, 17 (1), 111-117. Corrigendum (1975); Technometrics 17(4), 520.

Fischman HR (1981). Multiple sclerosis: A two stage process? Am J Epidem 114, 244-252.

Fleetwood AJ Furley CW (1990). Spongiform encephalopathy in an eland. Vet Rec 126, 408-409.

Foncin J-F Cardot J-L Martinet Y Arnott G (1982). Maladie de Gerstmannn-Straussler-Scheinker: Etude anatomoclinique et genealogique. Rev Neurol (Paris) 138, 123-135.

Foster JD Hope J Fraser H (1993). Transmission of bovine spongiform encephalopathy to sheep and goats. Vet Rec 133, 339-341.

Fraser H McConnell I Wells GAH Dawson M (1988). Transmission of bovine spongiform encephalopathy to mice. Vet Rec 123, 472.

Fraser H Bruce ME Davies D et al (1992). The lymphoreticular system in the pathogenesis of scrapie. In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub. Ellis Horwood. New York, London. 308-317.

Fraser H Pearson GR McConnell I et al (1994). Transmission of feline spongiform encephalopathy to mice. Vet Rec 134, 449.

Friedland GH Klein RS (1987). Transmission of the human immunodeficiency virus. N Engl J Med 317, 1125-1135.

Frosner GG (1988). Estimation of mean incubation time of HIV infected as derived from data of German hemophiliacs [abstract 7832]. In: Programs and abstracts of the Fourth International Conference on AIDS. Stockholm: Swedish Ministry of Health and Social Affairs.

Gail MH Brookmeyer R (1988). Methods for projecting course of acquired immunodeficiency syndrome epidemic. J Natl Cancer Inst 80, 900-911.

Gajdusek DC (1979). Observations on the early history of kuru investigation. In: Slow Transmissible Diseases of the Nervous System, Vol 1. Ed: SB Prusiner Hadlow WJ. Pub: Academic Press, New York & London, 7-36.

Gajdusek DC Zigas V (1959). Degenerative disease of the central nervous system in New Guinea. The endemic occurence of "kuru" in the native population. N Engl J Med 257, 974-978.

Gajdusek DC Gibbs CJ Alpers M (1965). Editors. Slow, Latent and Temperate Virus Infections. NINDB Monograph No 2. Pub: US Dept of Health Education and Welfare.

Gajdusek DC Gibbs CJ Alpers M (1966). Experimental transmission of a kuru-like syndrome in a chimpanzee. Nature 209, 794-796.

Gajdusek DC (1977) Unconventional viruses and the origin and disappearance of kuru. Science 197, 943-960.

Gallo RC Salahuddin SZ Popovic M et al (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science, 24, 500-3.

Galvez S Masters C Gajdusek DC (1980). Descriptive epidemiology of Creutzfeldt-Jakob disease in Chile. Arch Neurol 37, 11-14.

Gazzard B (1990). AIDS: an overview. Balliere's Clin Gastroenterol 4 , 59-289.

Genthon R Gray F Salama J et al (1992). Maladie de Gerstmann-Straussler-Sheinker: Etude pathalogique et genealogique. Rev Neurol (Paris) 148, 335-342.

Gerstmann J Straussler E Sheinker I (1936). Uber eine eigenartige hereditar-familiare Erkrankung des Zentralnervensystems. Z Neurol 154, 736-762.

Gibbons RA Hunter GD (1967). Nature of the scrapie agent Nature (London) 215, 1041-1043.

Gibbs CJ Gajdusek DC Morris JA (1965). Viral charachteristics of the scrapie agent in mice. NINDB Monograph No 2, Slow, Latent and Temperate Virus Infections, 195-202.

Gibbs CJ Gajdusek DC (1969). Science 165, 1023-1025.

Gibbs CJ Gajdusek DC Amyx H (1979). Strain variation in the viruses of Creutzfeldt-Jakob disease and kuru. In: Slow transmissible diseases of the nervous system: Volume 2. Eds: SB Prusiner WJ Hadlow. New York, Academic Press. 87-110.

Gibbs CJ Amyx HL Bacote A et al (1980). Oral transmission of kuru, Creutzfeldt-Jakob disease and scrapie to non-human primates. J Infect Dis 142, 205-208.

Gibbs C Joy A Heffner R et al (1985). Clinical and pathological features and laboratory confirmation of Creutzfeldt-Jakob disease in a recipient of pituitary derived human growth hormone. N Engl Med J 313, 734-739.

Gibbs CJ Safar J Ceroni M et al (1990). Experimental transmission of scrapie to cattle. Lancet 335, 1275.

Giesecke J Scalia-Tomba G Berglund O et al (1988). Incidence of symptoms and AIDS in 146 Swedish haemophiliacs and blood transfusion recipients infected with human immunodeficiency virus. BMJ 297, 99-102.

Goldblatt MW (1949). Vesical tumours induced by chemical compounds. Br J Industrial Medicine, 6, 65-81.

Goldfarb LG Mitrova E Brown P et al (1990). Mutation in codon 200 of scrapie amyloid protein gene in two clusters of Creutzfeldt-Jakob disease in Slovakia. Lancet 336, 514-515.

Goodall EW (1931). Incubation period of measles. (Letter to editor). BMJ 1, 73 (Cited in Sartwell, 1950).

Gordon WS (1960). Scrapie panel. Proc 63rd Ann Meet US Livestock San Assn. 286-294.

Gordon WS (1966). Transmission of scrapie and evidence of spread of infection in sheep at pasture. Report of scrapie seminar; US Dept of Agric. ARS 91(53), 8-18.

Gordon WS Brownlee A Wilson DR (1939). Studies in louping-ill, tick-borne fever and scrapie. Proc III Int Cong Microbiol, New York. 362-363.

Gordon WS Pattison IH (1957). The experimental production of scrapie in goats. Vet Rec 69, 1444.

Greig EDW (1939). Observations on the incubation period in cases of induced malaria. J Trop Med & Hyg 42, 325 (Cited in Sartwell, 1950).

Greig JR (1940a). Scrapie. Trans Highl Agric Soc Scot 52, 71-90.

Greig JR (1940b). Scrapie: Observations on the transmission of the disease by mediate contact. Vet J 96, 203-206.

Greig JR (1950). Scrapie in sheep. J Comp Path 60, 263-266.

Guess HA, Hoel DG (1977). The effect of dose on cancer latency period. J Environ Path Toxicol 1, 279-286.

