

APPENDIX I:

GROUPED DATA FROM THE ILLE-ET-VILAINE STUDY OF OESOPHAGEAL CANCER USED FOR ILLUSTRATION IN CHAPTERS 4 AND 6

AGE (YRS)	ALCOHOL (GM/DAY)	TOBACCO (GM/DAY)	CASES	CONTROLS
-----	-----	-----	-----	-----
25-34	0-39	0-9	0	40
×	×	10-19	0	10
×	×	20-29	0	6
×	×	30+	0	5
×	40-79	0-9	0	27
×	×	10-19	0	7
×	×	20-29	0	4
×	×	30+	0	7
×	80-119	0-9	0	2
×	×	10-19	0	1
×	×	20-29	0	0
×	×	30+	0	2
×	120+	0-9	0	1
×	×	10-19	1	0
×	×	20-29	0	1
×	×	30+	0	2
35-44	0-39	0-9	0	60
×	×	10-19	1	13
×	×	20-29	0	7
×	×	30+	0	8
×	40-79	0-9	0	35
×	×	10-19	3	20
×	×	20-29	1	13
×	×	30+	0	8
×	80-119	0-9	0	11
×	×	10-19	0	6
×	×	20-29	0	2
×	×	30+	0	1
×	120+	0-9	2	1
×	×	10-19	0	3
×	×	20-29	2	2
×	×	30+	0	0

AGE (YRS)	ALCOHOL (GM/DAY)	TOBACCO (GM/DAY)	CASES	CONTROLS
-----	-----	-----	-----	-----
45-54	0-39	0-9	1	45
×	×	10-19	0	18
×	×	20-29	0	10
×	×	30+	0	4
×	40-79	0-9	6	32
×	×	10-19	4	17
×	×	20-29	5	10
×	×	30+	5	2
×	80-119	0-9	3	13
×	×	10-19	6	8
×	×	20-29	1	4
×	×	30+	2	2
×	120+	0-9	4	0
×	×	10-19	3	1
×	×	20-29	2	1
×	×	30+	4	0
55-64	0-39	0-9	2	47
×	×	10-19	3	19
×	×	20-29	3	9
×	×	30+	4	2
×	40-79	0-9	9	31
×	×	10-19	6	15
×	×	20-29	4	13
×	×	30+	3	3
×	80-119	0-9	9	9
×	×	10-19	8	7
×	×	20-29	3	3
×	×	30+	4	0
×	120+	0-9	5	5
×	×	10-19	6	1
×	×	20-29	2	1
×	×	30+	5	1

AGE (YRS)	ALCOHOL (GM/DAY)	TOBACCO (GM/DAY)	CASES	CONTROLS
-----	-----	-----	-----	-----
65-74	0-39	0-9	5	43
✱	✱	10-19	4	10
✱	✱	20-29	2	5
✱	✱	30+	0	2
✱	40-79	0-9	17	17
✱	✱	10-19	3	7
✱	✱	20-29	5	4
✱	✱	30+	0	0
✱	80-119	0-9	6	7
✱	✱	10-19	4	8
✱	✱	20-29	2	1
✱	✱	30+	1	0
✱	120+	0-9	3	1
✱	✱	10-19	1	1
✱	✱	20-29	1	0
✱	✱	30+	1	0
75+	0-39	0-9	1	17
✱	✱	10-19	2	4
✱	✱	20-29	0	0
✱	✱	30+	1	2
✱	40-79	0-9	2	3
✱	✱	10-19	1	2
✱	✱	20-29	0	3
✱	✱	30+	1	0
✱	80-119	0-9	1	0
✱	✱	10-19	1	0
✱	✱	20-29	0	0
✱	✱	30+	0	0
✱	120+	0-9	2	0
✱	✱	10-19	1	0
✱	✱	20-29	0	0
✱	✱	30+	0	0

APPENDIX II:

GROUPED DATA FROM THE OXFORD CHILDHOOD CANCER SURVEY USED FOR ILLUSTRATION IN CHAPTERS 6 AND 7

YEAR OF DEATH	AGE AT DEATH	X-RAYED	CASES	CONTROLS
1944	9	YES	3	0
1944	9	NO	25	28
1945	9	YES	5	2
1945	9	NO	16	19
1945	8	YES	2	2
1945	8	NO	30	30
1946	9	YES	7	1
1946	9	NO	28	34
1946	8	YES	7	2
1946	8	NO	28	33
1946	7	YES	2	0
1946	7	NO	36	38
1947	9	YES	5	1
1947	9	NO	25	29
1947	8	YES	3	1
1947	8	NO	40	42
1947	7	YES	5	1
1947	7	NO	44	48
1947	6	YES	11	2
1947	6	NO	42	51
1948	9	YES	6	4
1948	9	NO	25	27
1948	8	YES	6	4
1948	8	NO	29	31
1948	7	YES	11	2
1948	7	NO	35	44
1948	6	YES	4	1
1948	6	NO	49	52
1948	5	YES	4	7
1948	5	NO	57	54
1949	9	YES	2	4
1949	9	NO	38	36
1949	8	YES	8	3
1949	8	NO	21	26
1949	7	YES	8	5
1949	7	NO	36	39
1949	6	YES	6	3
1949	6	NO	46	49
1949	5	YES	5	2
1949	5	NO	50	53
1949	4	YES	15	4
1949	4	NO	46	57

CHILDHOOD CANCER SURVEY

285

YEAR OF DEATH	AGE AT DEATH	X-RAYED	CASES	CONTROLS
1950	9	YES	4	1
1950	9	NO	27	30
1950	8	YES	9	4
1950	8	NO	39	44
1950	7	YES	9	5
1950	7	NO	35	39
1950	6	YES	4	2
1950	6	NO	38	40
1950	5	YES	12	7
1950	5	NO	41	46
1950	4	YES	8	5
1950	4	NO	48	51
1950	3	YES	8	8
1950	3	NO	63	63
1951	9	YES	6	2
1951	9	NO	37	41
1951	8	YES	8	3
1951	8	NO	35	40
1951	7	YES	12	5
1951	7	NO	31	38
1951	6	YES	4	4
1951	6	NO	36	36
1951	5	YES	7	1
1951	5	NO	37	43
1951	4	YES	16	6
1951	4	NO	54	64
1951	3	YES	12	5
1951	3	NO	63	70
1951	2	YES	9	11
1951	2	NO	62	60
1952	9	YES	4	5
1952	9	NO	33	32
1952	8	YES	7	2
1952	8	NO	24	29
1952	7	YES	8	5
1952	7	NO	34	37
1952	6	YES	11	8
1952	6	NO	35	38
1952	5	YES	5	5
1952	5	NO	42	42
1952	4	YES	12	6
1952	4	NO	43	49
1952	3	YES	8	6
1952	3	NO	55	57
1952	2	YES	17	10
1952	2	NO	74	81
1952	1	YES	9	7
1952	1	NO	34	36

APPENDIX II

YEAR OF DEATH	AGE AT DEATH	X-RAYED	CASES	CONTROLS
1953	9	YES	3	5
1953	9	NO	36	34
1953	8	YES	2	5
1953	8	NO	33	30
1953	7	YES	7	2
1953	7	NO	25	30
1953	6	YES	6	8
1953	6	NO	47	45
1953	5	YES	5	1
1953	5	NO	44	48
1953	4	YES	11	13
1953	4	NO	64	62
1953	3	YES	14	9
1953	3	NO	50	55
1953	2	YES	13	11
1953	2	NO	56	58
1953	1	YES	8	9
1953	1	NO	56	55
1953	0	YES	6	4
1953	0	NO	43	45
1954	9	YES	4	4
1954	9	NO	25	25
1954	8	YES	8	8
1954	8	NO	32	32
1954	7	YES	4	6
1954	7	NO	23	21
1954	6	YES	8	8
1954	6	NO	40	40
1954	5	YES	7	6
1954	5	NO	36	37
1954	4	YES	15	8
1954	4	NO	46	53
1954	3	YES	15	14
1954	3	NO	62	63
1954	2	YES	9	6
1954	2	NO	46	49
1954	1	YES	9	5
1954	1	NO	51	55
1954	0	YES	5	5
1954	0	NO	41	41
1955	9	YES	6	2
1955	9	NO	22	26
1955	8	YES	3	4
1955	8	NO	30	29
1955	7	YES	9	2
1955	7	NO	23	30

YEAR OF DEATH	AGE AT DEATH	X-RAYED	CASES	CONTROLS
1955	6	YES	12	9
1955	6	NO	34	37
1955	5	YES	14	5
1955	5	NO	43	52
1955	4	YES	16	6
1955	4	NO	40	50
1955	3	YES	17	7
1955	3	NO	61	71
1955	2	YES	8	5
1955	2	NO	50	53
1955	1	YES	8	10
1955	1	NO	44	42
1955	0	YES	9	3
1955	0	NO	22	28
1956	8	YES	5	2
1956	8	NO	23	26
1956	7	YES	9	1
1956	7	NO	37	45
1956	6	YES	11	7
1956	6	NO	31	35
1956	5	YES	6	9
1956	5	NO	39	36
1956	4	YES	14	13
1956	4	NO	49	50
1956	3	YES	21	9
1956	3	NO	50	62
1956	2	YES	16	11
1956	2	NO	53	58
1956	1	YES	6	4
1956	1	NO	37	39
1956	0	YES	9	8
1956	0	NO	41	42
1957	7	YES	8	2
1957	7	NO	23	29
1957	6	YES	9	3
1957	6	NO	25	31
1957	5	YES	8	7
1957	5	NO	46	47
1957	4	YES	4	4
1957	4	NO	42	42
1957	3	YES	11	7
1957	3	NO	47	51
1957	2	YES	11	5
1957	2	NO	51	57
1957	1	YES	6	6
1957	1	NO	46	46
1957	0	YES	9	6
1957	0	NO	32	35

APPENDIX II

YEAR OF DEATH	AGE AT DEATH	X-RAYED	CASES	CONTROLS
1958	6	YES	4	5
1958	6	NO	30	29
1958	5	YES	4	6
1958	5	NO	48	46
1958	4	YES	9	9
1958	4	NO	54	54
1958	3	YES	9	7
1958	3	NO	50	52
1958	2	YES	10	7
1958	2	NO	78	81
1958	1	YES	14	7
1958	1	NO	48	55
1958	0	YES	6	4
1958	0	NO	41	43
1959	5	YES	3	2
1959	5	NO	50	51
1959	4	YES	4	3
1959	4	NO	53	54
1959	3	YES	6	4
1959	3	NO	68	70
1959	2	YES	10	10
1959	2	NO	58	58
1959	1	YES	4	3
1959	1	NO	57	58
1959	0	YES	3	4
1959	0	NO	42	41
1960	4	YES	3	2
1960	4	NO	42	43
1960	3	YES	10	10
1960	3	NO	52	52
1960	2	YES	4	5
1960	2	NO	69	68
1960	1	YES	10	4
1960	1	NO	43	49
1960	0	YES	5	5
1960	0	NO	34	34
1961	3	YES	4	4
1961	3	NO	41	41
1961	2	YES	3	6
1961	2	NO	48	45
1961	1	YES	13	5
1961	1	NO	42	50
1961	0	YES	1	3
1961	0	NO	40	38

<u>YEAR OF DEATH</u>	<u>AGE AT DEATH</u>	<u>X-RAYED</u>	<u>CASES</u>	<u>CONTROLS</u>
1962	2	YES	7	2
1962	2	NO	46	51
1962	1	YES	5	2
1962	1	NO	46	49
1962	0	YES	7	4
1962	0	NO	35	38
1963	1	YES	6	6
1963	1	NO	40	40
1963	0	YES	3	4
1963	0	NO	51	50
1964	0	YES	7	1
1964	0	NO	25	31

APPENDIX III:

MATCHED DATA FROM THE LOS ANGELES STUDY OF ENDOMETRIAL CANCER USED FOR ILLUSTRATION IN CHAPTERS 5 AND 7

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBESITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CASE	74	NO	NO	YES	YES	3	96+	YES
CONTROL	75	NO	NO	UNK	NO	0	0	NO
CONTROL	74	NO	NO	UNK	NO	0	0	NO
CONTROL	74	NO	NO	UNK	NO	0	0	NO
CONTROL	75	NO	NO	YES	YES	1	48	YES
CASE	67	NO	NO	NO	YES	3	96+	YES
CONTROL	67	NO	NO	NO	YES	3	5	NO
CONTROL	67	NO	YES	YES	NO	0	0	YES
CONTROL	67	NO	NO	NO	YES	2	53	NO
CONTROL	68	NO	NO	NO	YES	2	45	YES
CASE	76	NO	YES	YES	YES	1	9	YES
CONTROL	76	NO	YES	YES	YES	2	96+	YES
CONTROL	76	NO	YES	NO	YES	1	3	YES
CONTROL	76	NO	YES	YES	YES	2	15	YES
CONTROL	77	NO	NO	NO	YES	1	36	YES
CASE	71	NO	NO	UNK	YES	UNK	96+	NO
CONTROL	70	YES	NO	NO	YES	2	7	YES
CONTROL	70	NO	NO	NO	YES	0	0	YES
CONTROL	71	NO	YES	YES	YES	2	7	YES
CONTROL	70	NO	NO	YES	YES	2	27	YES
CASE	69	YES	NO	YES	YES	2	36	YES
CONTROL	69	NO	YES	NO	YES	1	96+	YES
CONTROL	69	NO	NO	YES	YES	2	1	YES
CONTROL	69	NO	NO	NO	YES	0	0	YES
CONTROL	68	NO	NO	UNK	NO	0	0	NO
CASE	70	NO	YES	YES	YES	2	71	YES
CONTROL	71	NO	NO	NO	NO	0	0	NO
CONTROL	71	NO	YES	YES	YES	3	5	YES
CONTROL	70	NO	NO	YES	NO	0	0	NO
CONTROL	71	NO	NO	UNK	NO	0	0	NO
CASE	65	YES	NO	NO	YES	1	96+	YES
CONTROL	65	NO	NO	UNK	NO	0	0	NO
CONTROL	64	NO	NO	NO	YES	3	91	YES
CONTROL	64	NO	NO	NO	YES	2	96+	YES
CONTROL	65	NO	NO	YES	YES	2	60	NO
CASE	68	YES	YES	YES	YES	1	36	YES
CONTROL	68	NO	YES	UNK	NO	0	0	YES
CONTROL	68	NO	NO	YES	NO	0	0	YES
CONTROL	68	YES	NO	UNK	YES	0	0	NO

ENDOMETRIAL CANCER STUDY

291

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBESITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CONTROL	68	NO	NO	UNK	YES	1	1	YES
CASE	61	NO	NO	UNK	NO	0	0	YES
CONTROL	61	NO	NO	YES	NO	0	0	YES
CONTROL	61	NO	NO	NO	YES	1	24	YES
CONTROL	61	NO	NO	UNK	NO	0	0	YES
CONTROL	60	YES	NO	NO	NO	0	0	NO
CASE	64	NO	NO	YES	YES	1	54	YES
CONTROL	64	NO	NO	UNK	NO	0	0	NO
CONTROL	65	NO	YES	UNK	YES	3	2	YES
CONTROL	64	NO	YES	YES	YES	3	10	YES
CONTROL	65	NO	NO	UNK	NO	0	0	NO
CASE	68	YES	NO	YES	YES	3	96+	YES
CONTROL	69	NO	NO	UNK	NO	0	0	NO
CONTROL	69	NO	NO	YES	NO	0	0	YES
CONTROL	69	YES	YES	YES	YES	0	0	YES
CONTROL	69	YES	NO	YES	YES	1	35	NO
CASE	74	NO	NO	NO	YES	2	96+	YES
CONTROL	74	NO	YES	NO	YES	3	4	YES
CONTROL	73	NO	YES	NO	YES	2	11	YES
CONTROL	74	NO	NO	YES	YES	1	6	YES
CONTROL	74	NO	NO	YES	YES	1	12	NO
CASE	67	YES	NO	YES	YES	0	0	YES
CONTROL	68	NO	YES	NO	YES	0	0	YES
CONTROL	68	NO	YES	YES	YES	3	65	YES
CONTROL	68	NO	NO	UNK	NO	0	0	NO
CONTROL	68	NO	NO	YES	YES	2	96+	YES
CASE	62	YES	NO	NO	YES	1	UNK	YES
CONTROL	62	YES	NO	NO	NO	0	0	NO
CONTROL	63	NO	NO	YES	NO	0	0	NO
CONTROL	63	NO	NO	UNK	NO	0	0	NO
CONTROL	63	NO	NO	YES	YES	2	UNK	NO
CASE	71	YES	NO	YES	YES	2	59	YES
CONTROL	70	NO	YES	YES	NO	0	0	YES
CONTROL	71	NO	NO	YES	YES	UNK	UNK	YES
CONTROL	71	NO	YES	YES	NO	0	0	YES
CONTROL	71	NO	YES	YES	YES	2	84	YES
CASE	83	NO	YES	YES	YES	3	96+	YES
CONTROL	82	NO	NO	YES	NO	0	0	NO
CONTROL	82	NO	YES	NO	YES	3	4	YES
CONTROL	82	NO	YES	NO	NO	0	0	YES
CONTROL	82	NO	NO	UNK	NO	0	0	NO
CASE	70	NO	NO	YES	NO	0	0	YES
CONTROL	70	YES	YES	YES	YES	2	55	YES
CONTROL	70	NO	YES	YES	YES	2	14	YES
CONTROL	70	NO	YES	YES	YES	1	39	YES

