Hum Genet (1988) 81:71–75

Human Genetics

Klinefelter's syndrome in Sardinia and Scotland

Comparative studies of parental age and other aetiological factors in 47,XXY

Andrew D. Carothers[1] **and Giorgio Filippi**[2]

[1] Medical Research Council, Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH42XU, UK
[2] Cattedra di Genetica Medica e Istituto per l'Infanzia, Università degli Studi di Trieste, I-34137, Trieste, Italy

Summary. Data on 151 non-mosaic 47,XXY males from Sardinia, previously reported by Filippi (1986), were analysed for associations with parental ages at birth, sib order and sex ratio among siblings. The results confirm those of earlier Scottish-based studies in that: (1) there was a significant increase in risk of 47,XXY livebirths at advanced parental ages; (2) maternal age, and maternal age alone, was sufficient to explain the effect; (3) there were no independent effects of paternal age or sib order once maternal age had been taken into account; (4) there was no evidence of any distortion of the sex ratio among siblings. Estimates of relative risk at different maternal ages were compatible with those from the Scottish studies, and pooled estimates are therefore derived. They suggest, for example, that the risk at maternal age 40 years is 2–3 times that at age 30 years. In 33 cases, the parental origin of the supernumerary X chromosome was determined by analysing the segregation of genetic markers. The mean parental ages of 19 'maternal' cases were significantly raised above those of controls, whereas those of 14 'paternal' cases were slightly, and marginally significantly, reduced. The conclusions were essentially unaffected by whether the Sardinian population, the siblings of cases or a group of 94 unrelated Sardinian males were used as controls.

Introduction

The incidence of 47,XXY Klinefelter's syndrome is associated with increased parental age at birth (Lenz et al. 1959; Ferguson-Smith et al. 1964; Court Brown et al. 1969; Carothers et al. 1978, 1984; Schreinemachers et al. 1982; Ferguson-Smith and Yates 1984). The two studies by Carothers et al. suggested that the association is primarily with maternal age, and that paternal age has little or no independent effect. They are also the only ones to have provided quantitative estimates of the increase in risk with maternal age indicating, for example, that the risk at maternal age 40 is 2–4 times, that at age 30. Both were based on essentially the same Scottish data from the MRC Cytogenetic Registry at Edinburgh, ascertained through such sources as infertility clinics, mental hospitals, penal institutions and liveborn surveys. The present study was motivated by the wish to establish whether the same conclusions could be drawn from an entirely independent data set. Our source is a group of some 150 47,XXY males from the 265 hypogonadic Sardinian males, previously reported by Filippi (1986). These were individuals who were medically examined for military service, then referred for cytogenetic investigation following diagnosis of bilateral testicular atrophy. Since *all* Italian males are required to undergo this examination, there should be little or no ascertainment bias.

Filippi (1986) compared the 47,XXY males with normal controls and found a highly significant increase in the means of both maternal and paternal ages at birth. However, a potential problem was that the controls were not closely matched to the cases for year of birth. He also reported that in about 25% of cases if was possible to determine the parental origin of the supernumerary X chromosome, by studying the segregation of G6PD deficiency, Xg type and/or colour blindness. We have therefore been able to investigate whether the parental age effects are confined to cases arising from errors occurring in a particular parental sex. Finally, data on sibs were also available, enabling us to examine possible associations with birth order and sex ratio. Neither was found to be a significant aetiological factor in the Scottish studies.

Offprint requests to: A. D. Carothers

Materials and methods

Cases

The case group consisted of 151 males with non-mosaic 47,XXY karyotype identified among the hypogonadic Sardinian males reported by Filippi (1986) (to which reference should be made for full details of ascertainment). One individual had not been karyotyped in time for the earlier study, but was later found to be non-mosaic 47,XXY and has therefore been included here. Relevant data recorded for each case were: date of birth, parents' dates of birth or ages at ascertainment, parental source of supernumerary X chromosome (where known from analysis of G6PD deficiency, Xg type and/or colour blindness), details of siblings (see below).

The following discussion of controls is largely concerned with the problems of matching to the cases for year of birth.