Hadlow WJ (1959). Scrapie and kuru. Lancet 2, 289-290.

Hadlow WJ (1961). The pathology of experimental scrapie in the dairy goat. Res Vet Sci 2, 289-314.

Hadlow WJ (1990). An overview of scrapie in the United States. JAVMA, 196, 1676-1977.

Hadlow WJ Eklund CM Kennedy RC et al (1974). Course of experimental scrapie virus infection in the goat. J Inf Dis 129, 559-567.

Hadlow WJ Kennedy RC Race RE (1982). Natural infection of Suffolk sheep with scrapie virus. J Infect Dis, 146,657-664.

Harvey PH Clutton-Brock TH (1985). Life history variation in primates. Evolution 39, 559-581.

Hamar WH (1906). Epidemic disease in England. Lancet, 1, 733-9.

Harcourt RA Anderson MA (1974). Naturally occuring scrapie in goats. Vet Rec 94, 504.

Harris C Small CB Klein RS et al (1983). Immunodeficiency in female sexual partners of men with the acquired immunodeficiency syndrome. N Engl J Med, 308, 1181-4.

Hartsough GR Burger D (1965). Encephalopathy of mink. 1. Epizootiological and clinical observations. J Infect Dis 115, 387-392.

Hendriks JCM Medley GF Heisterk SH et al (1992). Short-term predictions of HIV prevalence and AIDS incidence. Epidemiol Infect 109, 149-160.

Hendriks JC Medley GF van-Griensven GJ et al (1993). The treatment-free incubation period of AIDS in a cohort of homosexual men. AIDS 7, 231-9.

Hoinville LJ (1994). Decline in the incidence of BSE in cattle born after the introduction of the 'feed ban'. Vet Rec 134, 274-275.

Hope-Simpson RE (1952). Infectiousness of communicable diseases in the household. Lancet, ii, 733-739.

Horner RD (1987). Age at onset of Alzheimer's disease: Clue to the relative importance of etiologic factors? Am J Epidemiol, 126, 409-414.

Hornick RB Woodward TE (1966). Appraisal of typhoid vaccine in experimentally infected human subjects. Trans Am Clin Climatol Assoc 78, 70-80.

Hourrigan J Klingsporn A McDaniel HA Reimenschneider MN (1969). Natural Scrapie in a goat. JAVMA 154, 538-539.

Hourrigan J Klingsporn A Clark WW de Camp M (1979). Epidemiology of scrapie in the United States. In: Slow Transmissible Diseases of the Nervous System: Volume 1. Eds: SB Prusiner & WJ Hadlow. New York, Academic Press. 331-356.

Hsiao K Baker HF Crow TJ et al (1989). Linkage of a prion protein missense variant to Gerstmann-Straussler syndrome. Nature 338, 342-345.

Hsiao KK Scott M Foster D et al (1990). Spontaneous neurodegenration in transgenic mice with mutant prion protein. Science 250, 1587-1590.

Hsiao K Prusiner SB (1990). Inherited human prion diseases. Neurology 40, 1820-1827.

Hsiao KK Cass C Schellenberg GD et al (1991a). A prion protein variant in a family with the telencephalic form of Gerstmann-Straussler-Scheinker syndrome. Neurology 41, 681-684.

Hsiao K Meiner Z Kahana E et al (1991b). Mutation of the prion protein in Libyan Jews with Creutzfeldt-Jakob disease. New Engl J Med 324, 1091-1097.

Hsiao K Dlouhy SR Farlow MR et al (1992). Mutant prion proteins in Gerstmann-Straussler-Scheinker disease with neurofibrillary tangles. Nature genetics 1, 68-71.

Huckstep RL (1962). In: Typhoid fever and other salmonella infections. Edinburgh & London, Livingstone, Ltd. p211.

Hunter GD (1965). Progress towards the characterisation of the scrapie agent. NINDB Monograph No 2, 259-262.

Hunter GD Millson GC (1964). Studies on the heat stability and chromatographic behaviour of the scrapie agent. J Gen Microbiol, 37, 251-258.

Hunter GD Millson GC (1967). Attempts to release the scrapie agent from tissue debris. J Comp Path 77, 301-307.

Hunter G (1992). The search for the scrapie agent 1961-1981.In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London. 23-29.

Hunter N Foster JD Hope J (1992). Natural scrapie in British sheep: breeds, ages and PrP gene polymorphisms. Vet Rec 130, 389-392.

Hunter N Goldmann W Smith G Hope J (1994a). The association of a codon 136 PrP gene variant with the occurrence of natural scrapie. Arch Virol 137, 171-177.

Hunter N Goldmann W Smith G Hope J (1994b). Frequencies of PrP gene variants in healthy cattle and cattle with BSE in Scotland. Vet rec, 135, 400-403.

Ingraham HS. Unpublished data on an epidemic of milk-borne streptococcal sore throat, Carskill, New York. Personal communication cited by Sartwell, 1950.

Irgens LM (1985). Secular trends in leprosy: Increase in age at onset associated with declining rates and long incubation periods. Int J Leprosy 53, 610-617.

Jakob A (1929). Z Ges Neurol Psychiatr 64, 147-228.

Jaffe HW Darrow WW Echenberg DF et al (1985). The acquired immunodeficiency syndrome in a cohort of homosexual men; a six year follow-up study. Annals Int Med 103, 210-214.

Jarrett JT Lansbury PT (1993). Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? Cell 73, 1055-1058.

Jeffrey M Wells GAH (1988). Spongiform encephalopathy in a nyala (*Tragelaphus angasi*). Vet Pathol 25, 398-399.

Jones HB, Grendon A (1975). Environmental factors in the origin of cancer and estimation of the possible hazard to man. Fd Cosmet Toxicol 13, 251-268. (Cited in Guess & Hoel, 1977).

Kahana E Alter M Braham J Sofer D (1974). Creutzfeldt-Jakob disease: Focus among Libyan Jews in Israel. Science 183, 90-91.

Kalbfleish JD Lawless JF (1988). Estimating the incubation period for AIDS patients. Letter. Nature 333, 504-505.

Kalbfleish JD Lawless JF (1989a). Inference based on retrospective ascertainment: An analysis of the data on transfusion-related AIDS. J Am Stat Assoc 84, 360-372.