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBESITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CONTROL	70	NO	YES	YES	NO	0	0	YES
CASE	74	NO	NO	NO	YES	0	0	YES
CONTROL	75	YES	YES	NO	YES	2	6	YES
CONTROL	74	NO	NO	YES	NO	0	0	YES
CONTROL	74	NO	YES	NO	YES	2	46	YES
CONTROL	75	NO	NO	UNK	NO	0	0	NO
CASE	70	NO	NO	UNK	YES	0	0	YES
CONTROL	70	NO	YES	NO	YES	1	96+	YES
CONTROL	70	NO	NO	UNK	NO	0	0	NO
CONTROL	70	NO	NO	UNK	NO	0	0	YES
CONTROL	70	NO	NO	UNK	NO	0	0	NO
CASE	66	NO	YES	YES	YES	3	48	YES
CONTROL	66	NO	NO	UNK	YES	1	96+	YES
CONTROL	66	NO	NO	UNK	NO	0	0	YES
CONTROL	66	NO	NO	YES	NO	0	0	NO
CONTROL	66	NO	YES	YES	YES	1	12	YES
CASE	77	NO	NO	YES	YES	3	4	YES
CONTROL	77	YES	YES	YES	YES	0	0	YES
CONTROL	77	NO	YES	NO	YES	2	24	YES
CONTROL	77	NO	NO	YES	NO	0	0	NO
CONTROL	78	NO	YES	YES	YES	2	9	YES
CASE	66	NO	YES	NO	YES	3	29	YES
CONTROL	67	NO	YES	NO	NO	0	0	YES
CONTROL	66	NO	NO	YES	NO	0	0	YES
CONTROL	67	NO	NO	YES	NO	0	0	YES
CONTROL	69	NO	YES	YES	YES	2	10	YES
CASE	71	NO	YES	YES	YES	1	96+	NO
CONTROL	72	NO	NO	UNK	NO	0	0	NO
CONTROL	72	NO	NO	NO	NO	0	0	YES
CONTROL	71	NO	NO	UNK	NO	0	0	NO
CONTROL	71	NO	YES	YES	NO	0	0	YES
CASE	80	NO	NO	NO	YES	2	UNK	YES
CONTROL	79	NO	NO	UNK	NO	0	0	NO
CONTROL	79	NO	NO	NO	NO	0	0	NO
CONTROL	79	NO	NO	YES	NO	0	0	NO
CONTROL	80	NO	NO	NO	NO	0	0	NO
CASE	64	NO	NO	YES	YES	2	UNK	YES
CONTROL	64	NO	NO	NO	YES	0	0	YES
CONTROL	63	NO	NO	YES	YES	1	60	YES
CONTROL	64	NO	YES	NO	YES	1	6	YES
CONTROL	66	NO	YES	YES	YES	1	UNK	YES
CASE	63	NO	NO	NO	YES	1	60	YES
CONTROL	63	NO	YES	NO	YES	1	96+	YES
CONTROL	65	NO	NO	NO	YES	1	25	NO
CONTROL	65	NO	NO	NO	NO	0	0	YES
CONTROL	64	NO	NO	NO	YES	1	96+	YES

ENDOMETRIAL CANCER STUDY

293

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBEILITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CASE	72	YES	NO	YES	NO	0	0	YES
CONTROL	72	NO	YES	UNK	NO	0	0	NO
CONTROL	72	NO	YES	YES	NO	0	0	YES
CONTROL	72	NO	YES	NO	YES	1	48	YES
CONTROL	72	NO	YES	YES	YES	0	0	YES
CASE	57	NO	NO	NO	YES	3	12	NO
CONTROL	57	NO	YES	YES	YES	0	0	YES
CONTROL	58	NO	NO	YES	YES	1	36	YES
CONTROL	57	NO	NO	NO	YES	1	36	NO
CONTROL	57	NO	NO	NO	YES	0	0	NO
CASE	74	YES	NO	YES	NO	0	0	YES
CONTROL	74	NO	NO	YES	NO	0	0	YES
CONTROL	73	NO	NO	YES	YES	2	2	YES
CONTROL	75	NO	NO	YES	NO	0	0	YES
CONTROL	75	NO	NO	UNK	NO	0	0	NO
CASE	62	NO	YES	YES	YES	2	6	YES
CONTROL	62	NO	NO	YES	YES	2	37	YES
CONTROL	62	NO	NO	YES	YES	2	63	YES
CONTROL	63	NO	NO	UNK	NO	0	0	NO
CONTROL	61	YES	YES	YES	YES	3	96+	YES
CASE	73	NO	YES	YES	YES	1	4	YES
CONTROL	72	NO	NO	NO	YES	2	90	YES
CONTROL	73	NO	NO	NO	YES	3	5	YES
CONTROL	73	NO	YES	NO	YES	1	15	YES
CONTROL	73	NO	YES	NO	NO	0	0	NO
CASE	71	NO	YES	YES	YES	1	UNK	YES
CONTROL	71	NO	NO	UNK	NO	0	0	NO
CONTROL	71	NO	NO	NO	NO	0	0	YES
CONTROL	71	NO	NO	NO	NO	0	0	YES
CONTROL	71	NO	YES	NO	YES	UNK	UNK	YES
CASE	64	NO	YES	YES	NO	0	0	YES
CONTROL	65	NO	NO	YES	YES	3	96+	YES
CONTROL	64	NO	NO	YES	YES	3	96+	YES
CONTROL	64	NO	NO	YES	YES	2	36	YES
CONTROL	64	NO	NO	YES	YES	3	96+	NO
CASE	63	NO	NO	NO	YES	UNK	96+	YES
CONTROL	64	NO	NO	YES	NO	0	0	YES
CONTROL	62	NO	NO	YES	NO	0	0	YES
CONTROL	64	YES	NO	NO	YES	1	18	NO
CONTROL	64	NO	YES	YES	YES	3	UNK	YES
CASE	79	YES	YES	YES	YES	1	96+	YES
CONTROL	78	YES	YES	YES	YES	1	96+	YES
CONTROL	79	NO	NO	YES	NO	0	0	YES
CONTROL	79	NO	YES	NO	YES	0	0	YES
CONTROL	78	NO	NO	YES	YES	1	24	YES

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBESITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CASE	80	NO	NO	YES	YES	1	15	YES
CONTROL	81	NO	YES	YES	NO	0	0	YES
CONTROL	81	NO	YES	NO	YES	1	18	YES
CONTROL	80	NO	NO	YES	YES	2	74	YES
CONTROL	80	NO	YES	YES	NO	0	0	YES
CASE	82	NO	YES	YES	YES	2	6	YES
CONTROL	82	NO	NO	YES	NO	0	0	YES
CONTROL	81	NO	YES	UNK	NO	0	0	YES
CONTROL	81	NO	YES	YES	YES	1	12	YES
CONTROL	82	NO	YES	YES	YES	2	13	YES
CASE	71	NO	YES	NO	YES	UNK	84	YES
CONTROL	71	NO	YES	YES	NO	0	0	YES
CONTROL	71	YES	NO	YES	NO	0	0	YES
CONTROL	71	YES	NO	YES	YES	1	96+	YES
CONTROL	71	NO	NO	NO	YES	1	30	YES
CASE	83	NO	YES	YES	YES	3	14	YES
CONTROL	83	NO	YES	YES	NO	0	0	YES
CONTROL	83	NO	NO	NO	NO	0	0	YES
CONTROL	83	NO	YES	YES	YES	2	16	YES
CONTROL	83	NO	NO	NO	NO	0	0	YES
CASE	61	NO	YES	NO	YES	3	96+	YES
CONTROL	60	NO	NO	NO	NO	0	0	YES
CONTROL	61	NO	NO	NO	YES	1	24	YES
CONTROL	62	NO	NO	YES	NO	0	0	YES
CONTROL	61	NO	NO	NO	YES	0	0	YES
CASE	71	NO	NO	NO	YES	1	96+	YES
CONTROL	71	NO	NO	YES	NO	0	0	NO
CONTROL	71	NO	YES	YES	NO	0	0	NO
CONTROL	70	NO	NO	NO	NO	0	0	NO
CONTROL	71	NO	YES	YES	YES	1	3	YES
CASE	69	NO	YES	YES	YES	2	40	YES
CONTROL	69	YES	NO	YES	NO	0	0	YES
CONTROL	70	NO	YES	NO	YES	0	0	YES
CONTROL	70	NO	YES	NO	YES	1	32	YES
CONTROL	70	NO	NO	YES	YES	UNK	UNK	YES
CASE	77	NO	NO	YES	YES	3	73	YES
CONTROL	76	NO	YES	NO	YES	0	0	YES
CONTROL	76	NO	YES	YES	YES	0	0	YES
CONTROL	77	YES	YES	YES	YES	0	0	YES
CONTROL	77	NO	YES	NO	NO	0	0	YES
CASE	64	NO	NO	YES	YES	1	37	NO
CONTROL	64	NO	NO	YES	YES	3	6	NO
CONTROL	63	YES	NO	YES	NO	0	0	NO
CONTROL	63	NO	YES	NO	YES	UNK	UNK	YES
CONTROL	63	NO	YES	YES	NO	0	0	YES

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBEILITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CASE	79	YES	NO	NO	NO	0	0	NO
CONTROL	82	NO	NO	YES	YES	1	UNK	YES
CONTROL	78	NO	NO	NO	NO	0	0	NO
CONTROL	80	NO	NO	YES	NO	0	0	NO
CONTROL	81	NO	NO	NO	NO	0	0	NO
CASE	72	NO	NO	NO	YES	0	0	YES
CONTROL	72	NO	NO	YES	YES	2	57	YES
CONTROL	73	NO	NO	UNK	NO	0	0	NO
CONTROL	73	YES	YES	NO	YES	2	96+	YES
CONTROL	73	NO	NO	NO	NO	0	0	YES
CASE	82	YES	YES	YES	YES	3	96+	YES
CONTROL	81	NO	NO	UNK	NO	0	0	NO
CONTROL	81	NO	NO	YES	NO	0	0	NO
CONTROL	81	NO	NO	YES	YES	0	0	YES
CONTROL	81	NO	YES	YES	NO	0	0	YES
CASE	73	NO	YES	YES	YES	2	60	YES
CONTROL	74	NO	NO	YES	YES	1	1	YES
CONTROL	75	NO	NO	NO	NO	0	0	YES
CONTROL	75	NO	YES	YES	YES	1	96+	YES
CONTROL	74	NO	NO	NO	NO	0	0	NO
CASE	69	NO	NO	UNK	YES	UNK	UNK	YES
CONTROL	68	NO	NO	NO	NO	0	0	YES
CONTROL	68	NO	NO	YES	YES	2	48	YES
CONTROL	68	NO	NO	NO	YES	1	96+	NO
CONTROL	70	NO	NO	NO	NO	0	0	NO
CASE	79	NO	YES	YES	YES	1	67	YES
CONTROL	79	NO	YES	YES	NO	0	0	YES
CONTROL	79	NO	YES	YES	NO	0	0	YES
CONTROL	78	YES	NO	YES	YES	1	UNK	YES
CONTROL	79	NO	NO	YES	NO	0	0	YES
CASE	72	NO	NO	YES	YES	3	60	NO
CONTROL	71	NO	NO	NO	YES	0	0	YES
CONTROL	72	NO	NO	NO	NO	0	0	YES
CONTROL	72	NO	YES	YES	YES	3	96+	YES
CONTROL	71	NO	YES	YES	YES	3	12	YES
CASE	72	NO	YES	YES	YES	1	27	YES
CONTROL	72	NO	YES	YES	YES	1	3	YES
CONTROL	71	NO	NO	UNK	NO	0	0	NO
CONTROL	72	NO	YES	YES	NO	0	0	YES
CONTROL	72	NO	YES	YES	NO	0	0	YES
CASE	65	NO	YES	YES	YES	2	16	YES
CONTROL	67	NO	NO	NO	NO	0	0	NO
CONTROL	67	NO	NO	UNK	NO	0	0	NO
CONTROL	66	NO	NO	YES	NO	0	0	YES
CONTROL	66	NO	NO	NO	YES	2	3	NO

CASE OR CONTROL	AGE	GALL BLADDER DISEASE	HYPER TENSION	OBESEITY	ESTROGEN (ANY) USE	CONJUGATED DOSE (SEE CODE)	ESTROGEN DURATION (MONTHS)	NON ESTROGEN DRUG
CASE	67	NO	YES	YES	YES	2	96+	YES
CONTROL	66	NO	NO	YES	YES	2	56	YES
CONTROL	66	NO	NO	YES	NO	0	0	NO
CONTROL	67	NO	NO	YES	YES	1	UNK	YES
CONTROL	67	NO	NO	YES	YES	2	34	YES
CASE	64	YES	NO	YES	YES	3	96+	YES
CONTROL	63	NO	NO	YES	NO	0	0	YES
CONTROL	64	NO	NO	YES	YES	1	4	YES
CONTROL	63	NO	NO	YES	NO	0	0	YES
CONTROL	65	NO	NO	UNK	NO	0	0	NO
CASE	62	NO	NO	UNK	YES	2	36	NO
CONTROL	63	NO	NO	YES	NO	0	0	NO
CONTROL	62	NO	NO	NO	NO	0	0	YES
CONTROL	62	NO	NO	UNK	YES	3	UNK	NO
CONTROL	62	NO	NO	UNK	NO	0	0	NO
CASE	83	YES	YES	UNK	NO	0	0	YES
CONTROL	83	YES	NO	UNK	NO	0	0	NO
CONTROL	82	NO	NO	NO	YES	2	6	YES
CONTROL	83	NO	NO	UNK	NO	0	0	YES
CONTROL	83	YES	NO	UNK	NO	0	0	YES
CASE	81	NO	NO	YES	YES	0	0	YES
CONTROL	79	NO	NO	UNK	NO	0	0	NO
CONTROL	80	NO	NO	YES	NO	0	0	YES
CONTROL	82	NO	NO	YES	NO	0	0	YES
CONTROL	80	NO	NO	NO	NO	0	0	NO
CASE	67	NO	NO	YES	YES	2	96+	YES
CONTROL	66	NO	NO	YES	YES	2	40	YES
CONTROL	68	NO	NO	UNK	NO	0	0	YES
CONTROL	65	NO	NO	NO	NO	0	0	YES
CONTROL	66	NO	YES	YES	YES	1	96+	YES
CASE	73	YES	YES	YES	YES	1	UNK	YES
CONTROL	72	NO	NO	YES	YES	1	12	YES
CONTROL	71	NO	NO	YES	YES	1	96+	YES
CONTROL	73	YES	NO	NO	YES	2	96+	YES
CONTROL	72	NO	NO	YES	NO	0	0	YES
CASE	67	YES	NO	NO	YES	3	96+	YES
CONTROL	67	YES	NO	YES	YES	2	96+	YES
CONTROL	68	NO	NO	YES	NO	0	0	NO
CONTROL	67	NO	NO	YES	NO	0	0	NO
CONTROL	67	NO	NO	YES	NO	0	0	YES
CASE	74	NO	YES	YES	YES	2	9	YES
CONTROL	75	NO	NO	NO	NO	0	0	YES
CONTROL	75	NO	NO	UNK	NO	0	0	NO
CONTROL	75	YES	YES	NO	NO	0	0	YES
CONTROL	75	NO	NO	NO	YES	2	41	YES
CASE	68	YES	NO	YES	YES	3	18	YES
CONTROL	69	NO	NO	YES	YES	2	96+	YES
CONTROL	70	NO	NO	UNK	NO	0	0	NO
CONTROL	69	NO	YES	YES	YES	2	92	YES
CONTROL	69	NO	YES	NO	YES	3	59	YES

APPENDIX IV

LISTING OF PROGRAM MATCH

The following main program and subroutine perform the conditional logistic regression analysis for matched sets consisting of a single case and a variable number of controls, as described in § 7.4. To illustrate how the program works we have input the data from the Los Angeles endometrial cancer study shown in Appendix III. Four different models were fit in hierarchical fashion using the variables GALL, OB, EST and $GALL \times EST$ (see Table 7.2). The last two of these correspond precisely to Models 4 and 8 of Table 7.7. Note the use of the SCORE statistic, calculated at the first iteration of each fit with the estimated regression coefficients from the previous model, to test the significance of the added variable(s).

The subprogram calls the IBM scientific subroutine SINV for inversion of symmetric matrices.

MAIN PROGRAM

(READS DATA, SETS UP RISK VARIABLES, CALLS SUBROUTINE)

```

C   MASTER MAIN
C   DIMENSION NR(63),IVAR(6),W(6),EZB(6),
C   * Z(6,5,63),B(6),SCORE(6),COV(21),COVI(21)
C
C   DIMENSION IVAR(NM),NCA(NS),NCT(NS),IP(NMAX2),W(NM),WWW(NM),
C   *Z(NM,NRMAX,NS),B(NM),SCORE(NM),COV(NM1),COVI(NM1)
C   SEE SUBROUTINE FOR DEFINITIONS
C
C   DATA NR/63*5/,B/6*0.0/,IVAR/1,2,3,4,5,6/,NP/6/
C   DATA NM,NRMAX,NIT,EPS/6,5,10,0.0001/
C   NM1=NM*(NM+1)/2
C   I=0
C   READ(1,100) GALL,OB,EST
1000 CONTINUE
100  FORMAT(10X,F5.0,5X,F5.0,15X,F5.0)
C
C   I IS THE ORDER NUMBER OF THE SET
C
C   I=I+1
C   K=0
C   DO 6 KK=1,5
C   K=K+1
C   Z(1,K,I)=2-GALL
C   IF (OB.EQ.9) OB=2
C   Z(2,K,I)=2-OB
C   Z(3,K,I)=2-EST
C   Z(4,K,I)=(2-GALL)*(2-EST)
C   Z(5,K,I)=(2-OB)*(2-EST)
C   Z(6,K,I)=(2-GALL)*(2-OB)
C   READ(1,100,END=2000) GALL,OB,EST
6   CONTINUE
C   GOTO 1000
2000 CONTINUE
C   NS = I
C   DO 2 I=1,NP
C   B(I)=0.0
2   CALL MATCH(NS,NR,NRMAX,NM,I,NIT,B,Z,SCORE,COVI,COV,
C   * EZB,IVAR,NM1,EPS,W)
C   STOP
C   END

