Birth dates of men with cancer of the testis

E G KNOX AND C CUMMINS

From the Health Services Research Centre, Medical School, University of Birmingham, Edgbaston, Birmingham B15 2TJ

SUMMARY A recent rise in the incidence of cancer of the testis in affluent countries and affluent social strata suggests either an iatrogenic effect analogous with adenocarcinoma of the vagina or a prenatal infection analogous with congenital rubella syndrome. Three samples were assembled, two based on registration and one on death certification, and dates of birth were examined for the presence of steps, cycles, and irregular clusters. There was no evidence of a sudden step, and only equivocal evidence of clustering, but there was strong evidence of a temporal cycle. The cycle has a four-month period, analogous with a school-term-related cyclical component identified in other diseases. A prenatal infection is strongly suggested, but the pattern enables us to exclude many of the common childhood infectious fevers.

The incidence of testicular cancer has recently increased, especially in countries with high standards of living, and in affluent social strata within those countries. Two questions arise. What environmental circumstances could cause such an increase? At what ages might these causes operate?

Swerdlow et al., in a study of childhood testicular cancer, noted that the increased incidence in adults had not been matched by an increase in children. They suggested that the causes of childhood and adult cancer might operate at different ages and that the cause of the increase in adults might be found in the postnatal period. However, there are a priori grounds at least for suspecting that many cancers of the testis, in adults as well as in children, might have early (including prenatal) origins.

Firstly, there is an increasing tally in the literature of reported occurrences in father and son, in brothers, and in twins. In these cases, the causes must operate early in development, before birth, or before conception. Abnormal appearances of chromosomes have also been reported from tumour tissues, in non-familial cases, with characteristic defects in chromosomes-1. Secondly, cancer of the testis is generally a disease of young men, ages at onset following quite closely upon the hormonal stimuli of puberty and upon adolescent growth of the testis. If clinical onsets are triggered hormonally, as this pattern suggests, and if tumour growth is dependent upon the hormonal stimulus, as we might suppose, then the origins of the latent tumours upon which the hormones act might indeed be early.

There are also several compelling analogies between the social and geographical distributions of the increased incidence of testicular cancer and increases in several other diseases of early determination, which also demand that we seek early causes for this one. The first springs from the increase of iatrogenic diseases in wealthy countries and in affluent social strata; the particular analogy is the epidemic of clear-cell adenocarcinoma of the vagina in adolescent girls and young women following the administration of stilboestrol to their mothers during pregnancy. A high standard of living postpones the age at which some infections are first encountered, and this can cause an increase in the incidence of particular complications. There are two concrete analogies here. The first is congenital rubella syndrome, the recent epidemic properties of which can be attributed to postponement of the mean age of infection with a consequent increase in the proportion of susceptible young women. The second example is poliomyelitis: infection was postponed to ages at which the infection induced identifiable clinical consequences and the increased mobility of the infected persons enabled the disease to reach epidemic proportions.

A third possible mechanism might be through a lower mean age of puberty due to better nutrition, although this is better documented in girls than in...
immune responses until it encountered stimuli or abrupt steplike show putative dates of birth. We seek evidence of irregular disturbances of the dates of birth, of kinds which might suggest the influence of each of these determinants. In particular, we seek evidence of an abrupt steplike increase of the kind which would reflect the introduction of a new drug, or the consequences of radioactive fallout following the first atomic bombs in 1945 and the test series of the next decade. Secondly, we look for cyclical variations, including both annual variations and oscillations with periods of two or more years, such as those encountered in measles and mumps. Thirdly, we seek evidence of irregular temporal clusters, such as might be caused by prenatal or perinatal infection by agents which follow an irregularly epidemic pattern rather than a systematic cycle. Causes characterised by a gradual and steady increase in exposure, such as the cited example of puberty advancement, or alterations in birthweight affecting the incidence of cryptorchidism, or altered exposures to other putative postnatal risk factors would probably not show any of these temporal features in relation to dates of birth.