Kalbfleish JD Lawless JF (1989b). Estimating the incubation time distribution and expected number of cases of transfusion-associated acquired immune deficiency syndrome. Transfusion 29, 672-676.

Kennaway EL (1957). The incubation period of cancer in man. In: Cancer, Vol 1. Ed: RW Raven. Butterworths, London.

Kimberlin RH (1976). Editor. Slow Virus Diseases of Animals and Man. Pub: North-Holland Publishsing Company, Amsterdam & Oxford, and American Elselvier Publishing Co Inc, New York.

Kimberlin RH Walker CA (1978). Evidence that the transmission of one source of scrapie agent to hamsters involves separation of agent strains from a mixture. J Gen Virol 39, 487-496.

Kimberlin RH (1990). Transmissible spongiform encephalopathies in animals. Can J Vet Res 54, 30-37.

Kimberlin RH (1992). Bovine spongiform encephalopathy. Rev Sci Tech Off Int Epiz, 11, 347-390.

Kitamoto T Iizuka R Tateishi J (1993). An amber mutation of prion protein in Gerstmann-Straussler syndrome with mutant PrP plaques. Biochem Biophys Res Comm 192, 525-531.

Kirkwood BR (1988). Transformations. In: Essentials of medical statistics. Oxford, Blackwell Scientific Publications. 145.

Kirkwood JK Wells GAH Wilesmith JW et al (1990). Spongiform encephalopathy in an arabian oryx (*Oryx leucoryx*) and a greater kudu (*Tragelaphus strepsiceros*). Vet rec 127, 418-420.

Kirkwood JK Wells GAH Cunningham AA et al (1992). Scrapie-like encephalopathy in a greater kudu (Tragelaphus strepsiceros) which had not been fed ruminant derived protein. Vet Rec 130, 365-367.

Kirkwood JK Cunningham AA Wells GAH et al (1993). Spongiform encephalopathy in a herd of greater kudu (*Tragelaphus strepsiceros*): Epidemiological observations. Vet Rec 133, 360-364.

Kirkwood JK Cunningham AA Austin AR et al (1994). Spongiform encephalopathy in a greater kudu (*Tragelaphus strepsiceros*) introduced into an affected group. Vet Rec 134, 167-168.

Kirkwood JK Cunningham AA (1994). Epidemiological observations on spongiform encephalopathies in captive wild animals in the British Isles. Vet Rec 135, 296-303.

Klitzman RL Alpers MP Gajdusek DC (1984). The natural incubation period of kuru and the episodes of transmission in three clusters of patients. Neuroepidemiology 3, 3-20.

Kocisko DA Come JH Priola SA (1994). Cell-free formation of protease-resistant prion protein. Nature 370, 471-474.

Kondo K (1977). The lognormal distribution of the incubation time of exogenous disease; Genetic interpretations and a computer simulation. Jap J Hum Genet 21, 217-237.

Kretzschmar HA (1993). Neuropathology of human prion disease (spongiform encephalopathies). In: Transmissible spongiform encephalopathies - Impact on animal and human health. Ed: F Brown. Dev Biol Stand, vol 80. Karger, Basel. 71-90.

Kretzschmar HA Honold G Seitelberger F et al (1991). Prion protein mutation in family first reported by Gerstmann, Straussler, and Scheinker. Lancet 337, 1160.

Kretzschmar HA Kufer P Reithmuller G et al (1992). Prion protein mutation at codon 102 in an Italian family with Gerstmann-Straussler-Scheinker syndrome. Neurology 42, 809-810.

Latunde Odeku E Osuntokun BO (1968). The clinical neurology of Burkitt's neoplasm: A preliminary evaluation based on 105 cases. W Afr Med J, 17, 263-267.

Lechat MF (1971). An epidemiometric approach for planning and evaluating leprosy control activities. Int J Lepr, 39, 603-7.

Leggett MM Dukes J Pirie HM (1990). A spongiform encephalopathy in a cat. Vet Rec 127, 586-588.

Lemp GF Hessol NA Rutherford GW et al. (1988). Projections of AIDS morbidity and mortality in San Francisco using epidemic models[abstract 4682]. In: Programs and abstracts of the Fourth International Conference on AIDS. Stockholm: Swedish Ministry of Health and Social Affairs cited in Bacchetti & Moss, 1989.

Lifson AR Rutherford GW Jaffe HW (1988). The natural history of human immunodeficiency virus infection. J Inf Dis 158, 1360-1367.

Longini IM Clark WS (1989). Statistical analysis of the stages of HIV infection using a Markov model. Stats in Med 8, 831-843.

Lovestone S (1990). Delusional bovine spongiform encephalopathy. Lancet 336, 565.

Lui K-J Lawrence DN Morgan WM et al (1986). A model-based approach for estimating the mean incubation period of transfusion acquired immunodeficiency syndrome. Proc Natl Acad Sci USA 83, 3051-3055.

Lui K-J Peterman TA Lawrence DN (1987). Comments on the sombre view of AIDS. Letter. Nature, 329, 207.

Lui K-J Peterman TA Lawrence DN Allen JR (1988a). A model-based approach to characterise the incubation period of Paediatric transfusion-associated acquired immunodeficiency syndrome. Stats in Med 7, 359-401.

Lui K-J Darrow WW Rutherfored GW (1988b). A model-based estimate of the of the mean incubation period for AIDS in homosexual men. Science 240, 1333-5.

Mackay JMK Smith W (1961). A case of scrapie in an uninoculated goat - a natural occurence or a contact infection? Vet Rec 73, 394-396.

Mahoney JF van Slyke CJ Cutler JC Blum HL (1946). Experimental gonococcic urethritis in human volunteers. Am J Syph, Gonor and Ven Dis 30, 1-39.

Manuelidis E Manuelidis L (1979). Observations on Creutzfeldt-Jakob disease propagated in small rodents. In: Slow transmissible diseases of the nervous system: Volume 1. Eds: SB Prusiner WJ Hadlow. New York, Academic Press. 147-173.

Manuelidis EE Kim JH Mericangas JR Manuelidis L (1985). Transmission to animals of Creutzfeldt-Jakob disease from human blood. Lancet 2, 896-897.

Marsh RF Burger D Eckroade R et al (1969). A prelliminary report on the experimental host range of the transmissible mink encephalopathy agent. J Inf Dis 120, 713-719.

Marsh RF Hartsough GR (1985). Is there a scrapie-like disease in cattle? Proc US Anim Health Assoc 89, 8-9.