```

SUBPROGRAM MATCH

```

C
C REFERENCES :
C N. E. BRESLOW,N.E. DAY,K. T. HALVORSEN, R. L. PRENTICE,C. SABAI :
C ESTIMATION OF MULTIPLE RELATIVE RISK FUNCTIONS IN MATCHED
C CASE-CONTROL STUDIES. AMERICAN JOURNAL OF EPIDEMIOLOGY
C VOL108,NO4, P 299-307 , 1978
C
C THIS SUBROUTINE COMPUTES A LINEAR LOGISTIC REGRESSION ANALYSIS FOR
C MATCHED SETS OF 1 CASE A VARIABLE NO. OF CONTROLS PER CASE
C THE VARIABLES APPEARING IN THE CALL STATEMENT ARE DEFINED AS FOLLO
C NS NUMBER OF MATCHED SETS
C NR VECTOR OF NO. OF CONTROLS IN EACH SET + 1
C NRMAX (MAX NO. OF CONTROLS PER CASE)+1
C NM MAXIMUM NUMBER OF VARIABLES TO BE ANALYZED
C NP NUMBER OF VARIABLES ANALYZED IN THIS RUN
C NIT MAXIMUM NUMBER OF ITERATIONS OF THE NEWTON-RAPHSON TYPE
C B PARAMETER VECTOR OF LENGTH NM
C Z NM BY NRMAX BY NS MATRIX CONTAINING COVARIATES
C SCORE FIRST DERIVATIVE OF THE LN-LIKELIHOOD OF LENGTH NM
C COVI INFORMATION MATRIX(2ND DERIVATIVE OF LN-LIKELIHOOD)
C COV INVERSE INFORMATION MATRIX(ESTIMATED COVARIANCE MATRIX)
C EZB WORKING VECTOR OF LENGTH NRMAX
C IVAR VECTOR OF VARIABLES USED IN THIS RUN (DIMENSION NM)
C NM1 = NM*(NM+1)/2 DIMENSION OF COVI AND COV
C EPS CHANGE IN LIKELIHOOD BELOW WHICH ITERATION STOPS
C W WORKING VECTOR OF LENGTH NM
C
C NOTE(Z,J,K,I) IS THE VALUE OF THE JTH COVARIATE FOR THE KTH
C MEMBER IN THE ITH SET.IT IS ASSUMED THAT THE FIRST MEMBER IS THE
C CASE AND THAT THE REMAINING NR(I)-1 MEMBERS ARE CONTROLS.
C
C NOTE(Z MUST BE DIMENSIONED TO HAVE NM ROWS,NRMAX COLUMNS
C AND AT LEAST NS SLICES IN THE MAIN PROGRAM,
C COVI AND COV ARE ARRAYS OF LENGTH NM*(NM+1)/2 SINCE THEY USE
C THE SYMMETRIC STORAGE MODE.
C
C SUBROUTINE MATCH(NS,NR,NRMAX,NM,NP,NIT,B,Z,SCORE,COVI,COV,
* EZB,IVAR,NM1,EPS,W)
C DIMENSION Z(NM,NRMAX,NS),B(NM),EZB(NRMAX),SCORE(NM),COV(NM1),
* COVI(NM1),IVAR(NM),NR(NS),W(NM)
C REAL LOGLIK
C DATA TEST/1.0/,ISUB/1/
C WRITE(6,100)
100 FORMAT(///' LOGISTIC REGRESSION ANALYSIS FOR MATCHED SETS',/)
C WRITE(6,101)NS
101 FORMAT(1H , 'NUMBER OF MATCHED SETS',I4)
C WRITE(6,102)((I,NR(I)),I=1,NS)
102 FORMAT(1H , 'SET NUMBER AND NUMBER OF MEMBERS',
* ' IN EACH SET,(INCLUDING CASE)',/,50(1X,10(I4,I3)),/)
C WRITE(6,103)NP,(IVAR(J),J=1,NP)

```

```

103  FORMAT(1H , 'NUMBER OF VARIABLES IN THIS ANALYSIS=', I3,
      */, ' THESE VARIABLES ARE NUMBERS', 30I3)
      WRITE(6,104)NIT
104  FORMAT(1H , 'MAXIMUM NUMBER OF ITERATIONS', I4)
      WRITE(6,106)
106  FORMAT('  ITER  LOG-LIKELIHOOD   SCORE', 8X, 'PARAMETER ESTIMATES')
      ITS=0
      IF(ISUB.NE.1)GOTO 1
C     SUBTRACT VALUE OF COVARIATES FOR THE CASE FROM THOSE OF CONTROLS
C     THIS ONLY NEEDS TO BE DONE AT THE FIRST CALL OF THE SUBROUTINE
      L=1
      DO 99 I=1,NS
      NRI=NR(I)
      DO 99 J=2,NRI
      DO 99 K=1,NM
      Z(K,J,I)=Z(K,J,I)-Z(K,L,I)
99   CONTINUE
      ISUB=0
1    CONTINUE
      ITS= ITS+1
C   CLEAR ARRAYS PRIOR TO NEXT ITERATION
      LOGLIK=0.0
      K=0
      DO 2 J=1,NP
      SCORE(J)=0.0
      DO 2 JJ=J,NP
      K=K+1
2    COVI(K)=0.0
C   CALCULATE LOGLIK, SCORE, COVI
      DO 7 I=1,NS
      DENOM=0.0
      NRI=NR(I)
      DO 4 K=2,NRI
      ZB=0.0
      DO 3 J=1,NP
      L=IVAR(J)
3     ZB=ZB+B(J)*Z(L,K,I)
      EZB(K)=EXP(ZB)
4     DENOM=DENOM+EZB(K)
      DENOM=1.0+DENOM
      LOGLIK=LOGLIK-ALOG(DENOM)
      KJ=0
      DO 7 J=1,NP
      L=IVAR(J)
      EZJ=0.0
      NRI=NR(I)
      DO 5 K=2,NRI
5     EZJ=EZJ+EZB(K)*Z(L,K,I)
      SCORE(J)=SCORE(J)-EZJ/DENOM

```

```

DO 7 JJ=1,J
LL=IVAR(JJ)
EZJJ=0.0
EZZ=0.0
NRI=NR(I)
DO 6 K=2,NRI
EZJJ=EZJJ+EZB(K)*Z(LL,K,I)
6  EZZ=EZB(K)*Z(L,K,I)*Z(LL,K,I)+EZZ
KJ=KJ+1
COVI(KJ)=COVI(KJ)+(EZZ/DENOM)-EZJ*EZJJ/(DENOM*DENOM)
7  CONTINUE
DO 500 I=1,NM1
COV(I)=COVI(I)
500 CONTINUE
CALL SINV(COV,NP,EPS,IER)
IF (IER.NE.0) WRITE (6,501)IER
501 FORMAT(' IER = ',I3)
TEMP=0.0
K=0
DO 51 J=1,NP
DO 51 JJ=1,J
K=K+1
G=1.0
IF (J.NE.JJ) G=2.0
TEMP=TEMP+G*SCORE(J)*SCORE(JJ)*COV(K)
51 CONTINUE
C  WRITE OUT ITERATION NUMBER AND LOG-LIKELIHOOD AND B'S
WRITE(6,107)ITS,LOGLIK,TEMP,(B(J),J=1,NP)
107  FORMAT(1X,I4,F14.4,F11.3,2X,8F12.4,20(/,21X,8F12.4)
C  TEST FOR CONVERGENCE
TEST=ABS(TEST-LOGLIK)
IF(TEST.LE.EPS.OR.ITS.GE.NIT)GO TO 9
TEST=LOGLIK
C  CALCULATE NEW VALUE OF PARAMETER ESTIMATE AND REPEAT
C  CALL SYMINV(COVI,NP,COV,W,NULLTY,IFALT,NM1)
DO 180 I=1,NP
W(I)=0.0
DO 181 J=1,I
K=I*(I-1)/2+J
181  W(I)=W(I)+SCORE(J)*COV(K)
I1=I+1
IF(I1.GT.NP)GOTO 180
DO 182 K=I1,NP
J=K*(K-1)/2+I
182  W(I)=W(I)+SCORE(K)*COV(J)
180  CONTINUE
DO 183 I=1,NP
183  SCORE(I)=W(I)
DO 8 J=1,NP
8  B(J)=B(J)+SCORE(J)

```

```
C  UPON CONVERGENCE OF MAXIMUM ITERATIONS WRITE OUT RESULTS
      GO TO 1
9      WRITE(6,108)(B(J),J=1,NP)
108     FORMAT(' ESTIMATED PARAMETER VECTOR',(/1X,10F12.6))
      WRITE(6,109)(SCORE(J),J=1,NP)
109     FORMAT(' FIRST DERIVATIVE LOG-LIKELIHOOD',(/1X,10F12.6))
      WRITE(6,110)
110     FORMAT(' INFORMATION MATRIX')
      DO 10 J=1,NP
      K=J*(J-1)/2+1
      JJ=J*(J+1)/2
10      WRITE(6,111) (COVI(I),I=K,JJ)
111     FORMAT(1X,10F12.6)
C  INVERT INFORMATION MATRIX AND WRITE OUT
C      CALL SYMINV(COVI,NP,COV,W,NULLTY,IFAUULT,NM1)
      DO 600 I=1,NM1
      COV(I)=COVI(I)
600     CONTINUE
      CALL SINVC(COV,NP,EPS,IER)
      IF (IER.NE.0) WRITE (6,501)IER
      WRITE(6,112)
112     FORMAT(' ESTIMATED COVARIANCE MATRIX')
      DO 11 J=1,NP
      K=J*(J-1)/2+1
      JJ=J*(J+1)/2
      SCORE(J)=B(J)/SQRT(COV(JJ))
11      WRITE(6,111)(COV(I),I=K,JJ)
      WRITE(6,113)(SCORE(J),J=1,NP)
113     FORMAT(' STANDARDIZED REGRESSION COEFFICIENTS',(/1X,100F12.6))
      RETURN
      END
```

MODEL WITH SINGLE VARIABLE: GALL

LOGISTIC REGRESSION ANALYSIS FOR MATCHED SETS

NUMBER OF MATCHED SETS 63

SET NUMBER AND NUMBER OF MEMBERS IN EACH SET,(INCLUDING CASE)

1	5	2	5	3	5	4	5	5	5	6	5	7	5	8	5	9	5	10	5
11	5	12	5	13	5	14	5	15	5	16	5	17	5	18	5	19	5	20	5
21	5	22	5	23	5	24	5	25	5	26	5	27	5	28	5	29	5	30	5
31	5	32	5	33	5	34	5	35	5	36	5	37	5	38	5	39	5	40	5
41	5	42	5	43	5	44	5	45	5	46	5	47	5	48	5	49	5	50	5
51	5	52	5	53	5	54	5	55	5	56	5	57	5	58	5	59	5	60	5
61	5	62	5	63	5														

NUMBER OF VARIABLES IN THIS ANALYSIS = 1

THESE VARIABLES ARE NUMBERS 1

MAXIMUM NUMBER OF ITERATIONS 10

ITER	LOG-LIKELIHOOD	SCORE	PARAMETER ESTIMATES
1	-101.3945	13.829	0.0
2	-95.6566	0.512	1.5714
3	-95.4043	0.000	1.3011
4	-95.4043	0.000	1.3061

ESTIMATED PARAMETER VECTOR

1.306142

FIRST DERIVATIVE LOG-LIKELIHOOD

0.000010

INFORMATION MATRIX

7.232518

ESTIMATED COVARIANCE MATRIX

0.138264

STANDARDIZED REGRESSION COEFFICIENTS

3.512652

MODEL WITH 2 VARIABLES: GALL + OB

LOGISTIC REGRESSION ANALYSIS FOR MATCHED SETS

NUMBER OF MATCHED SETS 63

SET NUMBER AND NUMBER OF MEMBERS IN EACH SET, (INCLUDING CASE)

1	5	2	5	3	5	4	5	5	5	6	5	7	5	8	5	9	5	10	5
11	5	12	5	13	5	14	5	15	5	16	5	17	5	18	5	19	5	20	5
21	5	22	5	23	5	24	5	25	5	26	5	27	5	28	5	29	5	30	5
31	5	32	5	33	5	34	5	35	5	36	5	37	5	38	5	39	5	40	5
41	5	42	5	43	5	44	5	45	5	46	5	47	5	48	5	49	5	50	5
51	5	52	5	53	5	54	5	55	5	56	5	57	5	58	5	59	5	60	5
61	5	62	5	63	5														

NUMBER OF VARIABLES IN THIS ANALYSIS = 2

THESE VARIABLES ARE NUMBERS 1 2

MAXIMUM NUMBER OF ITERATIONS 10

ITER	LOG-LIKELIHOOD	SCORE	PARAMETER ESTIMATES	
1	-95.4043	5.031	1.3061	0.0
2	-92.8562	0.016	1.2673	0.7044
3	-92.8483	0.000	1.3085	0.7254
4	-92.8484	0.000	1.3086	0.7255

ESTIMATED PARAMETER VECTOR

1.308584 0.725526

FIRST DERIVATIVE LOG-LIKELIHOOD

-0.000000 0.000001

INFORMATION MATRIX

7.115846

-0.413175 9.369903

ESTIMATED COVARIANCE MATRIX

0.140892

0.006213 0.106999

STANDARDIZED REGRESSION COEFFICIENTS

3.486248 2.218012

MODEL WITH 3 VARIABLES: GALL + OB + EST
(model 4, Table 7.7)

LOGISTIC REGRESSION ANALYSIS FOR MATCHED SETS

NUMBER OF MATCHED SETS 63

SET NUMBER AND NUMBER OF MEMBERS IN EACH SET,(INCLUDING CASE)

1	5	2	5	3	5	4	5	5	5	6	5	7	5	8	5	9	5	10	5
11	5	12	5	13	5	14	5	15	5	16	5	17	5	18	5	19	5	20	5
21	5	22	5	23	5	24	5	25	5	26	5	27	5	28	5	29	5	30	5
31	5	32	5	33	5	34	5	35	5	36	5	37	5	38	5	39	5	40	5
41	5	42	5	43	5	44	5	45	5	46	5	47	5	48	5	49	5	50	5
51	5	52	5	53	5	54	5	55	5	56	5	57	5	58	5	59	5	60	5
61	5	62	5	63	5														

NUMBER OF VARIABLES IN THIS ANALYSIS = 3

THESE VARIABLES ARE NUMBERS 1 2 3

MAXIMUM NUMBER OF ITERATIONS 10

ITER	LOG-LIKELIHOOD	SCORE	PARAMETER ESTIMATES		
1	-92.8484	26.837	1.3086	0.7255	0.0
2	-78.4750	1.213	1.0256	0.4851	1.6166
3	-77.8264	0.017	1.2543	0.5086	1.9860
4	-77.8179	0.000	1.2746	0.5113	2.0394
5	-77.8179	0.000	1.2748	0.5113	2.0403

ESTIMATED PARAMETER VECTOR

1.274839 0.511342 2.040298

FIRST DERIVATIVE LOG-LIKELIHOOD

0.000004 0.000002 0.000004

INFORMATION MATRIX

5.996700

-0.295261 7.917870

-0.581349 0.554122 5.222640

ESTIMATED COVARIANCE MATRIX

0.168775

0.005016 0.127390

0.018255 -0.012958 0.194881

STANDARDIZED REGRESSION COEFFICIENTS

3.103139 1.432659 4.621776

MODEL WITH 4 VARIABLES: GALL + OB + EST + GALL×EST

(model 8, Table 7.5)

LOGISTIC REGRESSION ANALYSIS FOR MATCHED SETS

NUMBER OF MATCHED SETS 63

SET NUMBER AND NUMBER OF MEMBERS IN EACH SET, (INCLUDING CASE)

1	5	2	5	3	5	4	5	5	5	6	5	7	5	8	5	9	5	10	5
11	5	12	5	13	5	14	5	15	5	16	5	17	5	18	5	19	5	20	5
21	5	22	5	23	5	24	5	25	5	26	5	27	5	28	5	29	5	30	5
31	5	32	5	33	5	34	5	35	5	36	5	37	5	38	5	39	5	40	5
41	5	42	5	43	5	44	5	45	5	46	5	47	5	48	5	49	5	50	5
51	5	52	5	53	5	54	5	55	5	56	5	57	5	58	5	59	5	60	5
61	5	62	5	63	5														

NUMBER OF VARIABLES IN THIS ANALYSIS = 4

THESE VARIABLES ARE NUMBERS 1 2 3 4

MAXIMUM NUMBER OF ITERATIONS 10

ITER LOG-LIKELIHOOD SCORE PARAMETER ESTIMATES

1	-77.8179	4.392	1.2748	0.5113	2.0403	0.0
2	-75.9031	0.221	3.0331	0.4859	2.5096	-2.2027
3	-75.7909	0.000	2.8467	0.4901	2.6172	-2.0003
4	-75.7910	0.000	2.8446	0.4901	2.6206	-1.9975

ESTIMATED PARAMETER VECTOR

2.844563 0.490130 2.620620 -1.997451

FIRST DERIVATIVE LOG-LIKELIHOOD

-0.000001 0.000004 0.000020 0.000006

INFORMATION MATRIX

6.692919			
-0.287091	7.667604		
-1.389089	0.432648	4.504435	
4.758762	-0.133908	0.584442	4.981014

ESTIMATED COVARIANCE MATRIX

0.774399			
-0.003830	0.131275		
0.340368	-0.014950	0.376377	
-0.779885	0.008943	-0.369744	0.989473

STANDARDIZED REGRESSION COEFFICIENTS

3.232460 1.352757 4.271620 -2.008047

APPENDIX V

LISTING OF PROGRAM STRAT

The following main program and subroutine perform the conditional logistic regression analysis for matched sets consisting of a variable number of cases and a variable number of controls, using the general likelihood (7.1). It was written by Mr Peter Smith as a generalization of the program MATCH. To illustrate its operation we have used exactly the same data as used with MATCH, namely those corresponding to several of the models in Table 7.7. Slight differences in the numerical output are due to the use of different computing formulae.

The program is completely self-contained, with the subroutines SYMINV, CHOL, and TWIDL which it calls being listed also.