This is a potentially important consideration since in other European populations (though not necessarily in Sardinia, as we will show), there have been large changes in the distribution of parental age at birth with time, particularly in the years following World War II when most of our cases were born. For example, over the period 1942–1971 the mean maternal age at birth fell from 30.5 years to 28.1 years in Italy, and from 29.4 years to 26.4 years in England and Wales (ISTAT 1980, OPCS 1975). It is therefore relevant to the discussion to note that the years of birth of the cases were between 1927 and 1961 with a median of 1948.

Controls

Control group 1 (CG1). This contained the 94 normal controls described by Filippi (1986), consisting of members of families collected for linkage studies in Sardinia. Date (or years) of birth and both parental dates (or years) of birth were recorded for each control. The years of birth of the controls were between 1949 and 1958 with a median of 1954. They were therefore well-matched only to the younger cases, but could be used for analysis of both maternal and paternal age effects.

Control group 2 (CG2). This consisted of population data on the distribution of maternal ages at birth in Sardinia 1951–1961 inclusive, provided by the Istituto Centrale di Statistica (ISTAT), Rome. Data on legitimate live births were used, these being the most appropriate controls for cases with both parental ages known. Unfortunately, population data are not available for the years before 1951, nor are data on paternal ages for any of the period relevant to this study. Like CG1, therefore, CG2 is well-matched only to the younger cases. Also, CG2 can be used only for analysis of maternal age effects.

Control group 3 (CG3). This consisted of all known siblings of the cases. Sex and dates (or years) of birth, or ages at ascertainment of the corresponding case, were recorded for siblings known to be alive at the time of ascertainment. In a few sibships, where the ages of one or more siblings were not known, their dates of birth were estimated by interpolating them evenly between the known dates of birth of a younger and an older sibling. In one sibship, where the age of the youngest sibling was not known, it was estimated by extrapolating by the average spacing between the remaining siblings. A similar method was used for another sibship in which the age of the oldest sibling was not known. Siblings are, of course, well-matched for many environmental and genetic factors, and Carothers et al. (1978) have shown how to correct for any bias arising from imperfect matching for dates of birth. However, a disadvantage is that they do not provide an efficient means of distinguishing between the effects of maternal and paternal age, because the two factors are perfectly correlated within a sibship.

Analyses

Because of the differing merits associated with each type of control, the following analyses were carried out.

Analysis A. The 94 youngest cases, who were born between 1945 and 1961 with a median of 1953, were ranked in order of birth and individually matched to the control with the same rank order in CG1. Both maternal and paternal age effects were examined.

Analysis B. The 94 youngest cases were matched to the Sardinian population data for the nearest year of birth. Maternal age only was examined.

Analysis C. The 94 youngest cases were matched to the combined Sardinian population data for the years 1951–1961 inclusive. Maternal age only was examined.

Analysis D. All 151 cases were matched to the combined Sardinian population data for the years 1951–1961 inclusive. Maternal age only was examined.

Analysis E. All 151 cases were individually matched to their siblings.

Analysis F. The 19 cases with a maternal supernumerary X chromosome were individually matched to their siblings.

Analysis G. The 14 cases with a paternal supernumerary X chromosome were individually matched to their siblings.

In analyses E, F and G, maternal age, paternal age and sib order were examined. The same method of analysis was used for each type of comparison. The method is based on logistic regression, in which the relative risk of being affected (in this case with 47,XXY) is modelled as

$$r = \exp\{f(x,y,z)\},$$

where f is a polynomial function, and x, y and z denote respectively maternal age, paternal age and sib order. The parameters of the model can be estimated by maximum likelihood, and the goodness-of-fit of different models compared by standard likelihood ratio tests (Carothers et al. 1978, 1984). For present purposse, we define a 'best' model, though not necessarily uniquely, as one for which the likelihood is not increased significantly by adding any term in x (and/or y and/or z, where relevant) of order equal to or higher than the highest-order terms already included (up to a maximum of order 3), and for which the likelihood is reduced significantly by removing any of the highest-order terms already included. As explained above, analysis E requires a correction for bias arising from imperfect matching of siblings for date of birth. The correction involves multiplying the relative risk of each individual born in a particular year by a factor proportional to the probability of ascertaining an affected individual born in that year (Carothers et al. 1978). Provided that the annual population incidence is approximately constant, this factor can be estimated by the number of cases (probands) born in the given year. In practice, we 'smooth' the estimate by taking a 5-year moving average.