Marsh RF Hadlow WJ (1992). Transmissible mink encephalopathy. Rev sci tech Off int Epiz 11, 539-550.

Martinez-Lage JF Sola J Poza M Esterban JA (1993). Pediatric Creutzfeldt-Jakob disease - probable transmission by a dural graft. Childs Nerv Syst 9, 239-242.

Masden T (1937). Mechanism of bacterial infection. In: Lectures on the epidemiology and control of syphilis, tuberculosis and whooping cough, and other aspects of infectious disease. Baltimore, Williams & Wilkins Co for Vanderbilt University. Chapter 11, 33-83.

Masters CL Gajdusek DC Gibbs JR et al (1979). Familial Creutzfeldt-Jakob disease and other familial dementias: An inquiry into possible modes of transmission of virus-induced familial disease. In: Slow Transmissible Diseases of the Nervous System, Volume 1, 143-194. Eds: SB Prusiner WJ Hadlow. New York, Academic Press.

Masters CL Harris JO Gajdusek DC et al (1979). Creutzfeldt-Jakob disease: Patterns of worldwide occurrence and the significance of familial and sporadic clustering. Ann Neurol 5, 177-18.

Matondo P (1992). Case definition for AIDS surveillance in Africa. BMJ, 304, 54.

Matthews WB (1990). Bovine spongiform encephalopathy. The safety of beef has not yet been tested and may not be testable. BMJ 300, 412.

Mayer V Orolin D Mitrova E (1977). Cluster of Creutzfeldt-Jakob disease and presenile dementia. Lancet 2, 256.

Mayer V Mitrova E Orolin D (1979). Creutzfeldt-Jakob disease in Czechoslovakia and a working concept of its surveillance. In: Slow Transmissible Diseases of the Nervous System: Volume 1, 288-303. Eds: SB Prusiner WJ Hadlow. New York, Academic Press.

McEvoy M Tillet HE (1985). Some problems on the prediction of future numbers of cases of the acquired immunodeficiency syndrome in the UK. Lancet 2, 541-542.

McKinley MP Bolton Dc Prusiner SB (1993). A protease-resistant protein is a structural component of the scrapie prion. Cell 35, 57-62.

Medley GF Anderson RM Cox DR Billard L (1987). Incubation period of AIDS in patients infected via blood transfusion. Nature 328, 719-721.

Medley GF Billard L Cox DR Anderson RM (1988a). The distribution of the incubation period for the acquired immunodeficiency syndrome (AIDS). Proc R Soc Lond B 233, 367-377.

Medley GF Anderson RM Cox DR Billard L (1988b). Estimating the incubation period for AIDS. Nature 333, 505.

Medori R Tritschler H-J LeBlanc A et al. (1992). Fatal familial insomnia, a prion disease with a mutation at codon 178 of the prion protein gene. N Engl J Med 326, 444-449.

Meynell GG, Meynell EW (1958). The growth of micro-organisms in vivo with particular reference to the relation between dose and latent period. J Hyg Camb 56, 323-346.

M'Fadyean J (1918). Scrapie. J Comp Path 31, 102-131.

Millson GC Hunter GD Kimberlin RK (1976). The physico-chemical nature of the scrapie agent. In: Slow Virus Diseases of Animals and Man. Ed: RH Kimberlin Pub: North-Holland Publishsing Company, Amsterdam & Oxford, and American Elselvier Publishing Co Inc, New York. 243-266.

Miner JR (1922). The incubation period of typhoid fever. J Infect Dis 31, 296-301.

Minitab statistical software, release 7.2; Minitab reference manual, release 7 (1989). Minitab Inc., 3081, Enterprise drive, State college, PA 16801-2756, USA.

Mitchell B Stamp JT (1983). Scrapie. In: Diseases of sheep, 71-75. Ed: WB Martin Pub: Blackwell's Scientific, Oxford.

Monari L Chen SG Brown P et al (1994). Fatal familial insomnia and familial Creutzfeldt-Jakob disease: Different prion proteins determined by a DNA polymorphism. Proc Natl Acad Sci USA 91, 2839-2842.

Mosher WE Wheeler SM Chant HL Hardy AV (1941). Studies of the acute diarrheal diseases. V. An outbreak due to *Salmonella typhi murium*. Pub Health Rep 56, 2415. (Cited in Sartwell, 1950).

Munoz A Wang M-C Bass S et al (1989). Aquired immunodeficiency syndrome (AIDS)-free time after human immunodeficiency virus type 1 (HIV-1) seroconversion in homosexual men. Am J Epidem 130, 530-539.

Murdoch GH Sklaviadis T Manuelidis EE Manuelidis L (1990). Potential retroviral RNAs in Creutzfeldt-Jakob disease. J Virol 64, 1477-1486.

Music SI Libonati JP Wenzel RP et al (1970). Induced human cholera. Antimicrobial Agents and Chemotherapy (United States) 10, 462-466.

NAG Fortran Library Manual, Mark 14 (1990). The Numerical Algorithms Group Ltd., Wilkinson House, Jordan Hill Road, Oxford, OX2 8DR, UK.

Nakazato Y Hirato J Ishida Y et al (1990). Swollen cortical neurons in CJD contain a phosphorylated neurofilament epitope. J Neuropath Exp Neurol 49, 197-205.

Nakazato Y Ohno R Negishi T et al (1991). An autopsy of Gerstmann-Straussler-Sheinker's disease with spastic paraplegia as its principal feature. Clin Neurol 31, 987-992.

Napier JR Napier PH (1967). The handbook of living primates. London, Academic Press.

Narang HK (1990). Detection of single-stranded DNA in scrapie-infected brain by electron-microscopy. J Mol Biol 216, 469-473.

Narang HK (1992) Evidence that the scrapie agent contains a single-stranded DNA (ssDNA) genome. Clin Neuropathol 11, 229.

Narang HK (1993). Evidence that scrapie-associated tubulofilamentous particles contain a single-stranded DNA. Intervirology 36, 1-10.

Narang HK Millar NS Asher DM Gajdusek DC (1991). Increased multimeric mitochondrial DNA in the brain of scrapie-infected hamsters. Intervirology 32, 316-324.

National Institute of Health (1936). Bulletin no. 16. Epidemic amebic dysentery: The Chicago outbreak of 1933. Washington, Govenment Printing Office. (Cited in Sartwell, 1950).