MAIN PROGRAM

```

    DIMENSION NCT(63),NCA(63),IVAR(6),IP(7),W(20),WW(20),
* Z(6,5,63),B(6),SCORE(6),COV(21),COVI(21)
    DATA NCA/63*1/,NCT/63*4/,NS,NM,NMAX,NIT,EPS/63,6,5,10,0.0001/
    DATA B/6*0.0/,IVAR/1,2,3,4,5,6/,NP/6/
    NMAX2=NMAX+2
    NM1=NM*(NM+1)/2
    DO 6 I=1,NS
    N=NCT(I)+NCA(I)
    DO 6 J=1,N
    L=N+1-J
    READ(1,100) GALL,OB,EST
100  FORMAT(10X,F5.0,5X,F5.0,15X,F5.0)
    Z(1,L,I)=2-GALL
    IF (OB.EQ.9) OB=2
    Z(2,L,I)=2-OB
    Z(3,L,I)=2-EST
    Z(4,L,I)=(2-GALL)*(2-EST)
    Z(5,L,I)=(2-OB)*(2-EST)
    Z(6,L,I)=(2-GALL)*(2-OB)
6    CONTINUE
    DO 2 I=1,NP
    B(I)=0.0
2    CALL STRAT(NS,NCA,NCT,NMAX,NM,I,NIT,B,Z,SCORE,COVI,COV,
* IVAR,EPS,NM1,W,WW,IP,NMAX2)
99   STOP
    END

```

SUBPROGRAM STRAT

```
SUBROUTINE STRAT(NS,NCA,NCT,NMAX,NM,NP,NIT,B,Z,SCORE,COVI,COV,  
*IVAR,EPS,NM1,W,WW,IP,NMAX2)
```

```
C  THIS SUBROUTINE COMPUTES A LINEAR LOGISTIC REGRESSION ANALYSIS  
C  FOR STRATIFIED SETS CONSISTING OF NCA(I) CASES AND NCT(I) CONTROLS  
C  THE ITH STRATUM  
C  THE VARIABLES APPEARING IN THE CALL STATEMENT ARE DEFINED AS FOLLO  
C  NS  NUMBER OF STRATA  
C  NCA  VECTOR OF THE NUMBER OF CASES IN EACH STRATUM  
C  NCT  VECTOR OF THE NUMBER OF CONTROLS IN EACH STRATUM  
C  NMAX  MAXIMUM NUMBER OF CASES + CONTROLS IN ANY STRATUM  
C  NM  MAXIMUM NUMBER OF VARIABLES TO BE ANALYSED  
C  NP  NUMBER OF VARIABLES TO BE ANALYSED IN THIS RUN  
C  NIT  MAXIMUM NUMBER OF ITERATIONS OF THE NEWTON-RAPHSON TYPE  
C  B  PARAMETER VECTOR OF LENGTH NM  
C  Z  NM BY NMAX BY NS MATRIX CONTAINING COVARIATES  
C  SCORE  FIRST DERIVATIVE OF THE LN-LIKELIHOOD OF LENGTH NM  
C  COVI  INFORMATION MATRIX (2ND DERIVATIVE OF LN-LIKELIHOOD)  
C  COV  INVERSE INFORMATION MATRIX : ESTIMATED COVARIANCE MATRIX  
C  IVAR  VECTOR OF VARIABLE NUMBERS USED IN THIS RUN  
C  EPS  CHANGE IN LIKELIHOOD BELOW WHICH ITERATION STOPS  
C  NM1   $NM*(NM+1)/2$   
C  WW  WORKING VECTOR OF LENGTH NM  
C  IP  WORKING VECTOR OF LENGTH NMAX2  
C  NMAX2  NMAX+2  
C  NOTE: IN THE ARRAY Z INDIVIDUALS ARE INDEXED BY THE 2ND  
C  SUBSCRIPT, WITHIN EACH STRATUM DATA ON CONTROLS MUST BE STORED  
C  FIRST FOLLOWED BY THOSE ON CASES.
```

C REFERENCES:

C N. E. BRESLOW, N. E. DAY, K. T. HALVORSEN, R. L. PRENTICE, C. SABAI:

C ESTIMATION OF MULTIPLE RELATIVE RISK FUNCTIONS IN MATCHED

C CASE CONTROL STUDIES . AM. J EPIDEMIOL 108:299-307, 1978

C DIMENSION Z(NM,NMAX,NS),B(NM),SCORE(NM),COV(NM1),COVI(NM1)

C *,IVAR(NM),NCA(NS),NCT(NS),W(NM),WW(NM),IP(NMAX2)

REAL LOGLIK

DATA TEST/1.0/

WRITE(6,100)

100 FORMAT(1H1,///' LOGISTIC REGRESSION ANALYSIS IN STRATA',/)

WRITE(6,101)NS

101 FORMAT(1H,'NUMBER OF STRATA',I4)

WRITE(6,102)(I,NCA(I),NCT(I),I=1,NS)

102 FORMAT(1H,'STRATUM NUMBER AND NUMBERS OF CASES AND CONTROLS',

*50(1X,10(I4,I3,I3),/))

WRITE(6,103)NP,(IVAR(J),J=1,NP)

103 FORMAT(1H,'NUMBER OF VARIABLES IN THIS ANALYSIS',I3,/,

* ' THESE VARIABLES ARE',30I3)

WRITE(6,104)NIT

104 FORMAT(1H,'MAXIMUM NUMBER OF ITERATIONS',I4)

WRITE(6,106)

106 FORMAT(1H,'ITER LOG-LIKELIHOOD SCORE PARAMETER ESTIMATES')

ITS=0

1 ITS=ITS+1

LOGLIK=0.0

K=0

DO 2 J=1,NP

SCORE(J)=0.0

DO 2 JJ=1,J

K=K+1

2 COVI(K)=0.0

C CALCULATE LOGLIK, SCORE AND COVI

DO 27 I=1,NS

IFG=0

SX=0.0

K=0

DO 8 J=1,NP

W(J)=0.0

WW(J)=0.0

DO 8 JJ=1,J

K=K+1

8 COV(K)=0.0

M=NCA(I)

N=M+NCT(I)

XX=1.0

X=0.0

IP(1)=N+1

IP(N+2)=-2

NMX=N-M+1

DO 30 J=2,NMX

```

30  IP(J)=0
    DO 31 J=1,M
31  IP(NMX+J)=J
    KK=N-M+1
    DO 12 J=KK,N
    DO 12 K=1,NP
    L=IVAR(K)
    W(K)=W(K)+Z(L,J,I)
12  X=X+B(K)*Z(L,J,I)
11  XX=EXP(X)
    SX=SX+XX
    IF(IFG.NE.0)GOTO 13
    LOGLIK=LOGLIK+ALOG(SX)
    DO 17 K=1,NP
17  SCORE(K)=SCORE(K)+W(K)
    IFG=1
13  L=0
    DO 14 K=1,NP
    WW(K)=WW(K)+XX*W(K)
    DO 14 KK=1,K
    L=L+1
14  COV(L)=COV(L)+XX*W(K)*W(KK)
    ID=0
    CALL TWIDL(IPS,IM,IZ,ID,IP,NMAX2)
    IF(ID.EQ.1)GOTO 20
    DO 15 K=1,NP
    L=IVAR(K)
    ZC=Z(L,IPS,I)-Z(L,IM,I)
    W(K)=W(K)+ZC.
15  X=X+B(K)*ZC
    GOTO 11
20  LOGLIK=LOGLIK-ALOG(SX)
    L=0
    DO 21 J=1,NP
    SCORE(J)=SCORE(J)-WW(J)/SX
    DO 21 K=1,J
    L=L+1
21  COVI(L)=COVI(L)+(SX*COV(L)-WW(J)*WW(K))/SX**2
27  CONTINUE

```

```

CALL SYMINV(COVI,NP,COV,W,NULLTY,IFAUULT,NM1)
TEMP=0.0
K=0
DO 51 J=1,NP
DO 51 JJ=1,J
K=K+1
G=1.0
IF (J.NE.JJ) G=2.0
TEMP=TEMP+G*SCORE(J)*SCORE(JJ)*COV(K)
51 CONTINUE
C WRITE OUT ITERATION NUMBER AND LOG-LIKELIHOOD AND B'S
WRITE (6,107)ITS,LOGLIK,TEMP,(B(J),J=1,NP)
107 FORMAT(1X,I4,F14.4,F11.3,2X,8F12.4,20(/,21X,8F12.4))
C TEST FOR CONVERGENCE
TEST=ABS(TEST-LOGLIK)
IF (TEST.LE.EPS.OR.ITS.GE.NIT)GOTO 9
TEST=LOGLIK
C CALCULATE NEW VALUE OF PARAMETER ESTIMATE AND REPEAT
DO 180 I=1,NP
W(I)=0.0
DO 181 J=1,I
K=I*(I-1)/2+J
181 W(I)=W(I)+SCORE(J)*COV(K)
I1=I+1
IF(I1.GT.NP)GOTO 180
DO 182 K=I1,NP
J=K*(K-1)/2+I
182 W(I)=W(I)+SCORE(K)*COV(J)
180 CONTINUE
DO 183 I=1,NP
183 SCORE(I)=W(I)
DO 18 J=1,NP
18 B(J)=B(J)+SCORE(J)
C UPON CONVERGENCE OR MAXIMUM ITERATIONS WRITE OUT RESULTS
GOTO 1
9 WRITE(6,108)(B(J),J=1,NP)
108 FORMAT(1H0,'ESTIMATED PARAMETER VECTOR',(/1X,10F12.6))
WRITE(6,109)(SCORE(J),J=1,NP)
109 FORMAT(1H0,'FIRST DERIVATIVE LOG-LIKELIHOOD',(/1X,10F12.6))
WRITE(6,110)
110 FORMAT(1H0,'INFORMATION MATRIX')
DO 10 J=1,NP
K=J*(J-1)/2+1
JJ=J*(J+1)/2
10 WRITE (6,111) (COVI(I),I=K,JJ)
111 FORMAT(1X,10F12.6)
C INVERT INFORMATION MATRIX AND WRITE OUT
CALL SYMINV(COVI,NP,COV,W,NULLTY,IFAUULT,NM1)
WRITE(6,112)

```

```
112  FORMAT(1H0,'ESTIMATED COVARIANCE MATRIX')
      DO 19 J=1,NP
      K=J*(J-1)/2+1
      JJ=J*(J+1)/2
      SCORE(J)=B(J)/SQRT(COV(JJ))
19   WRITE(6,111)(COV(I),I=K,JJ)
      WRITE(6,113)(SCORE(J),J=1,NP)
113  FORMAT(1H0,'STNDIZED REGRESSION COEFFICIENTS',(/1X,100F12.6))
      RETURN
      END
```

```

SUBROUTINE SYMINV(A,N,C,W, NULLTY, IFAULT, NFAC)
DIMENSION A(NFAC), C(NFAC), W(N)
C
C   ALGORITHM AS7 J.R. STATIST.SOC.C.(1968) VOL.17,NO.2
C
  NROW=N
  IFAULT=1
  IF(NROW.LE.0)GOTO 100
  IFAULT=0
  CALL CHOL(A,NROW,C, NULLTY, IFAULT, NFAC)
  IF(IFAULT.NE.0)GOTO 100
  NN=(NROW*(NROW+1))/2
  IROW=NROW
  NDIAG=NN
16  IF(C(NDIAG).EQ.0.90)GOTO 11
  L=NDIAG
  DO 10 I=IROW,NROW
    W(I)=C(L)
    L=L+1
10  CONTINUE
  ICOL=NROW
  JCOL=NN
  MDIAG=NN
15  L=JCOL
  X=0.0
  IF(ICOL.EQ.IROW)X=1.0/W(IROW)
  K=NROW
13  IF(K.EQ.IROW)GOTO12
  X=X-W(K)*C(L)
  K=K-1
  L=L-1
  IF(L.GT.MDIAG)L=L-K+1
  GOTO 13
12  C(L)=X/W(IROW)
  IF(ICOL.EQ.IROW)GOTO 14
  MDIAG=MDIAG-ICOL
  ICOL=ICOL-1
  JCOL=JCOL-1
  GOTO 15
11  L=NDIAG
  DO 17 J=IROW,NROW
    C(L)=0.0
    L=L+J
17  CONTINUE
14  NDIAG=NDIAG-IROW
  IROW=IROW-1
  IF(IROW.NE.0)GOTO 16
100 RETURN
END

```

```
SUBROUTINE CHOL(A,N,U, NULLTY, IFAULT, NFAC)
DIMENSION A(NFAC), U(NFAC)
REAL ETA
DATA ETA/0.000001/
C
C ALGORITHM AS6 J.R.STATIS. SOC. C. (1968). VOL.17, NO.2
C
  IFAULT=1
  IF(N.LE.0)GOTO 100
  IFAULT=2
  NULLTY=0
  J=1
  K=0
  DO 10 ICOL=1,N
    L=0
    DO 11 IROW=1,ICOL
      K=K+1
      W=A(K)
      M=J
      DO 12 I=1, IROW
        L=L+1
        IF(I.EQ.IROW-GOTO 13
        W=W-U(L)*U(M)
        M=M+1
12      CONTINUE
13      IF(IROW.EQ.ICOL)GOTO 14
        IF(U(L).EQ.0.0)GOTO 21
        U(K)=W/U(L)
        GOTO 11
21      U(K)=0.0
11      CONTINUE
14      IF(ABS(W).LT.ABS(ETA*A(K))) GOTO 20
        IF(W.LT.0.0)GOTO 100
        U(K)=SQRT(W)
        GOTO 15
20      U(K)=0.0
        NULLTY=NULLTY+1
15      J=J+ICOL
10      CONTINUE
        IFAULT=0
100     RETURN
        END
```

```

SUBROUTINE TWIDL(X,Y,Z,DONE,P,N2)
INTEGER X,Y,Z,DONE,N2,P(N2)

C
C   THIS SUBROUTINE IS A FORTRAN VERSION OF CACM 382
C   FOR SELECTING ALL POSSIBLE COMBINATIONS OF M THINGS
C   OUT OF N OBJECTS. OUTSIDE OF THE SUBROUTINE
C   INITIALLY SET P(1)=N+1,P(N+2)=-2,P(2) TO
C   P(N-M+1)=0,P(N-M+2) TO P(N+1)=1 TO M
C   RESPECTIVELY. IF M=0,P(2)=1
C   INITIALLY SET DONE =0 AND THE S/R SETS THIS TO
C   1 WHEN ALL COMBINATIONS HAVE BEEN SELECTED
C   SUPPOSE THE N OBJECTS ARE IN A(1:N) AND
C   SUCCESSIVE COMBINATIONS ARE STORED IN C(1:M)
C   INITIALLY SET C(1) TO C(M)=A(N-M+1) TO A(N)
C   THEN CALL TWIDL IF DONE=1 ALL COMBINATIONS
C   HAVE BEEN GENERATED, OTHERWISE SET C(Z)=A(X)
C   ALTERNATIVELY STORE (N-M) 0'S AND M 1'S IN
C   B(1:N) IN THAT ORDER THEN CALL TWIDL.
C   IF DONE.NE.1 SET B(X)=1 AND B(Y)=0
C   AND CONTINUE IN THIS WAY UNTIL DONE=1
C
      J=0
1     J=J+1
      IF(P(J+1).LE.0)GOTO 1
      IF(P(J).NE.0)GOTO 5
      J1=J-1
      IF(J1.LT.2)GOTO 6
      DO 7 I=2,J1
      K=J1-I+2
7     P(K+1)=-1
6     P(J+1)=0
      P(2)=1
      X=1
      Z=1
      Y=J
      GOTO 4
5     CONTINUE
      IF(J.GT.1)P(J)=0
2     J=J+1
      IF(P(J+1).GT.0)GOTO 2
      I=J-1
      K=J-1
3     I=I+1
      IF(P(I+1).NE.0)GOTO 8
      P(I+1)=-1
      GOTO 3
8     CONTINUE

```

```
      IF(P(I+1).NE.-1)GOTO 9
      P(I+1)=P(K+1)
      Z=P(K+1)
      X=I
      Y=K
      P(K+1)=-1
      GOTO 4
9      CONTINUE
      IF(I.NE.P(1)) GOTO 10
      DONE=1
      GOTO 4
10     Z=P(I+1)
      P(J+1)=Z
      P(I+1)=0
      X=J
      Y=I
4      RETURN
      END
```

MODEL WITH SINGLE VARIABLE: GALL

LOGISTIC REGRESSION ANALYSIS IN STRATA

NUMBER OF STRATA 63

STRATUM NUMBER AND NUMBERS OF CASES AND CONTROLS

1	1	4	2	1	4	3	1	4	4	1	4	5	1	4	6	1	4	7	1	4	8	1	4
9	1	4	10	1	4	11	1	4	12	1	4	13	1	4	14	1	4	15	1	4	16	1	4
17	1	4	18	1	4	19	1	4	20	1	4	21	1	4	22	1	4	23	1	4	24	1	4
25	1	4	26	1	4	27	1	4	28	1	4	29	1	4	30	1	4	31	1	4	32	1	4
33	1	4	34	1	4	35	1	4	36	1	4	37	1	4	38	1	4	39	1	4	40	1	4
41	1	4	42	1	4	43	1	4	44	1	4	45	1	4	46	1	4	47	1	4	48	1	4
49	1	4	50	1	4	51	1	4	52	1	4	53	1	4	54	1	4	55	1	4	56	1	4
57	1	4	58	1	4	59	1	4	60	1	4	61	1	4	62	1	4	63	1	4			

NUMBER OF VARIABLES IN THIS ANALYSIS 1

THESE VARIABLES ARE 1

MAXIMUM NUMBER OF ITERATIONS 10

ITER LOG-LIKELIHOOD SCORE PARAMETER ESTIMATES

1	-101.3945	13.829	0.0
2	-95.6567	0.512	1.5714
3	-95.4042	0.000	1.3011
4	-95.4041	0.000	1.3061
5	-95.4041	0.000	1.3061

ESTIMATED PARAMETER VECTOR

1.306143

FIRST DERIVATIVE LOG-LIKELIHOOD

0.000002

INFORMATION MATRIX

7.232529

ESTIMATED COVARIANCE MATRIX

0.138264

STNDIZED REGRESSION COEFFICIENTS

3.512656

MODEL WITH 2 VARIABLES: GALL + OB

LOGISTIC REGRESSION ANALYSIS IN STRATA

NUMBER OF STRATA 63

STRATUM NUMBER AND NUMBERS OF CASES AND CONTROLS

1	1	4	2	1	4	3	1	4	4	1	4	5	1	4	6	1	4	7	1	4	8	1	4
9	1	4	10	1	4	11	1	4	12	1	4	13	1	4	14	1	4	15	1	4	16	1	4
17	1	4	18	1	4	19	1	4	20	1	4	21	1	4	22	1	4	23	1	4	24	1	4
25	1	4	26	1	4	27	1	4	28	1	4	29	1	4	30	1	4	31	1	4	32	1	4
33	1	4	34	1	4	35	1	4	36	1	4	37	1	4	38	1	4	39	1	4	40	1	4
41	1	4	42	1	4	43	1	4	44	1	4	45	1	4	46	1	4	47	1	4	48	1	4
49	1	4	50	1	4	51	1	4	52	1	4	53	1	4	54	1	4	55	1	4	56	1	4
57	1	4	58	1	4	59	1	4	60	1	4	61	1	4	62	1	4	63	1	4			

NUMBER OF VARIABLES IN THIS ANALYSIS 2

THESE VARIABLES ARE 1 2

MAXIMUM NUMBER OF ITERATIONS 10

ITER LOG-LIKELIHOOD SCORE PARAMETER ESTIMATES

1	-95.4041	5.031	1.3061	0.0
2	-92.8559	0.016	1.2673	0.7044
3	-92.8481	0.000	1.3085	0.7254
4	-92.8483	0.000	1.3086	0.7255
5	-92.8483	0.000	1.3086	0.7255

ESTIMATED PARAMETER VECTOR

1.308584 0.725525

FIRST DERIVATIVE LOG-LIKELIHOOD

-0.000006 0.000004

INFORMATION MATRIX

7.115862	
-0.413175	9.369920

ESTIMATED COVARIANCE MATRIX

0.140892	
0.006213	0.106998

STNDIZED REGRESSION COEFFICIENTS

3.486253 2.218013

MODEL WITH 3 VARIABLES: GALL + OB + EST

(Model 4, Table 7.7)

LOGISTIC REGRESSION ANALYSIS IN STRATA

NUMBER OF STRATA 63

STRATUM NUMBER AND NUMBERS OF CASES AND CONTROLS

1	1	4	2	1	4	3	1	4	4	1	4	5	1	4	6	1	4	7	1	4	8	1	4
9	1	4	10	1	4	11	1	4	12	1	4	13	1	4	14	1	4	15	1	4	16	1	4
17	1	4	18	1	4	19	1	4	20	1	4	21	1	4	22	1	4	23	1	4	24	1	4
25	1	4	26	1	4	27	1	4	28	1	4	29	1	4	30	1	4	31	1	4	32	1	4
33	1	4	34	1	4	35	1	4	36	1	4	37	1	4	38	1	4	39	1	4	40	1	4
41	1	4	42	1	4	43	1	4	44	1	4	45	1	4	46	1	4	47	1	4	48	1	4
49	1	4	50	1	4	51	1	4	52	1	4	53	1	4	54	1	4	55	1	4	56	1	4
57	1	4	58	1	4	59	1	4	60	1	4	61	1	4	62	1	4	63	1	4			