Materials and methods

The only practicable approach is a retrospective one. For this it is necessary to identify all men with testicular cancers presenting (or dying) between two nominated dates, and then to seek information on their birth dates. Three separate sets of material were assembled in this manner consisting of (i) all testicular cancer deaths occurring in England and Wales and reported to OPCS on death certificates between 1 January 1978 and 31 December 1982; (ii) all testicular cancers occurring in the West Midlands Region and reported to the West Midlands Regional Cancer Registry ("anniversary dates") between 1 January 1965 and 31 December 1976; and (iii) all testicular cancers occurring in the North West Region reported to the North West Regional Cancer Registry between 1 January 1974 and 31 January 1979. In each data set, although with some differences in detail, the tumours were classified according to histological type.

For purposes of seeking cyclical phenomena, the full sets of material were used, but for seeking step-functions and irregular clusters the material was limited to registrations of deaths or onsets between the ages of 18 and 64 years. The different investigations were pursued separately for seminomas and for teratomas, as well as for all cases together, including unclassified cases.

Seasonal variations were sought and tested using Walter and Elwood's modification of Edwards' technique. Step-functions were sought using a modified Cusum technique, and statistical significance was examined using the Kolmogorov-Smirnov test. Temporal clusters were sought using the principles of the "moving window" test, as developed by Naus, Wallenstein, and Weinstock and an approximation devised by Knox and Lancashire.

Examinations for steps and for clusters are normally tested against the null expectation of a rectangular distribution, that is, it is supposed that in the absence of steps or clusters the events are evenly distributed across the designated period of time. Here, however, the distribution of dates of birth takes the form of an age of onset curve, set out in reverse direction from the point of ascertainment. Indeed, since there is a period of ascertainment, rather than a point of ascertainment, it is assembled from a series of overlapping curves. Our chief technical problem lay in representing this non-rectangular null distribution with sufficient precision to detect departures from it.

Our first attempts used a "smoothed" version of the observed distribution of birth dates. This, however, tended to "follow" suspected steps or clusters and thus to obscure them. The method finally adopted was to derive a modified log-normal function which fitted the age distribution of cancer of the testis, and then to apply it in a reverse sense in a series of overlapping curves. It was found that this could be done with reasonable accuracy provided that we limited our examination to cancers occurring between 18 and 64 years.

The generated frequency density (y) according to age in years (x) used three fitted parameters, as follows:

\[ b = \text{base line age (eg, 18.0)} \]
\[ a = \text{peak age (eg, 25.0), (>b)} \]
\[ c = \text{a scaling factor} \]

Intermediate derived variables were calculated as follows:

\[ g = \ln (a-b) \]
\[ h = \ln (x-b) \]
Then:

\[ y = h \exp \left\{ -\left(\frac{g-h}{c}\right)^2 \right\} \]

For each 30-day interval within the nominated ascertainment period these densities were distributed
Birth dates of men with cancer of the testis

backwards in time over a period extending from 18 years before that date to 65 years previously. The elements of the distribution were also distributed at 30-day intervals. The 30-day density elements were summated overall and then adjusted so that the total matched the observed number of births.

Iterative adjustments of the frequency density parameters were carried out in two stages. The first stage was based upon the age-at-registration distribution, and further refinements were carried out later to obtain the "best fit" between the cumulative observed and the cumulative expected distributions of birth dates. The fitting process was assisted partly by graphical display and partly through minimising the squared differences between the cumulative-observed and cumulative-expected distributions.

The chief effect of the second stage of fitting was to eliminate trends in the numbers of births and in date-of-birth-related incidence trends. This process sharpens the clarity with which we can detect non-linear changes but sacrifices the possibility of commenting upon the linear components and separating incidence trends from trends in numbers of births. It is possible in principle to correct for trends in the numbers of births directly, but there is a residual problem. This arises because the fitted distribution of ages at onset is not itself "pure" but itself depends in part upon changes in birth patterns. We intend to pursue additional methodological investigations, but for purposes of the present study we limit ourselves to the use of null distributions calculated as above.