Neugut RH Neugut AI Kahana E et al (1979). Creutzfeldt-Jakob disease: Familial clustering among Libyan-born Israelis. Neurology 29, 225-231.

Newton E Farley J Gayle C (1993). Back-projection of the HIV/AIDS epidemic in the Caribbean. Int Conf AIDS (Germany), June 6-11, 1993 [abstract PO-C33-3332]; 9 (2), 772.

Nisipeanu P El Ad B Korczyn AD (1990). Spongiform encephalopathy in an Israeli born to immigrants from Libya. Lancet 336, 686 .

Nochlin D Sumi SM Bird TD et al (1989). Familial dementia with PrP-positive amyloid plaques: A variant of Gerstmann-Straussler syndrome. Neurology 39, 910-918.

Oesch B Westaway D Walchli M et al (1985). A cellular gene encodes scrapie PrP 27-30 protein. Cell 40, 735-746.

OIE (1992). Scientific and technical review, volume 11, No 2. Transmissible spongiform encephalopathies of animals. Eds: R Bradley D Matthews. Pub: Office International des Epizooties, Paris.

OIE (1994). Report of the meeting of the OIE ad hoc group on bovine spongiform encephalopathy; Paris, 1-2 September, 1994. Pub: Office International des Epizooties, Paris.

Order (1988). The Bovine Spongiform Encephalopathy Order 1988. Statutory Instrument 1988, No 1039. HMSO, London.

Orskov J (1932). Der bacterielle infectionmechanismus. Acta path microbiol scand. suppl 11.

Outram GW (1976). The pathogenesis of scrapie in mice. In: Slow Virus Diseases of Animals and Man. Ed: RH Kimberlin. Pub: North-Holland Publishing Company, Amsterdam & Oxford. 325-357.

Palmer AC (1959) Scrapie: A reveiw of the literature. Vet Rev Annot 5, 1-15.

Palmer MS Dryden AJ Hughes JT Collinge J (1991). Homozygous prion protein genotype predisposes to sporadic Creutzfeldt-Jakob disease. Nature 352, 340-342.

Palsson PA (1979). Rida (scrapie) in Iceland and its epidemiology. In: Slow transmissible diseases of the nervous system, Volume 1, 357-366. Eds: SB Prusiner WJ Hadlow. Pub: New York, Academic Press.

Pan K-H Baldwin M Nguyen et al (1993). Conversion of α–helices into β–sheets features in the formation of the scrapie prion proteins. Proc Natl Acad Sci USA 90, 10962-10966.

Pape JW Liautaud B Thomas F et al (1983). Characteristics of the acquired immunodeficiency syndrome (AIDS) in Haiti. N Engl J Med, 309, 945-50.

Parr LW (1945). Host variation in the manifestation of disease: with particular reference to homologous serum jaundice in the army of the United States. Med Annals of the District of Columbia 14, 443-449.

Parry HB (1962). Scrapie: A transmissible and hereditary disease of sheep. Heredity 17, 75-105.

Parry HB (1983). Scrapie disease in sheep. Ed: DR Oppenheimer. London, Academic Press.

Pattison IH (1957). Myopathy in sheep. Lancet 1, 104.

Pattison IH (1964). The spread of scrapie by contact between affected and healthy sheep, goats, or mice. Vet rec 76, 333-336.

Pattison IH (1965a). Experiments with scrapie with special reference to the nature of the agent and the pathology of the disease. In: Slow, latent and temperate virus infections. Eds: Gajdusek DC Gibbs CJ Alpers M. Pub: NINDB Monograph 2. 249-257.

Pattison IH (1965b). Scrapie in Welsh Mountain breed of sheep and its experimental transmission to goats. Vet Rec 77, 1388-1390.

Pattison IH (1965c). Resistance of the scrapie agent to formalin. J Comp Path 75, 159 - 164.

Pattison IH (1974).Scrapie in sheep selectively bred for high susceptibility. Nature 248, 594 - 595.

Pattison IH (1988). Fifty years with scrapie: A personal reminiscence. Vet Rec 123, 661-666.

Pattison IH (1992). A sideways look at the scrapie saga: 1732-1991. In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London. 15-22.

Pattison IH Gordon WS Millson GC (1959). Experimntal production of scrapie in goats. J Comp Path 69, 300-312.

Pattison IH Millson GC (1960). Further observations on the experimental production of scrapie in goats and sheep. J Comp Path 70, 182-193.

Pattison IH Millson GC (1961a). Experimental transmission of scrapie to goats and sheep by the oral route. J Comp Path 71, 171-176.

Pattison IH Millson GC (1961b). Further experimental observations on scrapie. J Comp Path 71, 350-359.

Pattison IH Millson GC (1961c). Scrapie produced experimentally in goats with special reference to the clinical syndrome. J Comp Path 71, 101-108.

Pattison IH Millson GC (1962). Distribution of scrapie agent in the tissues of experimentaly inoculated goats. J Comp Path 72, 233-244.

Pattison IH Hoare MN Jebbett JN Watson WA (1972). Spread of scrapie to sheep and goats by oral dosing with foetal membranes from scrapie-affected sheep. Vet Rec 90, 465-467.

Pattison IH Hoare MN Jebbett JN Watson WA (1974). Further observations on the production of scrapie in sheep by oral dosing with foetal membranes from scrapie-affected sheep. Br Vet J 130, 65-67.

Pavord T Fisher R (1987). Sympathetic and parasympathetic system diseases; grass disease. In: The Equine veterinary manual. The Crowood Press, Ramsbury, Wiltshire. 142-143.

Pearlman RL Towfighi J Pezeshkpour GH et al (1988). Clinical significance of types of cerebellar amyloid plaques in human spongiform encephalopathies. Neurology 38, 1249-1254.

Pearson GR Wyatt JM Henderson JP Gruffydd-Jones TJ (1993). Feline spongiform encephalopathy: A review. Vet Annual 33, 1-10.

Peterman TA (1987). Transfusion -associated acquired immunodeficiency syndrome. World J Surg 11, 36-40.

Peterman TA Drotman DP Curran JW (1985). Epidemiology of the acquired immunodeficiency syndrome (AIDS). Epidemiol Rev 7, 1-21.

Peterson DA Wolfe LG Deinhardt FW (1978). Human spongiform encephalopathies in marmoset monkies (*Saguinus* sp.) Primates in medicine 10, 254-260.