NUMBER OF VARIABLES IN THIS ANALYSIS 3

THESE VARIABLES ARE 1 2 3

MAXIMUM NUMBER OF ITERATIONS 10

ITER	LOG-LIKELIHOOD	SCORE	PARAMETER	ESTIMATES
1	-92.8483	26.837	1.3086	0.7255 0.0
2	-78.4745	1.213	1.0256	0.4851 1.6166
3	-77.8259	0.017	1.2543	0.5085 1.9860
4	-77.8176	0.000	1.2746	0.5113 2.0394
5	-77.8177	0.000	1.2748	0.5113 2.0403
6	-77.8178	0.000	1.2748	0.5113 2.0403

ESTIMATED PARAMETER VECTOR

1.274838 0.511341 2.040295

FIRST DERIVATIVE LOG-LIKELIHOOD

-0.000002 -0.000016 -0.000012

INFORMATION MATRIX

5.996716		
-0.295262	7.917883	
-0.581353	0.554115	5.222651

ESTIMATED COVARIANCE MATRIX

0.168775		
0.005016	0.127390	
0.018255	-0.012958	0.194880

STNDIZED REGRESSION COEFFICIENTS

3.103141 1.432656 4.621776

MODEL WITH 4 VARIABLES: GALL + OB + EST + GALL×EST

(Model 8, Table 7.7)

LOGISTIC REGRESSION ANALYSIS IN STRATA

NUMBER OF STRATA 63

STRATUM NUMBER AND NUMBERS OF CASES AND CONTROLS

1	1	4	2	1	4	3	1	4	4	1	4	5	1	4	6	1	4	7	1	4	8	1	4
9	1	4	10	1	4	11	1	4	12	1	4	13	1	4	14	1	4	15	1	4	16	1	4
17	1	4	18	1	4	19	1	4	20	1	4	21	1	4	22	1	4	23	1	4	24	1	4
25	1	4	26	1	4	27	1	4	28	1	4	29	1	4	30	1	4	31	1	4	32	1	4
33	1	4	34	1	4	35	1	4	36	1	4	37	1	4	38	1	4	39	1	4	40	1	4
41	1	4	42	1	4	43	1	4	44	1	4	45	1	4	46	1	4	47	1	4	48	1	4
49	1	4	50	1	4	51	1	4	52	1	4	53	1	4	54	1	4	55	1	4	56	1	4
57	1	4	58	1	4	59	1	4	60	1	4	61	1	4	62	1	4	63	1	4			

NUMBER OF VARIABLES IN THIS ANALYSIS 4

THESE VARIABLES ARE NUMBERS 1 2 3 4

MAXIMUM NUMBER OF ITERATIONS 10

ITER LOG-LIKELIHOOD SCORE PARAMETER ESTIMATES

1	-77.8178	4.392	1.2748	0.5113	2.0403	0.0
2	-75.9028	0.221	3.0330	0.4859	2.5096	-2.2027
3	-75.7904	0.000	2.8467	0.4901	2.6172	-2.0003
4	-75.7906	0.000	2.8446	0.4901	2.6206	-1.9974
5	-75.7907	0.000	2.8446	0.4901	2.6206	-1.9974
6	-75.7904	0.000	2.8446	0.4901	2.6206	-1.9975
7	-75.7906	0.000	2.8446	0.4901	2.6206	-1.9974
8	-75.7906	0.000	2.8446	0.4901	2.6206	-1.9974

ESTIMATED PARAMETER VECTOR

2.844558 0.490128 2.620618 -1.997445

FIRST DERIVATIVE LOG-LIKELIHOOD

0.000001 -0.000002 0.000000 0.000002

INFORMATION MATRIX

6.692931			
-0.287092	7.667618		
-1.389091	0.432643	4.504441	
4.758771	-0.133909	0.584442	4.981022

ESTIMATED COVARIANCE MATRIX

0.774393			
-0.003830	0.131275		
0.340365	-0.014950	0.376375	
-0.779880	0.008943	-0.369741	0.989468

STNDIZED REGRESSION COEFFICIENTS

3.232466 1.352753 4.271628 -2.008047

APPENDIX VI

LISTING OF PROGRAM LOGODD

The following main program and subroutine perform the regression analysis of a set of 2×2 tables as described in § 7.5. For illustration we have used the data from the Oxford Childhood Cancer Survey published by Kneale (1971) and listed in Appendix II. Two different models were fit, the first with just the coefficient β representing the overall log relative risk for radiation, and the second with both β and the coefficient γ_1 representing the linear change in the log relative risk with year. The option for an 'exact' analysis was selected, giving numerical results as shown in § 7.5. (The last line in Table 6.17 shows the results which are obtained using the option for an 'approximate' analysis.) Following the fitting of the two parameter model, the programme was instructed (IPRINT = 1) to print out the observed and fitted entries for each table, the variance associated with that table (equation 4.13), the contribution it makes to the log likelihood, the estimated value of the odds ratio, and the associated covariates of which the first is the constant term and the second the coded year.

MAIN PROGRAM

```

      INTEGER P
      DOUBLE PRECISION INFO(10,10),LOGLIK,B(10),SCORE(10),SCORE1(10)
      DOUBLE PRECISION Z(10,120),T(4,120),PROB,E,V,DETER,CHIS,ULL
      REAL FMT(60),RT(4),RZ(10)
      DIMENSION IVAR(10)
      DATA NM/10/
      EQUIVALENCE(P,NP)
      EPS=1. E-4
1000 READ(5,150,END=900) P,NFORMT,ITS,ITAPE,IME,IBETA,IPRINT
      150 FORMAT(15I5)
      999 INF=20*NFORMT
      READ(5,155) (FMT(I),I=1,INF)
      155 FORMAT(20A4)
      IF(P.GT.10.OR.P.LT.1.OR.ITS.GT.20.OR.NFORMT.GT.3)GO TO 996
C NOW READ DATA ONE TABLE AND ASSOCIATED COVARIATES PER LINE
      K=0
      7 K=K+1
      READ(ITAPE,FMT,END=10)(RT(I),I=1,4)Z1,Z2
      RZ(1)=Z1
      RZ(2)=Z2
      86 IF(K.GT.120)GO TO 9
      IF(RT(1).EQ.0.0.AND.RT(2).EQ.0.0.AND.RT(3).EQ.0.0.AND.RT(4).EQ.0.0
      A)GO TO 10
      DO 8 I=1,4
      8 T(I,K)=RT(I)
      DO 85 I=1,2
      85 Z(I,K)=RZ(I)
      GO TO 7
      9 WRITE(6,154)
      154 FORMAT('ODATA EXCEEDS MAXIMUM OF 120 TABLES')
      GO TO 900
      10 NTAB = K-1
      NP=1
      DO 14 K=1,NP
      14 B(K)=0.0
      IVAR(1)=1
      JPRINT=0
      CALL LOGODD(NM,NP,NTAB,B,SCORE,ITS,EPS,IME,JPRINT,Z,INFO,T,IVAR,
      **SCORE1)
      IVAR(1)=1
      IVAR(2)=2
      NP=2
      CALL LOGODD(NM,NP,NTAB,B,SCORE,ITS,EPS,IME,IPRINT,Z,INFO,T,IVAR,
      **SCORE1)
      900 CONTINUE
      996 STOP
      END

```

SUBPROGRAM LOGODD

```

      SUBROUTINE LOGODD(NM,NP,NTAB,B,SCORE,ITS,EPS,IME,IPRINT,Z,INFO,T,
      * IVAR,SCORE1)
C   PROGRAM FOR THE REGRESSION ANALYSIS OF THE LOG ODDS RATIO
      DOUBLE PRECISION INFO(NM,1),LOGLIK,B(1),SCORE(1)
      DOUBLE PRECISION Z(NM,1),T(4,1),PROB,E,V,DETER,CHIS,ULL
      DOUBLE PRECISION TEMP,SCORE1(1)
      DIMENSION IVAR(1)
C   ITS = MAXIMUM NUMBER OF ITERATIONS
C   IME = 1 IF WANT EXACT CALCULATIONS
C   IME = 2 IF WANT APPROXIMATE CALCULATIONS
C   IPRINT = 1 IF WANT OBSERVED AND FITTED TABLES
C   PRINTED OUT AT 1ST AND LAST ITERATIONS
      WRITE(6,149)
149  FORMAT('1REGRESSION ANALYSIS OF LOG ODDS RATIO'/)
      WRITE(6,151)NTAB
151  FORMAT(' NUMBER OF TABLES = ',I3)
      IF(IME.EQ.1)WRITE(6,152)
      IF(IME.EQ.2)WRITE(6,153)
152  FORMAT(' EXACT ANALYSIS')
153  FORMAT(' APPROXIMATE ANALYSIS')
      WRITE(6,150)NP,(IVAR(K),K=1,NP)
150  FORMAT(' NUMBER OF VARIABLES IN THIS ANALYSIS =',I3,
      */, ' THESE VARIABLES ARE NUMBERS',30I3)
      WRITE(6,154)ITS
154  FORMAT(' MAXIMUM NUMBER OF ITERATIONS',I4)
C   INITIALIZE BETA TO ZERO
15  NFLAG=0
      DO 60 II=1,ITS
      IF(II.EQ.ITS)NFLAG=1
        II=0
      CALL CALC(NTAB,NP,NM,IME,II,SCORE,INFO,LOGLIK,B,CHIS,ULL,Z,T,IVAR)
      TEST=0.0
      DO 20 K=1,NP
      S=SNGI(SCORE(K))
      SCORE1(K)=SCORE(K)
20  TEST=TEST+ABS(S)
      IF(TEST.LT.EPS)NFLAG=1
      IF(NFLAG.EQ.1) GOTO 47
48  CALL INVR(INFO,NP,SCORE,1,DETER,NM,NP)
      TEMP=0.0
      DO 10 I=1,NP
      TEMP=TEMP+SCORE(I)*SCORE1(I)
10  CONTINUE
      IF(II.EQ.1)WRITE(6,116)
116  FORMAT(' ITER      LOG-LIKELIHOOD      SCORE      PARAMETER ESTIMA
      *TES')
47  CONTINUE

```

```

        WRITE(6,107)II,LOGLIK,TEMP,(B(J),J=1,NP)
107  FORMAT(1X,I4,F17.4,F13.4,2X,8F12.4,20(/,36X,8F12.4))
        IF(NFLAG.EQ.1)GO TO 99
CALCULATE NEW VALUE OF BETA
52  CONTINUE
    DO 55 K=1,NP
55  B(K)=B(K)-SCORE(K)
60  CONTINUE
99  CONTINUE
    WRITE(6,101)(B(J),J=1,NP)
101  FORMAT(' ESTIMATED PARAMETER VECTOR'/(1X,10F12.6)
    WRITE(6,102)(SCORE(J),J=1,NP)
102  FORMAT(' FIRST DERIVATIVE LOG-LIKELIHOOD'/(1X,10F12.6))
    WRITE(6,103)
103  FORMAT(' INFORMATION MATRIX')
    DO 11 I=1,NP
    WRITE(6,104)(INFO(I,J),J=1,I)
11  CONTINUE
104  FORMAT(1X,10F12.6)
    CALL INVR(INFO,NP,SCORE,1,DETER,NM,NP)
    WRITE(6,105)
105  FORMAT(' ESTIMATED COVARIANCE MATRIX')
    DO 12 I=1,NP
    SCORE1(I)=B(I)/DSQRT(INFO(I,I))
C  SCORE1 IS USED TO CALCULATE STANDARDISED COEFFICIENTS
    WRITE(6,104)(INFO(I,J),J=1,I)
12  CONTINUE
    WRITE(6,106)(SCORE1(I),I=1,NP)
106  FORMAT(' OSTDANDISED REGRESSION COEFFICIENTS'/(1X,10F12.6))
4  CONTINUE
    WRITE(6,156)CHIS
156  FORMAT(' OCHI-SQUARE GOODNESS OF FIT ',F12.4)
    WRITE(6,157)ULL
157  FORMAT(' O2XLR RATIO GOODNESS OF FIT ',F12.4)
    I=NTAB-NP
    WRITE(6,158)I
158  FORMAT(' ODEGREES OF FREEDOM ',I5)
    IF(IPRINT.NE.1)GO TO 899
    WRITE(6,290)(B(J),J=1,NP)
290  FORMAT(' 1 FITTED VALUES EACH TABLE FOR BETA VALUES :', 9F9.4/(43X,
    * 9F9.4))
    CALL CALC(NTAB,NP,NM,IME,1,SCORE,INFO,LOGLIK,B,CHIS,ULL,Z,T,IVAR)
899  CONTINUE
900  WRITE(6,667)
667  FORMAT(' 0 '// 'ORUN TERMINATES SUCCESSFULLY')
    RETURN
    END

```

```

SUBROUTINE CALC(NTAB,P,NM,IME,IPRINT,SCORE,INFO,LOGLIK,B,CHIS,ULL
* Z,T,IVAR)
DOUBLE PRECISION SCORE(1),INFO(NM,1),LOGLIK,Z(NM,1)B(1)
DOUBLE PRECISION T(4,1),SUM,PROB,E,V,CHIS,ULL,T1,T2,T3,T4,T5,
*T6,T7,T8
DIMENSION IVAR(1)
INTEGER P
CHIS=0.0
LOGLIK=0.0
ULL=0.0
IF(IPRINT.NE.1) GOTO 1
WRITE(6,290)
290 FORMAT(//'0 TABLE',5X,'A',11X,'B',11X,'C',11X,'D', 8X,'VARIANCE  L
*OG-LIKE',4X, 'PSI      COVARIATES'/6X,4(4X,'0',4X,'E',2X))
1 CONTINUE
DO 5 K=1,P
SCORE(K)=0.0
DO 5 KK=1,P
5 INFO(K,KK)=0.0
DO 20 I=1,NTAB
SUM=0.0
DO 10 K=1,P
L=IVAR(K)
10 SUM=SUM-Z(L,I)*B(K)
CALL MYST(T(1,I),SUM,PROB,E,V,P,IME)
PROB=DLOG(PROB)
LOGLIK=LOGLIK+PROB
CHIS=CHIS+((T(2,I)-E)*(T(2,I)-E))/V
DO 15 K=1,P
L=IVAR(K)
SCORE(K)=SCORE(K)+Z(L,I)*(T(2,I)-E)
DO 15 KK=1,P
LL=IVAR(KK)
15 INFO(K,KK)=INFO(K,KK)+Z(L,I)*Z(LL,I)*V
T1=T(1,I)
T2=T(2,I)
T3=T(3,I)
T4=T(4,I)
T6=E
T5=T(1,I)+T(2,I)-T6
T7=T(1,I)+T(3,I)-T5
T8=T(3,I) +T(4,I)-T7
IF(T1.GT.0.0)ULL=ULL+T1*DLOG(T5/T1)
IF(T2.GT.0.0)ULL=ULL+T2*DLOG(T6/T2)
IF(T3.GT.0.0)ULL=ULL+T3*DLOG(T7/T3)
IF(T4.GT.0.0)ULL=ULL+T4*DLOG(T8/T4)
IF(IPRINT.EQ.1)GO TO 21
GO TO 20
21 CONTINUE

```

```
I1=T1
I2=T2
I3=T3
I4=T4
WRITE(6,291) I,I1,T5,I2,T6,I3,T7,I4,T8,V,PROB,SUM,(Z(IVAR(K),I),K=
*1,P)
291 FORMAT(I4,2X,4(1X,I4,F7.2),3F10.3,8F6.2,20(/,84X,8F6.2))
20 CONTINUE
ULL=-2.0*ULL
RETURN
END
```

```

      SUBROUTINE MYST(T,SUM,PROB,E,V,P,IME)
C   THE EXTERNALS ***QUAD*** AND ***BINCOF*** ARE NECESSARY TO USE THIS
C   SUBROUTINE
      INTEGER P
      DOUBLE PRECISION T,PROB,E,V,TOT,BINCOF,PI
      DIMENSION T(4)
      DOUBLE PRECISION SUM,EM,EN,TEE,ENMT,A,B,C,U,G,H,HH
      PI= 3.14159 26535 89793 23846 26433
      U=0.
      H=0.
      G=0.
      HH=0.
      EM = T(1) + T(3)
      EN = T(2) + T(4)
      TEE = T(1) + T(2)
      ENMT = T(3) + T(4)
C   VARIABLE *IME* IS THE FLAG TO DETERMINE EXACT OR ASYMPTOTIC COMP
C   IME = 1 GIVES EXACT
C   IME = 2 GIVES ASYMPTOTIC
      IF(IME.LT.2) GO TO 100
      GO TO 101
C   EXACT CALCULATIONS
100  SAP=SNGL(DEXP(SUM))
      IF(SAP .EQ.1.0)GO TO 403
      A=BINCOF(EM,TEE-T(2))
      B=BINCOF(EN,T(2))
      C = DEXP(T(2) * SUM)
      FNUM = A*B*C
      KK = IDINT(TEE + 1.)
      DO 14 J=1,KK
      A=BINCOF(EM,TEE-U)
      B=BINCOF(EN,U)
      C =DEXP(U*SUM)
      H = A * B *C + H
      G = A*B*C*U + G
      HH = A*B*C*U**2 + HH
14  U = U + 1.
      PROB= FNUM/H
      E = G/H
      V = HH/H - E**2
      SUM=DEXP(SUM)
      GO TO 677
403  A=BINCOF(TEE,T(2))
      B=BINCOF(ENMT,T(4))
      TOT = ENMT + TEE
      C = BINCOF(TOT,EN)
      PROB = (A*B)/C
      EN = T(2) + T(4)
      E = TEE * (EN/(TOT))
      V = (EM*EN*TEE*ENMT)/(TOT*TOT*(TOT-1))
      SUM=DEXP(SUM)
      GO TO 677

```

```
C  ASYMPTOTIC METHOD
101 SUM =DEXP(SUM)
    SAP =SGL(SUM)
    IF(SAP-1.0)16,15,16
15  E = TEE * (EN/(ENMT + TEE))
    GO TO 102
16  A = (1. - SUM)
    B = EM - TEE + (SUM*TEE) + (SUM*EN)
    C = -(SUM*TEE*EN)
    E = QUAD(A,B,C,TEE,EN)
102 V = 1./((1./E) + (1./(TEE-E)) + (1./(EN-E)) + (1./(EM-TEE+E)))
    PROB= 1./(DSQRT(2.*PI*V))
    PROB=PROB*DEXP(-(0.50* ((T(2) - E)**2/V)))
677 CONTINUE
    RETURN
    END
```

```

DOUBLE PRECISION FUNCTION QUAD(A,B,C,TEE,EN)
DOUBLE PRECISION A,B,C,TEE,EN
C =      B**2 - (4.*A*C)
GA = SNGL(TEE)
GB = SNGL(EN)
G = SNGL(C)
IF(G) 32,40,40
40 C = DSQRT(C)
QUAD = ((-B) -C)/(2. * A)
G = SNGL(QUAD)
IF(G.GT.0.0.AND.G.LT.AMIN1(GA,GB)) RETURN
QUAD = ((-B) + C)/(2. * A)
G = SNGL(QUAD)
IF(G.GT.0.0.AND.G.LT.AMIN1(GA,GB)) RETURN
32 WRITE(6,100)
100 FORMAT(' ERROR IN QUAD COMPLEX SOLUTION')
STOP
END