Results

Table 1 summarises data for the various groups of cases and controls. Although formal statistical comparisons might be misleading in view of the fact that some pairs of groups are not closely matched for date of birth, several general conclusions can be drawn. Firstly, as reported by Filippi (1986), the mean parental ages of the main case group are 3–4 years higher than those of the control groups. Secondly, the effect appears to be confined to cases having a maternal supernumerary X chromosome; the 'paternal' cases differ little from the controls. Thirdly, the mean parental ages of the various control groups are very similar, in spite of differences in the dates of birth. These general impressions are confirmed by the results of the formal analyses summarised below.

Table 1. Data on cases and controls

Group	Mean date of birth	Maternal age (years)			Paternal age (years)		
		n	Mean	SD	*n*	Mean	SD
94 youngest cases	1953.5	94	34.02	7.94	94	39.62	8.29
All cases	1947.5	151	34.06	7.72	147	39.11	8.22
Maternal origin cases	1952.3	19	35.03	7.84	19	38.39	6.62
Paternal origin cases	1945.9	14	29.55	7.22	14	34.60	8.22
94 normal controls	1954.5	94	30.17	6.42	94	35.82	7.00
All normal siblings of cases	1943.8	680	31.53	6.27	666	36.58	7.17
Sardinian population 1951–61	1956.5	352210	31.27	6.34	–	–	–

Table 2. Estimates of model parameters and associated standard errors (SE)

Analysis	a	SE (a)	b	SE (b)	Corr (a, b)	Deviance
A	0.35	0.24	0.96	0.36	−0.33	19.8
B	0.27	0.18	0.53	0.16	−0.67	27.9
C	0.20	0.18	0.55	0.22	−0.50	26.5
D	0.27	0.15	0.47	0.13	−0.69	38.3
E	0.15	0.22	0.78	0.20	−0.52	25.7
Carothers et al. (1978)	0.82	0.19	0.49	0.19	−0.41	46.9
Carothers et al. (1984)	0.66	0.13	0.21	0.16	−0.43	42.0

Maternal age

The best model involving maternal age alone was found by all the analyses A–E to be of the form

$$f = a\left[\frac{x}{10} - 3\right] + b\left[\frac{x}{10} - 3\right]^2$$

For each analysis the estimates of a and b, together with their estimated standard errors and correlation, and the deviance (= −2 × log-likelihood ratio) are shown in Table 2. The latter is a measure of the improvement in goodness-of-fit over the 'null' model, f = constant, and is approximately distributed as chi-squared with n degrees of freedom, where n is the number of estimated parameters. Here n = 2, and the observed values of the deviance are therefore all highly significant ($P < 0.0001$). Corresponding values from the two Scottish studies are also shown. For the sake of comparability, the estimates from Carothers et al. (1984) are based on a quadratic model, even though the second-order term in that study just failed to reach statistical significance (chi-square = 3.52 on 1 *df*; $P = 0.06$). Table 3 shows the estimated relative risks, together with their estimated standard errors, at 2.5-yearly intervals of maternal age for analyses D and E and the two Scottish studies. Note that the relative risk at age 30 is defined to be one. Analysis D was chosen as representative of the four analyses (A–D) based on controls unrelated to the cases, since they all appeared to give mutually consistent results, and analysis D gave the lowest standard errors. Since analyses D and E are not independent of each other, nor are the two Scottish studies, it is difficult to apply formal tests of consistency to the estimates of risk in Table 3. Nonetheless, the four sets of results appear to be in reasonable agreement, provided that the magnitudes of the standard errors are taken into account. We have therefore derived pooled estimates of risk as follows. Let r_i, s_i denote respectively the estimated risk and its associated standard error at a particular maternal age derived from the ith analysis (i = 1,...,4). Let $w_i = r_i^2/s_i^2$. Then,

$$\text{Pooled estimate} = r(P) = \exp\left[\frac{\sum_1^4 w_i \log r_i}{\sum_1^4 w_i}\right]$$

$$\text{Estimated s.e. of } r(P) = \frac{r(P)}{\sqrt{\sum_1^4 w_i}}$$

$$90\%\ \text{C.L. for } r(P) = r(P)\exp\left[\frac{\pm 1.645}{\sqrt{\sum_1^4 w_i}}\right]$$

The pooled estimate is therefore a geometric mean of the individual estimates, weighted to take account of their standard errors. Although this is a reasonable method for obtaining the pooled estimate itself, the estimated standard error is likely to be negatively biased because of the dependencies between the various analyses, and should be interpreted with caution. These values are also shown in Table 3.