Results

STEP-PHENOMENA

Separate analyses were performed upon data sets for seminoma and for teratoma, and upon complete sets, irrespective of histological diagnosis. These investigations were limited to men dying (OPCS) or presenting (regional data) between 18 and 64 years of age. Null distributions were calculated separately for each set and subset. Observed and expected cumulative distributions of birth, in calendar time, were then displayed graphically at 300-day intervals and compared. An example is given in figure 1.

A sharp upward change in numbers, during the period spanned by the births, would have caused a progressive separation of the two curves up to the

---

**Fig 1** Cumulative number of births. Men with cancer of testis. Solid lines are mathematically fitted curves (see text). Dotted lines are observed curves of accumulation.
point of the change, and a progressive convergence thereafter. The point of maximum separation would indicate the point at which the rate of accession changed. These graphs do not represent true rates, so that abrupt changes in numbers of total births could produce the same effect, and any positive findings would have to be examined in more detail.

Table 1 displays the maximum deviations between observed and expected curves, in both directions, for several different data sets. These deviations were tested for statistical significance, using the Kolmogorov-Smirnov procedure. The last three columns of the table indicate the deviations which, if they had been achieved, would have been "significant" at the levels indicated at the heads of these columns. None of the observed deviations was in fact significant even at the level of $p = 0.1$. There were no correspondences between the dates of maximum deviations in different data sets to suggest any commonality of pattern.

**TEMPORAL CLUSTERS**

Temporal clusters were sought against the same hypothetical null backgrounds as those used for seeking step-phenomena. Different data sets were examined separately, and time windows of different widths were nominated and "scanned" across the birthdate series, single day by single day. A search was made within these windows for clusters in excess of expectations. The window sizes examined were 30, 60, 90, 120, 150, and 180 days.

Whenever a simple excess of observation over expectation was encountered, the difference was tested for statistical significance in three stages. Firstly it was tested as if the finding had been encountered at random, instead of as the result of a systematic search. If it was not significant at a level of at least $p < 0.01$ it was rejected, but if it was significant at this level, testing proceeded to the second stage. The second stage treated the interval as one of a large series of L intervals, L being calculated as the ratio between the total number of cases and the expected value within the current interval. This is known as a "disjoint" test, and the appropriate p-value, pd, is calculated from the earlier calculated value, ps, as follows:

$$pd = 1 - (1 - ps)^L$$

If the result was significant at least to the level $pd < 0.1$, the investigation proceeded to a third stage.

The third stage of testing relates to a window which can be moved arbitrarily in order to maximise the difference between observed and expected, rather than a series of windows whose boundaries fall randomly, as in the disjoint test. Tests for "moving window" examinations have been developed by a number of workers\(^{25-31}\) but, except in limited and special circumstances, solutions of the exact algebraic expressions are prohibitively in computing time. In some cases it would take more than a lifetime to reach the answer. Published tabulations of values relate only to the "easy" situations with small numbers and rather coarse subdivisions of the time period. An algebraic approximation to the tabulated findings has been developed by Knox and Lancashire\(^3\) and tested beyond the boundaries of the published tables using computer simulation techniques. It provides a reasonably serviceable initial screen. To a first level of approximation, the moving window probability value (pm) is as follows:

$$pm = 1 - (1 - ps)^{LVL}$$

The main results are given in table 2. Rows 1 to 4 show four possible clusters within the OPCS data. The clusters in rows 1, 2, and 4 are in fact the same cluster, detected in moving windows of different sizes. The most striking presentation is in row 4.
Birth dates of men with cancer of the testis

where 11 cases occurred within a window of 180 days against a null expectation of 2·08 cases. The random window and discrete window tests gave highly significant results, and the value pm for a moving window was 0·07, sufficiently small to be suggestive. Row 3 displays another possible cluster: pm was 0·11. None of the other four data sets shown in the table produced significant findings using the moving window procedure. The most suggestive of the outcomes are shown.