```

```

DOUBLE PRECISION FUNCTION BINCOF(X,Y)
DOUBLE PRECISION UPPER,BELOW,X,Y
UPPER=X
BELOW=Y
BINCOF=0.
A = SNGL(UPPER)
B = SNGL(BELOW)
IF(A-B.LT.B) BELOW= UPPER - BELOW
K = IDINT(BELOW)
IF(K.LT.0)GO TO 2
BINCOF = 1.
IF(K.EQ.0)GO TO 2
DO 41 I=1,K
BINCOF =(UPPER/BELOW) * BINCOF
UPPER = UPPER - 1.
41 BELOW = BELOW - 1.
2 RETURN
END

```

```

SUBROUTINE INVR(A,N,B,M,DETERM,ISIZE,JSIZE)
  DIMENSION IPIVOT(100),INDEX(100,2)
  DOUBLE PRECISION A(ISIZE,JSIZE),B(ISIZE,M),PIVOT(100),AMAX,T
  ASWAP,DETERM
  EQUIVALENCE (IROW,JROW),(ICOLUMN,JCOLUMN),(AMAX,T,SWAP)
C
  10 DETERM=1.0
  15 DO 20 J=1,N
  20 IPIVOT(J)=0
  30 DO 550 I=1,N
C
C   SEARCH FOR PIVOT ELEMENT
C
  40 AMAX=0.0
  45 DO 105 J=1,N
  50 IF (IPIVOT(J)-1) 60, 105, 60
  60 DO 100 K=1,N
  70 IF (IPIVOT(K)-L) 80, 100, 740
  80 IF(DABS(AMAX)-DABS(A(J,K)))85,100,100
  85 IROW=J
  90 ICOLUMN=K
  95 AMAX=A(J,K)
  100 CONTINUE
  105 CONTINUE
  110 IPIVOT(ICOLUMN)=IPIVOT(ICOLUMN)+1
C
C   INTERCHANGE ROWS TO PUT PIVOT ELEMENT ON DIAGONAL
C
  130 IF (IROW-ICOLUMN) 140, 260, 140
  140 DETERM=-DETERM
  150 DO 200 L=1,N
  160 SWAP=A(IROW,L)
  170 A(IROW,L)=A(ICOLUMN,L)
  200 A(ICOLUMN,L)=SWAP
  205 IF(M) 260, 260, 210
  210 DO 250 L=1, M
  220 SWAP=B(IROW,L)
  230 B(IROW,L)=B(ICOLUMN,L)
  250 B(ICOLUMN,L)=SWAP
  260 INDEX(I,1)=IROW
  270 INDEX(I,2)=ICOLUMN
  310 PIVOT(I)=A(ICOLUMN,ICOLUMN)
  320 DETERM=DETERM*PIVOT(I)
C
C   DIVIDE PIVOT ROW BY PIVOT ELEMENT
C
  330 A(ICOLUMN,ICOLUMN)=1.0
  340 DO 350 L=1,N
  350 A(ICOLUMN,L)=A(ICOLUMN,L)/PIVOT(I)
  355 IF(M) 380, 380, 360
  360 DO 370 L=1,M
  370 B(ICOLUMN,L)=B(ICOLUMN,L)/PIVOT(I)

```

C

C REDUCE NON-PIVOT ROWS

C

```
380 DO 550 L1=1,N
390 IF(L1-ICOLUM) 400, 550, 400
400 T=A(L1,ICOLUM)
420 A(L1,ICOLUM)=0.0
430 DO 450 L=1,N
450 A(L1,L)=A(L1,L)-A(ICOLUM,L)*T
455 IF(M) 550, 550, 460
460 DO 500 L=1,M
500 B(L1,L)=B(L1,L)-B(ICOLUM,L)*T
550 CONTINUE
```

C

C INTERCHANGE COLUMNS

C

```
600 DO 710 I=1,N
610 L=N+1-I
620 IF (INDEX(L,I)-INDEX(L,2)) 630, 710, 630
630 JROW=INDEX(L,1)
640 JCOLUM=INDEX(L,2)
650 DO 705 K=1,N
660 SWAP=A(K,JROW)
670 A(K,JROW)=A(K,JCOLUM)
700 A(K,JCOLUM)=SWAP
705 CONTINUE
710 CONTINUE
740 RETURN
END
```

MODEL WITH CONSTANT TERM ONLY

REGRESSION ANALYSIS OF LOG ODDS RATIO

NUMBER OF TABLES = 120

EXACT ANALYSIS

NUMBER OF VARIABLES IN THIS ANALYSIS = 1

THESE VARIABLES ARE NUMBERS 1

MAXIMUM NUMBER OF ITERATIONS 15

ITER.	LOG-LIKELIHOOD	SCORE	PARAMETER ESTIMATES
1	-261.4516	82.1955	0.0
2	-220.0785	0.0152	0.4982
3	-220.0708	0.0000	0.5051
4	-220.0708	0.0000	0.5051

ESTIMATED PARAMETER VECTOR

0.505080

FIRST DERIVATIVE LOG-LIKELIHOOD

-0.000000

INFORMATION MATRIX

313.718469

ESTIMATED COVARIANCE MATRIX

0.003147

STANDARDISED REGRESSION COEFFICIENTS

9.002886

CHI-SQUARE GOODNESS OF FIT 117.3498

2XLR RATIO GOODNESS OF FIT 124.1829

DEGREES OF FREEDOM 119

MODEL WITH LINEAR EFFECT OF YEAR

REGRESSION ANALYSIS OF LOG ODDS RATIO

NUMBER OF TABLES = 120

EXACT ANALYSIS

NUMBER OF VARIABLES IN THIS ANALYSIS = 2

THESE VARIABLES ARE NUMBERS 1 2

MAXIMUM NUMBER OF ITERATIONS 15

ITER	LOG-LIKELIHOOD	SCORE	PARAMETER ESTIMATES	
1	-220.0708	7.2135	0.5051	0.0
2	-216.4595	0.0013	0.5144	-0.0384
3	-216.4588	0.0000	0.5165	-0.0385
4	-216.4588	0.0000	0.5165	-0.0385

ESTIMATED PARAMETER VECTOR

0.516496 -0.038542

FIRST DERIVATIVE LOG-LIKELIHOOD

-0.000000 0.000000

INFORMATION MATRIX

316.476245

110.182468 4876.820615

ESTIMATED COVARIANCE MATRIX

0.003185

-0.000072 0.000207

STANDARDISED REGRESSION COEFFICIENTS

9.152142 -2.680950

CHI-SQUARE GOODNESS OF FIT 111.2280

2XLR RATIO GOODNESS OF FIT 116.8751

DEGREES OF FREEDOM 118

OBSERVED AND EXPECTED VALUES FOR MODEL 2

FITTED VALUES EACH TABLE FOR BETA VALUES : 0.5165 -0.0385

TABLE	A		B		C		D		VARIANCE	LOG-LIKE	PSI	COVARIATES
	O	E	O	E	O	E	O	E				
1	3	2.11	0	0.89	25	25.89	28	27.11	0.605	-1.068	2.464	1.00 -10.00
2	5	4.74	2	2.26	16	16.26	19	18.74	1.331	-1.094	2.371	1.00 -9.00
3	2	2.78	2	1.22	30	29.22	30	30.78	0.813	-1.295	2.371	1.00 -9.00
4	7	5.42	1	2.58	28	29.58	34	32.42	1.590	-1.816	2.281	1.00 -8.00
5	7	6.08	2	2.92	28	28.92	33	32.08	1.770	-1.390	2.281	1.00 -8.00
6	2	1.39	0	0.61	36	36.61	38	37.39	0.421	-0.738	2.281	1.00 -8.00
7	5	4.04	1	1.96	25	25.96	29	28.04	1.221	-1.304	2.195	1.00 -7.00
8	3	2.73	1	1.27	40	40.27	42	41.73	0.842	-0.892	2.195	1.00 -7.00
9	5	4.07	1	1.93	44	44.93	48	47.07	1.250	-1.280	2.195	1.00 -7.00
10	11	8.69	2	4.31	42	44.31	51	48.69	2.587	-2.357	2.195	1.00 -7.00
11	6	6.55	4	3.45	25	24.45	27	27.55	1.955	-1.375	2.112	1.00 -6.00
12	6	6.58	4	3.42	29	28.42	31	31.58	1.984	-1.392	2.112	1.00 -6.00
13	11	8.55	2	4.45	35	37.45	44	41.55	2.576	-2.498	2.112	1.00 -6.00
14	4	3.36	1	1.64	49	49.64	52	51.36	1.064	-1.084	2.112	1.00 -6.00
15	4	7.32	7	3.68	57	53.68	54	57.32	2.267	-3.633	2.112	1.00 -6.00
16	2	3.96	4	2.04	38	36.04	36	37.96	1.269	-2.495	2.112	1.00 -5.00
17	8	7.07	3	3.93	21	21.93	26	25.07	2.113	-1.466	2.032	1.00 -5.00
18	8	8.43	5	4.57	36	35.57	39	39.43	2.585	-1.462	2.032	1.00 -5.00
19	6	5.93	3	3.07	46	46.07	49	48.93	1.882	-1.259	2.032	1.00 -5.00
20	5	4.63	2	2.37	50	50.37	53	52.63	1.489	-1.163	2.032	1.00 -5.00
21	15	12.30	4	6.70	46	48.70	57	54.30	3.741	-2.509	2.032	1.00 -5.00
22	4	3.26	1	1.74	27	27.74	30	29.26	1.067	-1.145	1.956	1.00 -4.00
23	9	8.36	4	4.64	39	39.64	44	43.36	2.635	-1.469	1.956	1.00 -4.00
24	9	8.95	5	5.05	35	35.05	39	38.95	2.778	-1.443	1.956	1.00 -4.00
25	4	3.92	2	2.08	38	38.08	40	39.92	1.285	-1.081	1.956	1.00 -4.00
26	12	12.08	7	6.92	41	40.92	46	46.08	3.691	-1.585	1.956	1.00 -4.00
27	8	8.39	5	4.61	48	47.61	51	51.39	2.677	-1.472	1.956	1.00 -4.00
28	8	10.33	8	5.67	63	60.67	63	65.33	3.300	-2.361	1.956	1.00 -4.00
29	6	5.13	2	2.87	37	37.87	41	40.13	1.700	-1.364	1.882	1.00 -3.00
30	8	7.00	3	4.00	35	36.00	40	39.00	2.266	-1.514	1.882	1.00 -3.00

PROGRAM LOGODD

335

TABLE	A		B		C		D		VARIANCE	LOG-LIKE	PSI	COVARIATES	
	O	E	O	E	O	E	O	E					
31	12	10.64	5	6.36	31	32.36	38	36.64	3.269	-1.765	1.882	1.00	-3.00
32	4	5.12	4	2.88	36	34.88	36	37.12	1.691	-1.604	1.882	1.00	-3.00
33	7	5.13	1	2.87	37	38.87	43	41.13	1.703	-2.107	1.882	1.00	-3.00
34	16	13.89	6	8.11	54	56.11	64	61.89	4.392	-2.127	1.882	1.00	-3.00
35	12	10.84	5	6.16	63	64.16	70	68.84	3.534	-1.713	1.882	1.00	-3.00
36	9	12.68	11	7.32	62	58.32	60	63.68	4.055	-3.259	1.882	1.00	-3.00
37	4	5.67	5	3.33	33	31.33	32	33.67	1.883	-1.991	1.810	1.00	-2.00
38	7	5.64	2	3.36	24	25.36	29	27.64	1.846	-1.666	1.810	1.00	-2.00
39	8	8.12	5	4.88	34	33.88	37	37.12	2.631	-1.425	1.810	1.00	-2.00
40	11	11.73	8	7.27	35	34.27	38	38.73	3.639	-1.659	1.810	1.00	-2.00
41	5	6.31	5	3.69	42	40.69	42	43.31	2.117	-1.740	1.810	1.00	-2.00
42	12	11.22	6	6.78	43	43.78	49	48.22	3.602	-1.635	1.810	1.00	-2.00
43	8	8.82	6	5.18	55	54.18	57	57.82	2.944	-1.609	1.810	1.00	-2.00
44	17	16.87	10	10.13	74	74.13	81	80.87	5.469	-1.775	1.810	1.00	-2.00
45	9	9.92	7	6.08	34	33.08	36	36.92	3.135	-1.653	1.810	1.00	-2.00
46	3	4.99	5	3.01	36	34.01	34	35.99	1.717	-2.327	1.742	1.00	-1.00
47	2	4.37	5	2.63	33	30.63	30	32.37	1.507	-2.902	1.742	1.00	-1.00
48	7	5.57	2	3.43	25	26.43	30	28.57	1.867	-1.716	1.742	1.00	-1.00
49	6	8.67	8	5.33	47	44.33	45	47.67	2.912	-2.671	1.742	1.00	-1.00
50	5	3.77	1	2.23	44	45.23	48	46.77	1.333	-1.528	1.742	1.00	-1.00
51	11	14.77	13	9.23	64	60.23	62	65.77	4.842	-3.165	1.742	1.00	-1.00
52	14	14.10	9	8.90	50	49.90	55	55.10	4.550	-1.687	1.742	1.00	-1.00
53	13	14.73	11	9.27	56	54.27	58	59.73	4.775	-2.036	1.742	1.00	-1.00
54	8	10.53	9	6.47	56	53.47	55	57.53	3.529	-2.463	1.742	1.00	-1.00
55	6	6.24	4	3.76	43	42.76	45	45.24	2.142	-1.344	1.742	1.00	-1.00
56	4	4.89	4	3.11	25	24.11	25	25.89	1.678	-1.455	1.676	1.00	0.0
57	8	9.65	8	6.35	32	30.35	32	33.65	3.127	-1.947	1.676	1.00	0.0
58	4	6.06	6	3.94	23	20.94	21	23.06	1.998	-2.319	1.676	1.00	0.0
59	8	9.72	8	6.28	40	38.28	40	41.72	3.237	-1.983	1.676	1.00	0.0
60	7	7.92	6	5.08	36	35.08	37	37.92	2.676	-1.600	1.676	1.00	0.0

TABLE	A		B		C		D		VARIANCE	LOG-LIKE	PSI	COVARIATES	
	O	E	O	E	O	E	O	E					
61	15	13.90	8	9.10	46	47.10	53	51.90	4.535	-1.797	1.676	1.00	0.0
62	15	17.52	14	11.48	62	59.48	63	65.52	5.710	-2.362	1.676	1.00	0.0
63	9	9.16	6	5.84	46	45.84	49	49.16	3.127	-1.511	1.676	1.00	0.0
64	9	8.59	5	5.41	51	51.41	55	54.59	2.975	-1.495	1.676	1.00	0.0
65	5	6.14	5	3.86	41	39.86	41	42.14	2.146	-1.643	1.676	1.00	0.0
66	6	4.82	2	3.18	22	23.18	26	24.82	1.681	-1.542	1.613	1.00	1.00
67	3	4.25	4	2.75	30	28.75	29	30.25	1.522	-1.669	1.613	1.00	1.00
68	9	6.59	2	4.41	23	25.41	30	27.59	2.236	-2.569	1.613	1.00	1.00
69	12	12.44	9	8.56	34	33.56	37	37.44	3.985	-1.648	1.613	1.00	1.00
70	14	11.39	5	7.61	43	45.61	52	49.39	3.860	-2.443	1.613	1.00	1.00
71	16	13.11	6	8.89	40	42.89	50	47.11	4.325	-2.588	1.613	1.00	1.00
72	17	14.41	7	9.59	61	63.59	71	68.41	4.932	-2.365	1.613	1.00	1.00
73	8	7.87	5	5.13	50	50.13	53	52.87	2.794	-1.448	1.613	1.00	1.00
74	8	10.78	10	7.22	44	41.22	42	44.78	3.634	-2.623	1.613	1.00	1.00
75	9	7.16	3	4.84	22	23.84	28	26.16	2.383	-2.017	1.613	1.00	1.00
76	5	4.18	2	2.82	23	23.82	26	25.18	1.507	-1.322	1.552	1.00	2.00
77	9	5.98	1	4.02	37	40.02	45	41.98	2.175	-3.416	1.552	1.00	2.00
78	11	10.56	7	7.44	31	31.44	35	34.56	3.492	-1.574	1.552	1.00	2.00
79	6	8.88	9	6.12	39	36.12	36	38.88	3.070	-2.810	1.552	1.00	2.00
80	14	15.83	13	11.17	49	47.17	50	51.83	5.218	-2.080	1.552	1.00	2.00
81	21	17.60	9	12.40	50	53.40	62	58.60	5.813	-2.775	1.552	1.00	2.00
82	16	15.88	11	11.12	53	53.12	58	57.88	5.329	-1.762	1.552	1.00	2.00
83	6	5.97	4	4.03	37	37.03	39	38.97	2.160	-1.324	1.552	1.00	2.00
84	9	10.05	8	6.95	41	39.95	42	43.05	3.463	-1.721	1.552	1.00	2.00
85	8	5.85	2	4.15	23	25.15	29	26.85	2.079	-2.346	1.493	1.00	3.00
86	9	7.00	3	5.00	25	27.00	31	29.00	2.449	-2.142	1.493	1.00	3.00
87	8	8.80	7	6.20	46	45.20	47	47.80	3.175	-1.619	1.493	1.00	3.00
88	4	4.73	4	3.27	42	41.27	42	42.73	1.789	-1.399	1.493	1.00	3.00
89	11	10.52	7	7.48	47	47.48	51	50.52	3.741	-1.611	1.493	1.00	3.00
90	11	9.40	5	6.60	51	52.60	57	55.40	3.418	-1.883	1.493	1.00	3.00