Paternal age

Analyses A and E were the only ones relevant to the investigation of paternal age effects. In both cases, although paternal age was found to be a highly significant risk factor *on its own*, there was no evidence of any *independent* paternal age effect once maternal age had been included in the models. That is, taking the best model involving maternal age alone, no significant improvement occurred on fitting any terms in paternal age. Conversely, taking the best model involving paternal age alone, a highly significant improvement did occur on fitting terms in maternal age.

Sib order

The conclusions of analysis E in regard to sib order were exactly as described above for paternal age. That is, there was no evidence of an independent effect of sib order after includ-

Table 3. Estimated relative risks (×100) of 47,XXY by maternal age, ± standard errors. The pooled estimate is a weighted geometric mean of estimates from all four analyses (see text)

Maternal age	Analysis D	Analysis E	Carothers et al. (1978)	Carothers et al. (1984)	Pooled	90% confidence limits	
						Lower	Upper
17.5	149 ± 53	279 ± 145	77 ± 35	61 ± 21	106 ± 21	76	148
20.0	122 ± 31	187 ± 70	72 ± 23	64 ± 16	94 ± 14	74	120
22.5	106 ± 18	138 ± 34	71 ± 15	69 ± 11	88 ± 8	75	102
25.0	98 ± 10	113 ± 16	75 ± 9	76 ± 7	87 ± 5	79	95
27.5	96 ± 4	101 ± 6	84 ± 4	86 ± 3	90 ± 2	87	93
30.0	100 ± 0	100 ± 0	100 ± 0	100 ± 0	100 ± 0	100	100
32.5	110 ± 3	109 ± 6	126 ± 5	120 ± 4	116 ± 2	112	119
35.0	129 ± 7	131 ± 13	170 ± 15	147 ± 9	141 ± 5	133	149
37.5	159 ± 13	173 ± 25	243 ± 33	185 ± 19	179 ± 10	164	196
40.0	209 ± 23	253 ± 53	369 ± 75	239 ± 37	240 ± 18	212	272
42.5	292 ± 45	406 ± 119	596 ± 175	318 ± 73	344 ± 37	287	411
45.0	431 ± 92	719 ± 288	1022 ± 414	433 ± 140	523 ± 79	408	671
47.5	675 ± 196	1402 ± 748	1863 ± 1007	606 ± 264	847 ± 173	606	1185

Table 4. Sex ratio among sibs of cases

	Males	Females	Total	Males per 100 females
Immediately preceding sib	65	60	125	108.3
Other preceding sibs	166	141	307	117.7
All preceding sibs	231	201	432	114.9
Immediately following sib	49	38	87	128.9
Other following sibs	71	90	161	78.9
All following sibs	120	128	248	93.8
Total	351	329	680	106.7

ing maternal age, although it was a highly significant risk factor on its own.

Parental origin of supernumerary X chromosome

Analysis F revealed a significant ($P \sim 0.01$) increase in risk at advanced parental age for the 19 cases with a maternal supernumerary X chromosome. Analysis G revealed a marginally significant ($P \sim 0.05$) *decrease* in risk at advanced parental age for the 14 'paternal' cases. In neither analysis were there sufficient data to determine whether the observed effects were primarily due to maternal age, paternal age or birth order. Linear terms in all three factors gave an equally good fit.

Sex ratio

Table 4 summarises the sex ratios found in various subsets of siblings. None of the ratios differed significantly from each other, or from values in the range 104–107 that are typical of other European populations.