Phillipe P (1990). Twinning causative origin investigated by Sartwell's biometrical method. Am J Human Biology 2, 107-115.

Piot P Quinn TC Taelman Het al (1984). Acquired immunodeficiency syndrome in a heterosexual population in Zaire. Lancet 2, 65-9.

Pitchenic AE Fischl MA Dickinson GM et al (1983). Opportunistic infections and Kaposi's sarcoma among Haitians: evidence of a new acquired immunodeficiency state. Ann Intern Med 98, 277-84.

Polak BCP Wesseling H Schut D et al (1972). Blood dyscrasias attributed to chloramphenicol. Acta Med Scand 192, 409-414.

Prusiner SB Hadlow WJ (1979). Editors. Slow Transmissible Diseases of the Nervous System, Volumes 1 and 2. Pub: Academic Press, New York & London.

Prusiner SB (1982). Novel proteinaceous infectious particles cause scrapie. Science 216, 136-146.

Prusiner SB (1987). The biology of prion transmission and replication. In: Prions: Novel infectious pathogens causing scrapie and Creutzfeldt-Jakob disease. Eds: SB Prusiner M McKinley. Academic Press inc, New York, London. 83-112.

Prusiner SB (1989). Scrapie prions. Annu Rev Microbiol 43, 345-74.

Prusiner SB (1993). Genetic and infectious prion disease. Arch Neurol 50, 1129-1153.

Prusiner SB (1994). Inherited prion disease. Proc Natl Acad Sci USA 91, 4611-4614.

Prusiner SB (1995). The prion diseases. Sci Am 272, 30-37.

Prusiner SB Scott M Foster D et al (1990). Transgenic studies implicate interactions between homologous PrP isoforms in scrapie prion replication. Cell 63, 673-686.

Prusiner SB Westaway W (1991). Infectious and genetic manifestations of prion disease. Molecular plant-microbe interactions 4, 226-233.

Prusiner SB Collinge J Powell J Anderton B (1992). Editors Prion Diseases of Humans and animals. Pub: Ellis Horwood, New York, London.

Prusiner SB DeArmond SJ (1994). Prion diseases and neurodegeneration. Ann Rev Neurosci, 17, 311-339.

Quinn TC Mann JM Curran JW Piot P (1986). AIDS in Africa: An epidemiological paradigm. Science 234, 955-963.

Rees M (1987). The sombre view of AIDS. Nature 326, 343-345.

ReVelle C Lynn WR Feldmann F (1967). Mathematical models for the economic allocation of tuberculosis control activities in developing countries. Amer Rev Resp Dis 96, 893-909.

Ridley RM Frith CD Conneally PM (1992). Non-Mendelian features of the inheritance of Huntington's disease: a possible example of genomic imprinting. Seminars in Developmental Biology, 3 127-137.

Ridley RM Baker HF (1993). Genetics of human prion disease. In: Transmissible spongiform encephalopathies - Impact on animal and human health. Ed: F Brown. Dev Biol Stand, vol 80. Karger, Basel. 15-23.

Roberts GW Collinge J (1990). Bovine spongiform encephalopathy. BMJ 300, 943-944.

Robinson NJ Auvert B Mulder DW et al (1993). Unravelling factors that critically influence observed HIV/STD associations. Int Conf AIDS (Germany), June 6-11, 1993 [abstract PO-C03-2605]; 9 (2), 651.

Roche-Lubin M (1848). Memoire pratique sur la maladie connue sous les noms de prurigo lombaire, convulsive, trembleuse, tremblante. Rec Med Vet 25, 698-714.

Rodrigues L (1991). Epidemiology and control of communicable diseases. MSc course notes, LSH&TM, London.

Rogers MF Ewing EP Warfield D et al (1986). Virologic studies of HTLV-III/LAV in pregnancy: Case report of a woman with AIDS. Obstet gynecol 68 (3 Suppl), 2S-6S.

Rogers MF Thomas PA Starcher ETet al (1987). Aquired immunodeficiency syndrome in children: Report of the Centers for Disease Control National Surveillance, 1982 to 1985. Paediatrics 79, 1008-1014.

Ross R (1915). Some a priori pathometric equations. BMJ 1, 546-7.

Ross R (1916). An application of the theory of probabilities to the study of a priori pathometry, I. Proc Roy Soc A 92, 204-30.

Ross R (1917). An application of the theory of probabilities to the study of a priori pathometry, II. Proc Roy Soc A 93, 212-25.

Royston JP (1985). Algorithm AS209. The distribution function of skewness and kurtosis. Appl Statist 34, 87-94.

Sacher GA Staffeldt EF (1974). Relation of gestation time to brain weight for placental mammals. Amer Natur 108, 593-616.

Sack RB Carpenter CCJ (1968). Experimental Canine Cholera. I. Development of the model. J Infect Dis 119, 138-149.

Sadeh M Chagnac Y Goldhammer Y (1990). Cteutzfeldt-Jakob disease associated with peripheral neuropathy. Israeli J Med Sci 26, 220-222.

Safar J Wang W Padgett MP et al (1990). Molecular mass, biochemical composition, and physicochemical behaviour of the infectious form of the scrapie precursor protein monomer. Proc Natl Acad Sci 87, 6373-6377.

Sartwell PE (1950). The distribution of the incubation period of infectious disease. Am J Hyg 51, 310-318.

Sartwell PE (1952). The incubation period of poliomyelitis. Am J Public Health 42, 1403-1408.

Sartwell PE (1966). The incubation period and the dynamics of infectious disease. Am J Epidem 83, 204-216.

Sawyer WA (1914). Ninety-three persons infected by a typhoid carrier at a public dinner. J Am Med Assoc 63, 1537-1542.

Sawyer WA Meyer KF Eaton MD Bauer JH Putman P Schwentker FF (1944). Jaundice in army personnel in the western region of the United States and its relation to vaccination against yellow fever. Am J Hyg 39, 337 (cited by Sartwell, 1950).

Schechter MT Craib KJ LeTN et al (1989). Influence of zidovudine on progression to AIDS in cohort studies. Lancet 1, 1026-1027.

Schonberger LB Bregman DJ Sullivan-Bolyai JZ et al (1979). Guillain-Barre syndrome following vaccination in the national influenza immunization program, United States, 1976-1977. Am J Epidem 110, 105-123.

Schreuder BEC (1993). General aspects of transmissible spongiform encephalopathies and hypotheses about the agents. Veterinary Quarterly 15, 167-174).