TABLE	A		B		C		D		VARIANCE	LOG-LIKE	PSI	COVARIATES	
	O	E	O	E	O	E	O	E					
91	6	7.06	6	4.94	46	44.94	46	47.06	2.604	-1.642	1.493	1.00	3.00
92	9	8.73	6	6.27	32	32.27	35	34.73	3.032	-1.492	1.493	1.00	3.00
93	4	5.21	5	3.79	30	28.79	29	30.21	1.938	-1.651	1.437	1.00	4.00
94	4	5.82	6	4.18	48	46.18	46	47.82	2.226	-2.067	1.437	1.00	4.00
95	9	10.40	9	7.60	54	52.60	54	55.40	3.807	-1.862	1.437	1.00	4.00
96	9	9.25	7	6.75	50	49.75	52	52.25	3.412	-1.558	1.437	1.00	4.00
97	10	9.89	7	7.11	78	78.11	81	80.89	3.768	-1.594	1.437	1.00	4.00
98	14	12.08	7	8.92	48	49.92	55	53.08	4.312	-2.055	1.437	1.00	4.00
99	6	5.81	4	4.19	41	41.19	43	42.81	2.204	-1.337	1.437	1.00	4.00
100	3	2.89	2	2.11	50	50.11	51	50.89	1.175	-1.045	1.382	1.00	5.00
101	4	4.03	3	2.97	53	52.97	54	54.03	1.621	-1.194	1.382	1.00	5.00
102	6	5.75	4	4.25	68	68.25	70	69.75	2.297	-1.363	1.382	1.00	5.00
103	10	11.38	10	8.62	58	56.62	58	59.38	4.225	-1.882	1.382	1.00	5.00
104	4	4.03	3	2.97	57	56.97	58	58.03	1.626	-1.197	1.382	1.00	5.00
105	3	4.02	4	2.98	42	40.98	41	42.02	1.598	-1.510	1.382	1.00	5.00
106	3	2.84	2	2.16	42	42.16	43	42.84	1.173	-1.046	1.330	1.00	6.00
107	10	11.20	10	8.80	52	50.80	52	53.20	4.176	-1.820	1.330	1.00	6.00
108	4	5.10	5	3.90	69	67.90	68	69.10	2.090	-1.603	1.330	1.00	6.00
109	10	7.87	4	6.13	43	45.13	49	46.87	3.025	-2.187	1.330	1.00	6.00
110	5	5.63	5	4.37	34	33.37	34	34.63	2.177	-1.424	1.330	1.00	6.00
111	4	4.45	4	3.55	41	40.55	41	41.45	1.821	-1.306	1.280	1.00	7.00
112	3	5.01	6	3.99	48	45.99	45	47.01	2.047	-2.240	1.280	1.00	7.00
113	13	9.93	5	8.07	42	45.07	50	46.93	3.764	-2.805	1.280	1.00	7.00
114	1	2.24	3	1.76	40	38.76	38	39.24	0.950	-1.663	1.280	1.00	7.00
115	7	4.93	2	4.07	46	48.07	51	48.93	2.061	-2.263	1.231	1.00	8.00
116	5	3.84	2	3.16	46	47.16	49	47.84	1.631	-1.542	1.231	1.00	8.00
117	7	6.00	4	5.00	35	36.00	38	37.00	2.401	-1.557	1.231	1.00	8.00
118	6	6.45	6	5.55	40	39.55	40	40.45	2.625	-1.457	1.185	1.00	9.00
119	3	3.78	4	3.22	51	50.22	50	50.78	1.642	-1.377	1.185	1.00	9.00
120	7	4.23	1	3.77	25	27.77	31	28.23	1.772	-3.338	1.140	1.00	10.00

SUBJECT INDEX TO VOLUMES I AND II

A

Additive effect, I.55
 Additive models, I.55, I.58, II.122–31
 choice between additive and multiplicative models, II.142–46
 Age-incidence curves, II.55
 Age-specific rates, II.49–51, II.193–95
 Age-specific ratios, II.61, II.72
 Age standardization, I.254
 of mortality rates, II.51–70
 Age-standardized death rates, II.91
 Age-standardized mortality ratios. *See* Standardized mortality ratios (SMRs)
 Age/stratum-specific rates, II.61
 Age-time specific comparisons, II.48
 Alcohol consumption
 in relation to oesophageal cancer, I.216, I.218–20, I.223–24, I.227–35
 in relation to oral cancer, I.66, I.67, I.86, I.109, I.110
 Alternative explanation of observed relationships, I.89
 AMFIT program, II.175
 Analysis of variance, multiplicative model for SMRs, II.158
 Animal models, I.236
 Ankylosing spondylitis, irradiation for, I.62
 Annual incidence rates, I.43, I.47
 Armitage–Doll model, II.256, II.264
 see also Multistage models of carcinogenesis
 Asbestos exposure, I.21, I.90, II.31–34, II.38, II.103
 and lung cancer, II.242–44, II.262
 and mesothelioma, II.237–39, II.261

combined with cigarette smoking, I.66–68, II.352–53

Association. *See* Disease association
 Association strength, I.88–89
 Asymptotic normality, II.133–35
 Atomic bomb survivors, I.62, II.22
 life-span study, II.340–44
 Attained significance level, I.128
 Attributable risk, I.73–78, II.21
 for exposed persons, I.74
 population risk, I.74
 relative attributable risk, I.,76

B

Background rates
 incorporation into multiplicative model, II.151–53
 non-parametric estimation of, II.192–99
 Bandwidth, choice of kernel estimates, II.193–95
 Benzene exposure, risk of leukaemia, I.87
 Benzidine exposure and bladder cancer, II.252
 Benzo[*a*]pyrene, and incidence of skin tumours, I.237
 Bermuda Case-Control Symposium, I.19
 Bernoulli distribution, II.132
 Biases, I.22, I.35, I.73, I.84–85, I.89, I.105, I.113, II.9, II.16, II.73
 arising from unconditional analysis of matched data, I.249–51
 due to errors of measurement, II.41–42
 see also Recall bias; Selection bias
 Binomial coefficient, definition, I.125
 Binomial distribution, definition, I.125

- Biological monitoring, II.20
- Birth cohort analysis, I.48
- Bladder cancer, II.21, II.30
 - and benzidine exposure, II.252
 - in chemical industry, II.11
- Bone tumours, and radiation exposure, II.249–50
- Boston Drug Surveillance Program, I.22, I.115
- Breast cancer, II.21
 - age at first birth, I.64–66, I.77, I.86
 - age-specific incidence rates, I.49, I.50, I.59, I.60, II.129
 - and radiation exposure, II.247–49, II.262
 - bilateral, I.87
 - cohort analysis of Icelandic data, II.126–31
 - comparison of indirect standardization and multiplicative model fitting, II.130
 - example of negative confounding, I.93–94
 - influence of reproductive factors, I.66
 - irradiation-induced, I.62
 - relative risks for, I.92
 - reproductive experience in, I.17, I.66
- Breast disease, benign, II.187
- British doctors study, II.27, II.28, II.101, II.163–65, II.168–70, II.236, II.336–39
- C
- Calendar period-specific rates, II.49–51
- Calendar time, I.43
- Carcinogenesis, multistage models of, II.256–60
- Case-control sampling, II.205, II.289–302
 - see also* Risk set sampling
- Case-control studies
 - applicability, I.21
 - as related to cohort studies, II.3–22, II.35–36, II.42, II.44
 - chi-squared test statistic
 - 1 D.F. test for trend, I.147–50
 - combination of $2 \times K$ table, I.149
 - goodness-of-fit, I.209–10
 - in logistic regression, I.208–10
 - goodness-of-fit, I.208, I.222, I.273
 - matched pairs (McNemar test), I.165, I.184
 - for homogeneity of relative risk, I.166–69, I.185
 - 1 D.F. test for trend, I.184
 - matched samples (1:M), I.171, I.177
 - for homogeneity and trend in relative risk, I.173–76
 - series of 2×2 tables (Mantel–Haenszel statistic), I.138
 - for homogeneity of relative risk, I.142, I.143
 - for trend in relative risk, I.142
 - series of $2 \times K$ tables, I.149
 - trend test, I.149
 - summary chi-squared for combination of 2×2 table (Mantel–Haenszel statistic), I.138
 - test of homogeneity, I.166–69
 - 2×2 table, I.131–32
 - $2 \times K$ table, I.147
 - contradictory results, I.19
 - definition, I.14–16
 - design considerations, II.272, II.289–302
 - choice of case, I.23–25
 - choice of control, I.25–28
 - efficiency of, I.21
 - future role of, I.18
 - general considerations, I.14–40
 - history, I.17
 - limitations, I.22
 - low cost of, I.21
 - major strengths of, I.20–22
 - objectives of, I.17, I.19–20
 - planning, I.23–32
 - present significance, I.17–19
 - status of cases, I.24
 - unmatched design considerations, II.289–94, II.302–4
- Causality, I.84–85
 - criteria, I.36–37, I.86–90
 - evidence of, I.90
- Chi-squared test statistic. *See* under Case-control studies; Cohort studies
- Childhood cancers, I.239–42
- Chronological age, I.43
- Cigarette smoking. *See* Tobacco consumption
- Classification errors, I.114
- Coding of disease, II.30
- Coffee drinking, lack of dose response for bladder cancer, I.86,

- Cohort studies
 chi-squared test statistic, II.68, II.94, II.137
 comparison of two SMRs, II.94
 for SMR, II.68, II.69
 goodness-of-fit test for grouped data, II.129
 heterogeneity of SMR, II.96
 1 d.f. test for trend of SMR, II.96
 summary test for equality of relative risk, II.108, II.113
 test for heterogeneity of relative risk, II.112
 test for trend in relative risk, II.112
 definition, II.2
 design and execution, II.22
 design considerations, II.271–88
 further information from, II.28–29
 general considerations, II.2–46
 historical, II.2, II.5–11, II.19, II.21, II.32, II.33, II.35, II.37, II.42
 identification of cancer cases, II.28
 implementation, II.22–36
 interpretation, II.36–45
 limitations of, II.20–22
 present significance of, II.11–20
 problems in interpretation of, II.39–45
 prospective, II.2, II.20, II.33, II.35, II.37
 retrospective. *See* historical (*above*)
 sample size for
 comparison with external standard, 2.273–79
 comparison with internal control group, 2.279–85
 specific strengths of, 2.11–20
 Combined exposures, I.66–68
 see also Joint effects
 Comparative mortality figure (CMF), II.48, II.61–63, II.90, II.125, II.126
 instability of, II.63
 standard error of, II.64
 versus SMR, II.72
 Comparison groups, choice of, II.33–34, II.39–40, II.61
 Comparisons with several disease or control groups, I.111–12
 Composite variables, I.105
 Computer programs, II.175, II.192, II.206
 AMFIT, II.175
 GLIM. *See* GLIM
 LOGODDS, I.322–38, II.189
 MATCH, I.297–306
 PECAN, II.206
 Conditional analysis, I.249
 Conditional distribution
 for 2×2 table, I.125
 for series of 2×2 tables, I.138
 Conditional likelihood, I.204, I.209, I.248, I.251, I.253, I.255, I.270
 Conditional logistic regression analysis for matched sets, I.248–79, I.297–306
 Conditional maximum likelihood estimate
 for 2×2 table, I.127
 Confidence coefficient, definition, I.128
 Confidence intervals, I.134, I.165–67, I.182
 definition, I.128–29
 for common odds ratio in series of 2×2 tables, I.141–42
 for odds ratio in 2×2 table
 Cornfield, I.133–34
 exact, I.129
 logit, I.134
 test based, I.134
 for ratio of SMR, II.95
 for relative risk in matched pairs, I.163–67
 for relative risk in matched sets (1:M), I.172–76, I.182
 for the SMR, II.69–71
 logistic regression parameters. *See* Covariance matrix of logistic regression parameters
 test based, I.134, I.135
 see also Standard error of Mantel-Haenszel estimate
 Confounding, residual, I.100, I.101
 Confounding effects, I.84–85, I.93–108, II.87
 and misclassification, I.106
 control of, I.29–30, I.36, I.111, I.136–56, I.162, I.166
 effect of study design on, I.101–3
 negative, I.95
 of nuisance factors, I.225–26
 on sample size requirements, II.304–6
 statistical aspects of, I.94–97
 see also Logistic model and logistic regression; Standardized mortality ratios (SMRs), bias in the ratio of Stratification

Confounding risk ratio, I.76, I.96–97, I.99–101, II.33

Confounding score index, I.101

Conjugated oestrogen dose, I.178

Continuity correction, I.131–34, II.296, II.301

Continuous data, fitting models to II.178–229

Continuous data analysis, I.227–33
 choice of basic time variable, II.180–81
 comparison with grouped data analysis, II.211–12
 construction of exposure functions, II.181–82
 external standard rates, II.183–84
 fundamentals of, II.179–84
 in matched studies, I.265–68
 model equations, II.182–83
 of Montana smelter workers, II.206–18
 of South Wales nickel refiners – nasal sinus cancer, II.218–29

Continuous variables, I.92

Contour plot of deviances, II.169

Controls
 choice of control series, I.25–28
 selection procedure, II.205–6
 see also Case-control studies; Matching; Risk set sampling

Cornfield's limits, I.133–34

Coronary disease among British male doctors, II.112, II.145, II.146

Corrected chi-squared statistic, I.131

Covariance matrix of logistic regression parameters, I.207

Cross-classification. *See* Stratification

Cross-sectional analysis of incidence rates, I.48

Cross-tabulation. *See* Stratification

Cumulative background rates, estimation of, II.192–97, II.204–5

Cumulative incidence rates, I.49–53

Cumulative rate, II.57–58
 standard error, II.58–61

Cumulative ratio, kernel estimation of, II.193–95

Cumulative relative rates, II.204–5

Cumulative standardized mortality ratio, II.207

D

Data acquisition, II.36

Data collection, II.35, II.42

Data points, influence of, II.139–40

Death rates, US national, II.358–61

Denominator information, II.26–28

Design considerations. *See* under Case-control studies; Cohort studies; Matching

Deviances
 contour plot of, II.169
 see also Likelihood inference

Dichotomous exposure
 1:M matching, I.169–76
 in unmatched studies, I.124–46
 matched pairs, I.164–76
 variable number of controls, I.176–82

Dichotomous variables, I.91, I.94–97

Directly standardized rate, I.50, II.52–57, II.89–91
 standard error, II.58–61
 see also Comparative mortality figure (CMF)

Disease association models, I.53–59

Disease occurrence, measures of, I.42–47

Dose metameter selection, II.98–99

Dose-response analysis, II.105, II.115–18
 see also Logistic model and logistic regression; Regression analysis; Trend tests

Dose-response relationship, I.86, I.88, II.37, II.41, II.42, II.82, II.83, II.88, II.97, II.96, II.159, II.232
 multistage models, 2.262–63
 see also Joint effects

Dose-time relationships, II.233–55

Dose-time-response relationships, II.120

Dose transformations, II.159

Dummy variables for logistic regression models, I.196, I.214

E

Ecological studies, II.4

Effect modification. *See* Interaction (modifying) effects

Efficiency calculations for matched designs, II.302–4
 Empirical odds ratio, I.127
 Endometrial cancer, I.24, I.29, I.90, I.104, I.265
 Los Angeles study of, I.162–63, I.185, I.253, I.255, I.258, I.260, I.261, I.263, I.264, I.266, I.290–96
 Endometrial hyperplasia, I.30
 Epithelial tumours, age-specific rates and latent period, I.60, I.62, I.89
 Errors of classification and of measurement, I.114, II.41–42, II.265–66
 generating confounding, I.106
 Estimation of odds ratio
 combination of 2×2 table
 logit estimate, I.139
 Mantel-Haenszel estimate, I.140, I.141
 maximum likelihood estimate, I.140
 2×2 table
 asymptotic maximum likelihood, I.130
 exact conditional maximum likelihood, I.124
 see also Logistic regression
 Excess mortality ratio (EMR), II.174, II.175, II.268
 Excess risk, I.55, I.58, I.64, I.84, II.45
 Excess risk model
 fitting to grouped data, II.171–76
 see also Additive models
 Exponential distribution, II.132
 Exponential survival times, II.131–32
 Exposure functions, II.181–82
 Exposure index, II.172–73
 Exposure information, II.30–33, II.37
 Exposure probability, I.71
 Exposure variables, lagging of, II.48, II.87
 External standard rates, II.212–14
 Extra Poisson variability, II.99–100

F

Familial risk, I.87
 Fisher's exact test, I.128, I.129, I.133
 Fitted values in 2×2 table, I.130
 Follow-up losses, II.40–41, II.49
 Follow-up mechanisms, II.17, II.25–29
 Follow-up period, II.288–89

Follow-up schema, II.50
 Force of morbidity, I.45
 Force of mortality, I.45
 Forerunners of disease, II.44

G

Gall-bladder disease, I.22, I.168, I.254–59, I.262, I.264, I.265
 Gastric cancer, age-specific incidence rates, I.62
 GLIM computer program, I.206, I.208, I.214, I.253, II.128, II.136–37, II.139, II.141–143, II.160, II.162, II.163, II.167, II.174, II.175
 Global statistic for homogeneity test, I.142
 Global test, I.153
 Goodness-of-fit, I.142, II.161, II.190, II.199
 analysis of residuals, II.138–39, II.144–46
 in logistic regression, I.208, I.222, I.273
 influential data points, II.139–40, II.144–46
 multiplicative models, II.148, II.149
 statistics, II.141, II.144
 summary of measures of, II.137–38
 see also Case-control studies, chi-squared test statistic; Cohort studies, chi-squared test statistic
 Greenwood's formula, II.192
 Group-matching, I.122
 Grouped data, fitting models to, II.120–76
 Grouped data analysis
 case-control studies, I.122–59
 goodness-of-fit in logistic regression, I.208, I.222, I.273
 Ille-et-Vilaine study of oesophageal cancer, I.281–83
 Oxford Childhood Cancer Survey, I.284–89
 qualitative analysis of, I.213–19
 quantitative analysis of, I.221–24
 cohort studies, II.106–15
 comparison with continuous data analysis, II.211–12
 conservatism of indirect standardization, II.114–15
 extensions to $K > 2$ exposure classes, II.113–114

Grouped data analysis—contd.

cohort studies—contd.