Discussion

Although we have given careful consideration to the problem of matching for date of birth, the fact that similar results were obtained, irrespective of the choice of control group, suggests that the parental age distribution in Sardinia may have remained fairly constant over the period relevant to this study. Although there is no direct evidence on this point, there is indirect evidence from the parental ages of the cases themselves. Because a high proportion of them were born to older parents, their parental age distribution should be particularly sensitive to changes in that of the general population. In fact, the mean maternal ages of the cases born in the periods < 1935, 1935–1939, 1940–1944, 1945–1949, 1950–1954, 1955–1959 and ≥ 1960 were respectively 34.6, 31.2, 35.0, 33.9, 35.5, 32.7 and 33.9 years. Formal analysis confirms the absence of any significant trend. It therefore seems that, unlike other European populations, that of Sardinia did not change much over the period 1930–1960 in respect of parental ages at birth. The implication is that all the analyses A–E are equally valid.

The population of Sardinia during the period covered by this study would appear to have been particularly well-suited to aetiological studies of parental age-related genetic disease, because of the comparatively high proportion of births to older women. For example, some 10.7% of legitimate livebirths in Sardinia in 1951–1961 were to mothers aged 40 or over. Corresponding figures for Italy as a whole were 5.6% in 1955 and 2.0% in 1980, and for Scotland were 3.3% in 1950–1959 and 0.8% in 1980. It is therefore significant that, in spite of these differences, the present study confirms the results of the two earlier Scottish-based ones in that (1) maternal age, and maternal age alone, was sufficient to explain the observed increase in risk with parental age; (2) there was no evidence of independent effects of paternal age or birth order once maternal age had been taken into account; (3) the estimated relative risks at each maternal age were compatible; (4) there was no evidence of any distortion in the sex ratio of the siblings. The conclusions were essentially unaffected by the choice of control group.

The fact that similar conclusions have been drawn from such disparate groups of cases, ascertained by different methods in demographically dissimilar populations, is evidence in favour of their general applicability. We have therefore provided pooled estimates of relative risk (Table 3), based on both sets of data. In order to convert these into estimates of

absolute risk, they must be multiplied by a factor determined from a knowledge of the absolute incidence of 47,XXY in a population with a known maternal age distribution. Such figures are available from a study of 17,522 consecutive liveborn males between 1967 and 1978 in certain Scottish maternity hospitals (Ratcliffe et al. 1986). A total of 23 non-mosaic 47,XXY males were ascertained and, by applying the relative risks from Table 3 to the maternal age distribution for Scotland in 1970–1974, we can estimate the multiplicative factor as approximately 1.25×10^{-5}. This suggests that, in the absence of antenatal intervention, the livebirth incidence of 47,XXY increases from 1 in ca. 800 ($= 100 \times 1.25 \times 10^{-5}$) males at age 30, to 1 in ca. 570 at age 35,1 in ca. 330 at age 40 and 1 in ca. 150 at age 45. To our knowledge, there have been no other large-scale studies of liveborn 47,XXYs with which to compare these estimates. However, provided that prenatal losses of 47,XXYs are assumed to be minimal, comparable figures are available from studies of amniocenteses carried out solely for reasons of advanced maternal age. For example, Table 2 of Schreinemachers et al. (1982), based on some 20,000 prenatal diagnoses and 25 47,XXYs, implies risks of 1 in 820 males at maternal age 34.6 years (equivalent to 35 years at birth), 1 in 310 at age 39.6 and 1 in 120 at age 44.6. Similarly, Table 10 of Ferguson-Smith and Yates (1984), based on some 53,000 prenatal diagnoses and 65 47,XXYs, implies risks of 1 in 1,340 at age 34.6, 1 in 350 at age 39.6 in 1 in 90 at age 44.6. In view of the large random errors associated with all these figures (including our own), it is perhaps surprising that they agree as well as they do.

Contrary to our findings, Ferguson-Smith and Yates (1984) reported a significant *independent* effect of paternal age, in addition to that of maternal age. Their result was based on a subset of data containing only 23 47,XXYs, and may have been fortuitous. Alternatively, an affect at very advanced paternal ages might be more easily revealed by a study of amniocenteses than by one of liveborns. Such a possibility cannot be excluded at present.