Scott GB Fischl MA Klimas N et al (1985). Mothers of infants with the acquired immunodeficiency syndrome: Outcome of subsequent pregnancies. Conference on acquired immunodeficiency syndrom, Atlanta, April 1985, cited by Rogers et al, 1987.

Scott M Foster D Mirenda C et al (1989). Transgenic mice expressing hamster prion protein produce species-specific scrapie infectivity and amyloid plaques. Cell 59, 847-857.

Scott M Groth D Foster D et al (1993). Propogation with artificial properties in transgenic mice expressing chimeric PrP genes. Cell 73, 979-988.

Selik RM Haverkos HW Curran JW (1984). Acquired immune deficiency syndrome (AIDS) trends in the United States, 1978-1982. Am J Med 76, 493-500.

Selikoff IJ Hammond EC (1968). Community effects of nonoccupational environmental asbestos exposure. Am J public Health 58, 1658-1666.

Shapiro SS Wilk MB (1965). An analysis of variance test for normality (complete samples). Biometrika 52, 591.

Sheppard WF (1897). On the calculation of the most probable values of the frequency constants from the data arranged according to equidistant divisions of the scale. Proc Lond Math Soc 29, 353.

Siegel S (1956). Nonparametric statistics for behavioural sciences; p 184. McGraw Hill, New York.

Sigurdsson B (1954). Rida, a chronic encephalitis of sheep. Brit Vet J 110, 341-354.

Sklaviadis TK Manuelidis L Manuelidis EE (1989). Physical properties of the Creutzfeldt-Jakob disease agent. J Viro 63, 1212-1222.

Sklaviadis T Akowitz A Manuelidis EE Manuelidis L (1990). Nuclease treatment results in high specific purification of Creutzfeldt-Jakob disease infectivity with a density characteristic of nucleic acid-protein complexes. Arch Virol 112, 215-229.

Sklaviadis T Akowitz A Manuelidis EE Manuelidis L (1993). Nucleic acid binding proteins in highly purified Creutzfeldt-Jakob disease preparations. Proc Natl Acad Sci USA 90, 5713-5717.

Snedecor GW Cochran WG (1967). A test of skewness. In: Statistical Methods; 6th Edition, page 86. Iowa State University Press, Iowa, USA.

Soper HE (1929). The interpretation of periodicity in disease prevalence. J Royal Statistical Society, Vol XCII (New Series), 34-73.

Southwood R (1989). Chairman. Report of the working party on bovine spongiform encephalopathy. DOH & MAFF, London.

Stemshorn BW (1975). Un cas de tremlante naturelle chez une chevre. Can Vet J 16, 84-86.

Stewart AM Kneale GW (1970). Age-distribution of cancers caused by obstetric X-rays and their relevance to cancer latent periods. Lancet 1, 4-8.

Stillerman M Thalheimer W (1944). Attack rate and incubation period of measles. Am J Dis Childhood 67, 15-21.

Stockman S (1913). Scrapie: An obscure disease of sheep. J Comp Path 26, 317-327.

Stocks P (1931). Incubation period of measles. BMJ 11, 157.

Stocks P Karn MN (1928). A study of the epidemiology of measles. Annals of Eugenics 3, 361-398.

Sutherland I Springett VH (1987). Effectiveness of BGC vaccination in England and Wales in 1983. Tubercle 68, 81-92.

Swash M (1991). Clinical aspects of motor neurone disease. clues about cause. In: The aetiology of motor neurone disease: Conference report. Pub· MRC Environmental Epidemiology Unit, Southampton General Hospital, Southampton.

Tateishi J Kitamoto T Doh-ura K et al (1990). Immunochemical, molecular genetic, and transmission studies on a case of Gerstmann-Straussler-Sheinker syndrome. Neurology 40, 1578-1581.

Taylor A Santiago A Gonzales-Cortez A Gangarosa EJ (1974). Outbreak of typhoid fever in Trinidad in 1971 traced to a commercial ice-cream product. Am J Epidem 100, 150-157.

Taylor JMG (1989a). Models for the HIV infection and AIDS epidemic in the United States. Stats in Med 8, 45-58.

Taylor DM (1989b). Bovine spongiform encephalopathy and human health. Vet Rev 125, 413-415.

Taylor DM (1991). The control of bovine spongiform encephalopathy in Great Britain. Vet Rec 129, 522-526.

Taylor JMG Kuo J-M Detels R (1991). Is the incubation period of AIDS lengthening? J Acquir Immune Defic Syndr 4, 69-75.

Telo JMR (1994). Bovine spongiform encephalopathy: Ten cases discovered in Portugal. In: Report of the meeting of the OIE ad hoc group on bovine spongiform ncephalopathy; Paris, 1-2 September, 1994. Pub: Office International des Epizooties. 27-28.

Terao Y Hitoshi S Shimizu J et al (1992). Gerstmann-Straussler-Scheinker disease with heterozygous codon change at prion protein codon 129. Clin Neurol 32, 880-883.

Thiry L Sprecher-Goldberger S Jonckheer T et al (1985). Isolation of AIDS virus from cell-free breast milk of three healthy virus carriers. Letter. Lancet 2, 891-892.

Thomas PA Lubin K Milburg J et al (1987). Cohort comparison of children whose mothers have acquired immunodeficiency syndrome and children of well inner city mothers. Paediatr Infect Dis J 6, 247-51.

Thrusfield M (1986). Veterinary Epidemiology. Butterworths, London.

Tindall B Barker S Donovan B et al for Sydney AIDS study group. (1988). Characterization of the acute clinical illness associated with human immunodeficiency virus infection. Arch Intern Med 148, 945-9.

Tomonaga M (1962). Leukemia in Nagasaki atomic bomb survivors from 1945 through 1959. Bull WHO 26, 619-631.

Topley WWC (1919). The Goulstonian Lectures on the spread of bacterial infection. Lancet 2, 1-5; 45-49; 91-96.

Toumazos P Alley MR (1989). Scrapie in goats in Cyprus. NZ Vet J 37, 160.

Tranchant C Doh-ura K Steinmetz G et al (1991). Mutation du codon 117 du gene du prion dans une maladie de Gerstmann-Straussler-Sheinker. Rev Neurol (Paris) 147, 274-278.

Tyrrell DAJ (199). Chairman. Consultative committee on research into spong form encephalopathies. MAFF & DOH, London.