- heterogeneity of relative risk, II.110–13
- Mantel-Haenszel estimate, II.109–13
- Montana cohort, II.146–50, II.155–59
- restrictions on, II.178
- summary test of significance, II.108
- two dose levels, exposed versus unexposed, II.107–8
- maximum likelihood estimate, II.108–13

H

Hat matrix in residual analysis, II.138–40

Hazard rate, I.45

Hazards, proportional, I.201

Healthy worker effect, II.17, II.39–40, II.87, II.98

Heterogeneity, II.75–76

see also Case-control studies, chi-squared test statistic; Cohort studies, chi-squared statistic

Historical cohort studies, II.2, II.19

HLA antigen A2 and leukaemia, association with survival, I.25

HLA antigens and multiple comparison, I.115

Hodgkin's disease and tonsillectomy, I.16, I.31

Homogeneity of relative risk

in matched pairs, I.166–67

in matched sets (1:M), I.173–74

in series of 2×2 tables, I.137, I.142–43tests for homogeneity, *see under* Case-control studies; Cohort studies

Homogeneity test, global statistic for, I.142

Hypergeometric distribution

central, definition, I.127

K-dimensional, I.147

non-central, definition, I.127

Hypertension, I.168, I.169, I.254–59

I

Ille-et-Vilaine study of oesophageal cancer, I.122–24, I.162, I.210, I.213–33, I.238, I.281–83

Implementation

in case-control studies, I.32–35

in cohort studies, II.22–36

Incidence cohorts, II.25

Incidence rates, I.43, I.66, I.71

age-specific, I.44, I.47–48, I.59–61

calculation of, I.44

cumulative, I.49–53

directly standardized, I.50

estimation of, I.45

logarithmic transformation of, I.57

overall, I.76

time-specific, I.47–48

variations in, I.55

Indicator variables. *See* Dummy variables

Indirect standardization, II.48

see also Standardized mortality ratios (SMRs)

Influence of individual data points, II.139–40, II.144–46

Information matrix, I.207

Initial treatment of data, I.90–93

Instantaneous rate, I.45

Insulation workers, 2.103

Interaction (modifying) effects, I.108–11, I.167, II.110–13

definition of, I.108–11

effect on sample size requirement and matching, II.308–10

in conditional logit analysis, I.262–68, I.273

in logistic model

definition of, I.196–200

test for, I.221–24

in series of 2×2 tables, I.238–42

negative, I.196

see also Case-control studies; Cohort studies

Interaction parameter

in logistic regression, I.196

Internally standardized mortality ratios, II.103–6

International Classification of Diseases (ICD), II.30, II.355–57

Interpretation, basic considerations, I.35–37, I.112–15

Interviews, I.33–34

questionnaires, I.34

Iran, oesophageal cancer in Caspian littoral of, I.275, I.276

Irradiation

obstetric, and associated cancer risk,

I.239–42
 risk of cancer following, I.62–63
 Ischaemic heart disease, and cigarette smoking, I.68

J

Joint distribution, I.99
 Joint effects of multiple exposures, I.66–68, I.99, I.111, I.154–56, I.227, II.266–67
see also Interaction (modifying) effects

K

Kernael estimation of cumulative ratio, II.193–95

L

Lagging of exposure variables, II.48, II.87
 Large strata, unconditional logistic regression for, I.192–246
 Latency function, II.181–82, II.216–17, II.264–66
 Latent period, I.89
 Least-squares analyses, II.161
 Least-squares linear regression analysis, II.99–103
 Leukaemia, and radiation exposure, II.244–47
 Likelihood inference: outline of I.205–10
 likelihood ratio statistic, I.209
 likelihood ratio test, I.207
 log-likelihood analysis, I.206
 log-likelihood function, I.206, II.134, II.184–92, II.202–3
 see also Partial likelihood
 log-likelihood statistic, I.206
 log-normal distribution, II.181
 Log-linear models, definition, I.57
 Log odds. *See* Logit transform
 Log relative risks as logistic regression parameters, I.196
 Logistic model and logistic regression, I.142
 case-control studies, I.202–5
 general definition of, I.200–2

 introduction to, I.193–200
 results of fitting several versions of, I.212
 Logistic regression model, II.153–54
 dummy variable for, I.196, I.214
 Logit confidence limits from combination of 2×2 tables, I.134
 Logit estimate, I.139
 Logit limits, I.141
 Logit transform, I.194, I.196
 LOGODDS program listing, I.322–38, II.189
 Longitudinal studies, II.2
 Los Angeles study of endometrial cancer, I.162–63, I.185, I.253, I.255, I.258, I.260, I.261, I.263, I.264, I.266, I.290–96
 Losses to follow-up, II.40–41, II.49
 Lost-to-follow-up subjects, II.49
 Lung cancer, II.38, II.39, II.43, II.100, II.103
 age- and year-specific death rates, II.391–94
 and asbestos exposure, II.242–44, II.262
 and uranium miners, II.253–55
 British male doctors, II.163–65, II.168, II.169, II.170
 in relation to tobacco consumption, I.17, I.55, I.58, I.64, I.66–69, I.75, I.86–93, I.100, I.101, I.104, I.166, I.193, II.5–9, II.15, II.234–36, II.261
 combined with asbestos exposure, I.66–68
 relative risk of, II.235
 South Wales nickel refiners, II.171, II.174, II.268, II.347–48

M

McNemar's test, I.165
 Mantel test for trend, I.148
 Mantel-Haenszel analysis, II.82
 estimate for cohort studies, II.109–113, II.147, II.285
 for case-control studies, I.138–42, I.144, I.165, I.171, I.172, I.174, I.177, I.179, I.181, I.192, I.195–96
 test for cohort studies, II.189
 Mantel-Haenszel statistic, I.138
 MATCH program listing, I.297–306

- Matched case-control studies, II.297
 design
 comparison with unmatched design, II.306
 number of controls per case, II.304
 sample size requirements for
 dichotomous exposure, II.294–302
- Matched data analysis, I.162–89
 conditional logistic regression analysis for, I.248–79, I.297–306
 conditional logistic regression, I.248–79, I.297–306
 dichotomous variables
 1:1 matching, I.164–69
 1:M matching, I.169–76
 variable number of controls, I.176–82
 polytomous variables, I.182–87
- Matched designs, efficiency calculations for, I.270–76, II.302–4
- Matched versus unmatched analyses, I.102–6, I.249–51, I.270–76
- Matching
 by strata, I.30–31
 in choice of controls, I.28–32
 problems associated with, I.31
- Maximum likelihood estimate (MLE). *See* Estimation of odds ratio; Likelihood inference
- Mesothelioma
 and asbestos exposure, II.237–39, II.261
 of the pleura, I.21
- Misclassification, I.114
see also Biases; Errors of classification and of measurement
- Misclassification rates, II.42
see also Biases, due to errors of measurement
- Missing data, I.113–14
- Model selection, II.203–4
 biological basis for, II.125
- Modelling risk, I.111
- Models, disease association, I.53–59
- Modifying effect. *See* Interaction
- Montana smelter workers
 cohort studies, II.18, II.23, II.32, II.37, II.52, II.53, II.60, II.78, II.79, II.86–99, II.105, II.114, II.148, II.149, II.152, II.154, II.206–18, II.232, II.349–50
 grouped data analyses, II.146–50, II.155–59, II.363–65
 multiplicative models, II.146–50, II.211, II.213,
 numbers alive and under observation, II.202
 regression analyses, II.157–59
 respiratory cancer, II.157, II.158
- Mortality, proportional, II.45–46, II.76, II.115–18
- Mortality rates, I.43
 age-specific, I.65
 age standardization of, II.51–70
 estimation of, I.45
- Mortality ratios, standardized. *See* Standardized mortality ratios
- Mouth cancer. *See* Oral cancer
- Multiple comparison, I.115, II.43–44
- Multiple exposure levels, I.189
 matched studies, I.146–54, I.182–87
see also Logistic model and logistic regression
- Multiplicative models, I.57, I.58, I.67, II.122–31, II.135–42
 choice between additive and multiplicative models, II.142–46
 comparison with indirect standardization, II.125–31
 estimating base line rates under, II.195–99
 fitting of, II.148
 general form of, II.136
 goodness-of-fit, II.148, II.149
 incorporating external standard rates, II.151–53
- Montana smelter workers, II.146–50, II.211, II.213
- nasal sinus cancer in South Wales nickel workers, II.223, II.224, II.226, II.227
- partial likelihood for, II.185–86
 regression coefficients, II.158
- Multistage models of carcinogenesis, II.256–60
 dose-response relationship, II.262–63
 interpretation of epidemiological data in terms of, II.261–62
 metameters of dose when dose levels vary, II.263–65
 Welsh nickel refinery data, II.267–70

Multivariate analysis. *See* Logistic model and logistic regression
 Multivariate normality, I.204

N

Nasal sinus cancer in South Wales nickel refinery workers, II.105, II.106, II.142, II.172, II.218–29, II.268, II.367, II.369–74
 age- and year-specific death rates, II.391–94
 fitting relative and excess risk models to grouped data, II.171–76
 multiplicative model, II.223, II.224, II.226, II.227
see also Continuous data analysis
 Negative confounding, I.95
 Negative interaction, I.196
 Negative results, II.44–45
 Nested hierarchy of models, I.207
 Nickel workers. *See* South Wales nickel refiners
 Non-central hypergeometric distribution, definition, I.127
 Non-identifiability problem, II.128
 Non-multiplicative models and partial likelihood, II.191
 Non-oestrogen drug use, I.262, I.264, I.265
 Non-parametric estimation
 background rates, II.192–99
 relative mortality functions, II.197–99
 Normal approximation to exact distribution for 2×2 table, I.129
 Nuisance factors, confounding effects of, I.225–26
 Nuisance parameters, I.205

O

Obesity and risk of endometrial cancer, I.262, I.265
 Obstetric radiation and associated risk of cancer, I.239–42
 Odds ratio, I.70, I.73, I.94–96, I.99, I.102, I.103, I.106, I.108, I.130–31, I.135, I.139, I.140, I.196, I.250–10, I.241, I.252
 empirical, I.172

 equivalence to relative risk, I.70–72
 estimation of. *See* Estimation of odds ratio
 test for consistency, I.185–87
 test for homogeneity, I.142–46, I.167
 Oesophageal cancer, II.36, II.45, II.159
 among Singapore Chinese, I.274
 dose-response, II.263
 Ille-et-Vilaine study of, I.222–24, I.162, I.210, I.213–33, I.238, I.281–83
 in Caspian littoral of Iran, I.275, I.276
 in relation to alcohol consumption, I.216, I.218–20, I.223–24, I.227–35
 in relation to tobacco consumption, I.154, I.155, I.217–19, I.221, I.223–24, I.227–35, II.266
 log relative risk, I.216, I.217, I.220, I.221
 Oestrogen use, I.24, I.29, I.90, I.93, I.104, I.254–59, I.262, I.264, I.265
 Oral cancer in relation to alcohol and tobacco consumption, I.66, I.67, I.86, I.109, I.110
 Oral contraceptives, I.22
 Overmatching, I.104–6
 Oxford Childhood Cancer Survey, I.239–42, I.270, I.284–89, I.322

P

p-values as measure of degree of evidence, I.128
 Partial likelihood, II.186, II.188, II.189
 for multiplicative models, II.185–86
 Partial likelihood analysis, II.200, II.212–14, II.212
 PECAN program, II.206
 Person-years, algorithm for exact calculation, II.362
 Person-years allocation, II.49–51, II.83, II.85–86, II.88
 to time-dependent exposure categories, II.82–86
 Point prevalence I.42
 Poisson distribution, II.68, II.69, II.70, II.274
 Poisson models
 and the Poisson assumption, II.131–35
 for grouped data, II.185
 Poisson rates, fitting general models to, II.160–67

- Poisson variability, II.99, II.100
 Poolability of data. *See* Matched data analysis
 Population attributable risk, I.74, II.21
 Population controls, I.276
 Portsmouth (USA) Naval Shipyard workers, II.99
 Positive confounding, I.95, I.101
 Positive interaction, I.196
 Potential confounding, I.107
 Power considerations, II.34–35
 see also Case-control studies, design considerations; Cohort studies, design considerations
 Power to detect interaction, II.308–10
 Prevalence, point, I.42
 Prevalence cohorts, II.25
 Proportional hazards, I.201
 Proportional mortality, II.45–46, II.76, II.115–18
 analysis II.153–55, II.216
 incorporating standard rates, II.154–55
 risk functions for, II.168–71
 Proportionality assumption, II.93
 Prospective cohort studies, II.2
- Q**
- Questionnaires, I.34
 information management, I.34–35
- R**
- Radiation exposure
 and bone tumours, II.249–50
 and breast cancer, II.247–49, II.262
 and cigarette smoking, II.254
 and leukaemia, II.244–47
 Rate of occurrence, I.43
 Rates and rate standardization, II.48–79
 cumulative rate, II.57–58
 directly standardized rate, II.52–57
 standard error of cumulative or directly standardized rate, II.58–61
 standardized to world population, II.55–57
 summary measures, II.51
 see also Incidence rates; Mortality rate; Standardized mortality ratios (SMRs)
 Recall bias, I.22, I.35, I.84–85, I.113, II.16
- Regression adjustment for confounders, I.225–26
 Regression analysis, I.232, II.91, II.99, II.100
 Montana smelter workers, II.157–59
 see also Logistic model and logistic regression
 Regression coefficients, I.197, I.215, I.218, I.224, I.274, II.140, II.142
 interpretation of, I.233–36
 multiplicative model, II.158
 standardized, I.208
 Regression diagnostics, II.138–42, II.146, II.161, II.203–4
 Regression models, I.214, I.215, I.222, I.240
 Regression variables, I.239, I.254–59
 Relative attributable risk (RAR), I.76
 Relative mortality functions, nonparametric estimation, II.197–99
 Relative mortality index (RMI), II.75
 Relative risk, I.57–67, I.69–73, I.77, I.84, I.87–89, I.110, I.113, II.106–14, II.142
 additive, II.160
 see also Odds ratio
 Relative risk estimation, II.94–95, II.108–10, II.147
 general models of, II.159–71
 incorporating external standard rates, II.167
 see also Estimation of odds ratio; Mantel-Haenszel analysis
 Relative standardized mortality ratio (RSMR), II.77–78
 Reproductive factors in breast cancer, I.66
 Residual analysis, Hat matrix in, II.138–40
 Residual confounding, I.100, I.101
 Residuals, standardized, I.213
 Respiratory cancer, II.60, II.105, II.207
 standard death rates, II.88
 standard proportion of deaths due to, II.155
 Retirement, II.27
 Risk, I.51, I.53
 see also Excess risk; Relative risk
 Risk-dose-time relationship, modelling, II.232–70
 Risk factors, I.25, I.53, I.55, I.56, I.58, I.66, I.76, I.123, I.128
 binary, I.263

constellation of, I.199
 joint effects of, I.154–56
 more than two levels, I.198
 Risk ratio. *See* Relative risk; Odds ratio
 Risk set sampling, II.199–206, II.214–16,
 II.302–4

Risk specificity
 disease subgroups, I.86–87
 exposure subcategories, I.87
 Risk variables, I.123
 transforming continuous, I.236–38
 Rule of 5, I.139

S

Sample size. *See* Case-control studies, design
 considerations; Cohort studies,
 design considerations; Confounding
 effects; Interaction; Matching
 Sampling requirements, I.72
 Score statistic, I.207
 Second order interaction, I.199
 Selection bias, I.22, I.35, I.85, I.89, I.113,
 II.17, II.49
 Serial measurements, II.20
 Significance level
 attained, I.218
 two-sided, I.133
 Significance tests, I.127, I.131–33
 Single-tail test, I.133
 Skin cancer, case-control study of, I.200
 Skin tumours, I.236
 estimated cumulative incidence rates, I.237
 in mice, I.46, I.53, I.54
 South Wales nickel refiners, II.23–25, II.32,
 II.37, II.142, II.143, II.218–29,
 II.233
 continuous data II.374–90
 lung cancer in II.171, II.174, II.268,
 II.347–48
 mortality experiences, II.171
 multistage models, II.267–70
 nasal sinus cancer, II.268, II.365–67,
 II.369–74
 Spurious associations, I.89
 Standard error
 of CMF, 2.64
 of Mantel-Haenszel estimate, II.109
 of SMR, II.67

Standard populations, II.54–55
 Standardized mortality ratios (SMRs), II.49,
 II.65–68, II.83, II.88, II.125, II.126,
 II.128, II.151, II.152, II.157, II.158,
 II.173, II.175, II.197–98, II.268
 advantages over CMF, II.65–66
 approximate limits for, II.70
 bias in the ratio of, II.72–75, II.92
 by years since first employed, II.217–18
 comparison of II.91–103
 confidence intervals for, II.69–72
 testing for heterogeneity and trend in,
 II.96–97
 testing significance of, II.68–69
 versus CMF, II.72
 Standardized regression coefficient, I.208
 Standardized residuals, I.213
 Statistical inference, I.124–29, I.206
 approximate methods of, I.129
see also Likelihood inference
 Statistical interaction, definition, I.56
 Statistical modelling, advantages and
 limitations of, II.120
 Stomach cancer, age-specific incidence rates,
 I.60, I.61
 STRAT program listing, I.307–21
 Strata matching, I.30–31
 Stratification, I.89, I.105, I.111, I.122, I.225,
 I.242
see also Confounding, control of
 Stratification degree, I.99–101
 Summary chi-squared test for combination of
 2 × 2 tables. *See* under Case-control
 studies; Cohort studies
 Summary measures
 of goodness-of-fit, II.137–38
 of rates, II.51–61
 Survival rates, I.43
 Survival times, II.131–32

T

Tail probabilities, I.27
 Time-dependent exposure categories,
 person-years allocation to, II.82–86
 Time on study, I.43
 Time relationships, II.37–39
 Tobacco consumption, II.43, II.46, II.88,
 II.101, II.103

Tobacco consumption—contd.

and mortality, II.6

and radiation exposure, II.254

in relation to lung cancer, I.17, I.55, I.58,
I.64, I.66–69, I.75, I.86–93, I.100,
I.101, I.104, I.166, I.193, II.5–9,
II.15, II.234–36, II.261combined with asbestos exposure,
I.66–68, II.352–53in relation to oesophageal cancer, I.154,
I.155, I.217–19, I.221, I.223–24,
I.227–35, II.266in relation to oral cancer, I.66–69, I.86,
I.110Tonsillectomy and Hodgkin's disease, I.16,
I.31

Trend tests

for exposure effect versus trend test for
dose-response, II.97–98*see also* under Case-control studies,
chi-squared test statistic; Cohort
studies, chi-squared test statistic 2×2 table, I.126, I.146, I.148, I.154, I.169
approximate statistical inference for,
I.129–44combining results from, I.136–56,
I.210–13

combining sets of, I.268–70

conditional distribution for, I.125, I.138

conditional maximum likelihood estimate
for, I.127equivalence of odds ratio and relative risk,
I.70–72

exact statistical inference, I.124–29

interaction in, I.238–42

odds ratio in, I.248

 $2 \times K$ table, I.146–54

Two-sided significance level, I.133

Two-sided test, I.128

U

Unconditional analysis of matched data, bias
arising from, I.249–51Unconditional likelihood for logistic
regression, I.209, I.253

Unconditional likelihood function, I.269

Unconditional logistic regression, I.269
for large strata, I.192–46

Unconditional model, I.269

Unknown parameters, I.125

Unmatched analysis, I.271–76

Unmatched case-control studies, design
considerations, II.289–94, II.302–4

Unstratified analysis, I.146–47

Uranium miners, and lung cancer, II.253–55

Urinary tract tumour, I.52, I.86

US national death rates, II.358–61

Uterine bleeding, I.104

Uterine cancer, I.27

V

Vaginal adenocarcinoma, I.89

W

Weighted least squares regression, I.60

Welsh nickel refiners. *See* South Wales nickel
refinersWoelf estimate. *See* Logit estimate

Y

Yates correction. *See* Continuity correction