Early studies suggested that the parental age effect in 47,XXY was confined to cases with a maternal supernumerary X chromosome, but the samples were too small to provide statistical significante (Ferguson-Smith et al. 1964; Borgaonkar and Mules 1970). Recently Jacobs et al. (1988), using molecular probes to determine parental origin, were able to confirm the suggestion convincingly, and the present study provides further evidence in its support. In both the study by Jacobs et al. and the present one, there was a highly significant increase in mean parental age for the 'maternal' cases, and no increase, even a small decrease, for the 'paternal' cases. The latter result, though only marginally statistically significant in the present study and therefore possibly fortuitous, is of interest because similar findings have been reported for 47,XYY (Carothers et al. 1978), for 45,X (Kajii and Ohama 1979; Carothers et al. 1980; Ferguson-Smith and Yates 1984) and for trisomy 21 of paternal origin (Hook and Regal 1984). A negative trend with paternal age in the frequency of hyperhaploid sperm compliments has also been reported (Martin and Rademaker 1987). The present result therefore adds to a body of evidence suggesting that certain mechanisms leading to aneuploidy may be associated with some reduction in risk at higher parental ages.

Acknowledgements. We are indebted to ISTAT, Rome for providing Sardinian birth statistics, and to Marie-Luise Lee, Anna Raman-Rossi and Rosaria Scarfi for English-Italian translation.

References

Borgaonkar DS, Mules E (1970) Comments on patients with sex chromosome aneuploidy: dermatoglyphics, parental ages, X_g^a blood group. J Med Genet 7:345–350

Carothers AD, Collyer S, De Mey R, Frackiewicz A (1978) Parental age and birth order in the aetiology of some sex chromosome aneuploidies. Ann Hum Genet 41:277–287

Carothers AD, Frackiewicz A, De Mey R, Collyer S, Polani PE, Osztovics M, Horvath K, Papp Z, May HM, Ferguson-Smith MA (1980) A collaborative study of the aetiology of Turner syndrome. Ann Hum Genet 43:355–368

Carothers AD, Collyer S, De Mey R, Johnstone I (1984) An aetiological study of 290 XXY males, with special reference to the role of paternal age. Hum Genet 68:248–253

Court Brown WM, Law P, Smith PG (1969) Sex chromosome aneuploidy and parental age. Ann Hum Genet 33:1–14

Ferguson-Smith MA, Yates JRW (1984) Maternal age specific rates for chromosome aberrations and factors influencing them: report of a collaborative European study on 52,965 amniocenteses. Prenat Diagn 4:5–44

Ferguson-Smith MA, Mack WS, Ellis PM, Dickson M, Sanger R, Race RR (1964) Parental age and the source of the X chromosome in XXY Klinefelter's syndrome. Lancet I:46–47

Filippi G (1986) Klinefelter's syndrome in Sardinia. Clinical report of 265 hypogonadic males detected at the time of military check-up. Clin Genet 30:276–284

Hook EB, Regal RR (1984) A search for paternal-age effect upon cases of 47,+21 in which the extra chromosome is of paternal origin. Am J Hum Genet 36:413–421

ISTAT (1980) Annuario di statistiche. Istituto Centrale di Statistica, Rome

Jacobs PA, Hassold TJ, Whittington E, Butler G, Collyer S, Keston M, Lee M (1988) Klinefelter's syndrome: an analysis of the origin of the additional sex chromosome using molecular probes. Ann Hum Genet 52:93–109

Kajii T, Ohama K (1979) Inverse maternal age effect in monosomy X. Hum Genet 51:147–151

Lenz VW, Nowakowski H, Prader A, Schirren C (1959) Die Aetiologie des Klinefelter-Syndroms. Ein Beitrag zur Chromosomenpathologie beim Menschen. Schweiz Med Wochenschr 89:727–731

Martin RH, Rademaker AW (1987) The effect of age on the frequency of sperm chromosomal abnormalities in normal men. Am J Hum Genet 41:484–492

OPCS (1975) Birth statistics, series FM1 no 2. HMSO, London

Ratcliffe SG, Murray L, Teague P (1986) Edinburgh study of growth and development of children with sex chromosome abnormalities III. In: Ratcliffe SG, Paul N (eds) Prospective studies on children with sex chromosome aneuploidy. Liss, New York, pp 73–118

Schreinemachers DM, Cross PK, Hook EB (1982) Rates of trisomies 21, 18, 13 and other chromosome abnormalities in about 20,000 prenatal studies compared with estimated rates in live births. Hum Genet 61:318–324

Received May 25, 1988 / Revised July 15, 1988