Van der Perre P Rouvroy D Lepage Pet al (1984). Acquired immunodeficiency syndrome in Rwanda. Lancet, 2, 62-5.

Velasco-Hernandez JX Hsieh YH (1994). Modelling the effect of treatment and behavioural change in HIV transmission dynamics. J Math Biol 32, 233-49.

VIDA (1975 to date). Veterinary Investigation Diagnosis Analysis database. Epidemiology Unit, Central Veterinary Laboratory (MAFF Agency), New Haw, Addlestone, Surrey, KT15 3NB.

Vinters HV Hudson AJ Kaufmann JCE (1986). Gerstmann-Straussler-Scheinker disease: Autopsy study of a familial case. Ann Neurol 20, 540-543.

Volkow P Ponce-de-Leon S Calva J et al (1993). Transfusion associated AIDS in Mexico. Clinical spectrum, conditional latency distribution, and survival. Rev Invest Clin 45, 133-138.

Ward JW (1993). Transfusion-associated (TA)-AIDS in the United States. Dev Biol Stand 81, 41-43.

Wear DJ Rabin ER Richardson LS Rapp F (1968). Virus replication and ultrastructural changes after induction of encephalitis in mice by measles virus. Exp Molecular Path 9, 405-417.

Weissmann C (1989). Sheep disease in human clothing. Nature 338, 298-299.

Weissmann C (1991a). The prion's progress. Nature 349, 569-571.

Weissmann C (1991b). A 'unified' theory of prion propogation. Nature 352. 679-683.

Wells GAH (1993). Pathology of nonhuman spongiform encephalopathies : Variations and their implications for pathogenesis. In: Transmissible spongiform encephalopathies - Impact on animal and human health. Ed: F Brown. Dev Biol Stand, vol 80. Karger, Basel. 61-69.

Wells GAH Scott AC Johnson CT et al (1987). A novel progressive spongiform encephalopathy in cattle. Vet Rec 121, 419-420.

Wells AH Hawkins SAC Hadlow WJ Spencer YI (1992). The discovery of bovine spongiform encephalopathy and observations on the vacuolar changes. In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London. 256-274.

Wells GAH Dawson M Hawkins SAC et al (1994). Infectivity in the ileum of cattle challenged orally with bovine spongiform encephalopathy. Vet Rec 135, 40-41.

Westaway D Zuliani V Cooper CM et al (1994). Homozygosity for prion protein alleles encoding glutamine-171 renders sheep susceptible to natural scrapie. Genes Dev 8, 959-969.

WHO/UNICEF (1992). Global programme on AIDS. Concensus statement from the WHO/UNICEF consultation on HIV transmission and breast-feeding. Weekly Epidemiological Record, 67, 177-179.

Wijeratne WVS Curnow RN (1990). A study of the inheritance of susceptibility to bovine spongiform encephalopathy. Vet Rec 126, 5-8.

Wilesmith JW (1994). An epidemiologist's view of bovine spongiform encephalopathy. Phil Trans R Soc Lond B, 343, 357-361.

Wilesmith JW Wells GAH Cranwell MP Ryan JMB (1988). Bovine spongiform encephalopathy: Epidemiological studies. Vet Rec 123, 638-644.

Wilesmith JW Ryan JMB Atkinson MJ (1991). Bovine spongiform encephalopathy: Epidemiological studies on the origin. Vet Rec 128, 199-203.

Wilesmith JW Ryan JMB Heuston WD Hoinville LJ (1992). Bovine spongiform encephalopathy: Epidemiological features 1985 to 1990. Vet Rec 130, 90-94.

Wilesmith JW Hoinville LJ Ryan JMB Sayers AR (1992). Bovine spongiform encephalopathy: Aspects of the clinical picture and analyses of possible changes 1986-1990. Vet Rec 130, 197-201.

Will RG Esmonde TFG Matthews WB (1992). Creuttzfeldt-Jakob disease epidemiology. In: Prion diseases of humans and animals. Eds: SB Prusiner, J Collinge, J Powell, B Anderton. Pub: Ellis Horwood. New York, London. 188-199.

Will RG Matthews WB (1982). Evidence for case-to-case transmission of Creutzfeldt-Jakob disease. J Neurol Neurosurg Psychiat 45, 235-238.

Williams ES Young S (1980). Chronic wasting disease of of captive mule deer: a spongiform encephalopathy. J Wildl Dis 16, 89-98.

Williams ES Young S (1982). Spongiform encephalopathy of Rocky Mountain elk. J Wildl Dis 18, 465-471.

Williams ES Young S (1992). Spongiform encephalopathies in cervidae. Rev sci tech Off int Epiz 11, 551-567.

Willoughby K Kelly DF Lyon DG Wells GAH (1992). Spongiform encephalopathy in a captive puma (*Felis concolor*). Vet Rec 131, 431-434.

Wilson DR Anderson RD Smith W (1950). Studies in scrapie. J Comp Path 60, 267-282.

Wood JLN Done SH (1992). Natural scrapie in goats: Neuropathology. Vet Rec, 131, 93-96.

Wood JLN Done SH Pritchard GC Wooldridge MJA (1992). Scrapie in goats: Case histories and clinical signs. Vet Rec 131, 66-68.

Wooldridge MJA (1991). Preliminary studies on scrapie in sheep in Great Britain. MSc (Epidemiology) Thesis. London School of Hygiene and Tropical Medicine, London University.

Wooldridge MJA Wood J (1991). Some aspects of scrapie epidemiology in the goat. Goat Vet Soc J, 12, 4-7.

Wooldridge MJA Hoinville LJ Wilesmith JW (1992). A scrapie survey by postal questionnaire: Aims, problems and results. In: Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine, Edinburgh 1-3rd April, 1992. 78-89.

Wyatt JM Pearson GR Smerdon T et al (1990). Spongiform encephalopathy in a cat. Vet Rec 126, 513.

Wyatt JM Pearson GR Smerdon T et al (1991). Naturally occuring scrapie-like spongiform encephalopathy in five domestic cats. Vet Rec 129, 233-236.

Young (1799). In: Annals of Agriculture. Cited by Stockman (1913).

Zeigler JB Cooper DA Johnson RD et al (1985). Postnatal transmission of AIDS-associated retrovirus from mother to infant. Lancet 1, 896-897.

Zigas V Gajdusek DC (1959). Kuru. Papua and New Guinea Med J 3, 1-24.