We examined the possibility that the moving window test, as used, might be overconservative, failing to detect clusters that were truly present. Additional computer runs were therefore performed after addition/subtraction of random numbers from different ranges to the birthdates. Nothing was found to suggest that errors might have arisen from such causes.

**CYCLICAL PHENOMENA**

In these examinations we include cases with onset or death at any age, and without the limits employed above in relation to step-phenomena and in relation to temporal clusters. The dates of birth are distributed according to calendar month in the first panel of table 3. Statistical tests indicate a significant seasonal variation, but the pattern is rather one of a fluctuating distribution with a prominent peak in

---

**Table 2** Results of searches for temporal clusters

<table>
<thead>
<tr>
<th>Data</th>
<th>Window (days)</th>
<th>Number of onsets</th>
<th>Date of first</th>
<th>Date of last</th>
<th>p random</th>
<th>p disjoint</th>
<th>p moving</th>
<th>Row number</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPCS</td>
<td>692</td>
<td>120</td>
<td>9</td>
<td>27 Aug 60</td>
<td>24 Dec 60</td>
<td>1·469</td>
<td>&lt;0·0001</td>
<td>0·011</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>10</td>
<td>23 Jul 60</td>
<td>19 Dec 60</td>
<td>149</td>
<td>1·941</td>
<td>&lt;0·0001</td>
<td>0·013</td>
</tr>
<tr>
<td></td>
<td>180</td>
<td>12</td>
<td>31 Mar 60</td>
<td>25 Sept 60</td>
<td>174</td>
<td>2·727</td>
<td>&lt;0·0001</td>
<td>0·007</td>
</tr>
<tr>
<td>WM seminoma</td>
<td>336</td>
<td>60</td>
<td>24 Aug 49</td>
<td>12 Oct 49</td>
<td>20</td>
<td>0·859</td>
<td>0·0003</td>
<td>0·082</td>
</tr>
<tr>
<td>WM teratoma</td>
<td>273</td>
<td>9</td>
<td>1 Feb 22</td>
<td>15 Apr 22</td>
<td>74</td>
<td>1·422</td>
<td>0·0035</td>
<td>0·559</td>
</tr>
<tr>
<td>NW seminoma</td>
<td>198</td>
<td>9</td>
<td>9 Sep 37</td>
<td>7 Mar 38</td>
<td>161</td>
<td>2·879</td>
<td>0·0029</td>
<td>0·182</td>
</tr>
<tr>
<td>NW teratoma</td>
<td>140</td>
<td>10</td>
<td>11 Sep 48</td>
<td>6 Feb 49</td>
<td>177</td>
<td>3·623</td>
<td>0·0042</td>
<td>0·150</td>
</tr>
</tbody>
</table>

---

**Table 3** Cyclical distributions of dates of birth

**(a) Annual cycle**

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sept</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPCS</td>
<td>75</td>
<td>45</td>
<td>66</td>
<td>75</td>
<td>75</td>
<td>67</td>
<td>57</td>
<td>89</td>
<td>78</td>
<td>67</td>
<td>54</td>
<td>79</td>
<td>827</td>
</tr>
<tr>
<td>W Midlands Region</td>
<td>48</td>
<td>65</td>
<td>59</td>
<td>74</td>
<td>65</td>
<td>53</td>
<td>58</td>
<td>56</td>
<td>60</td>
<td>55</td>
<td>49</td>
<td>52</td>
<td>694</td>
</tr>
<tr>
<td>NW Region</td>
<td>30</td>
<td>30</td>
<td>32</td>
<td>39</td>
<td>50</td>
<td>23</td>
<td>35</td>
<td>29</td>
<td>41</td>
<td>32</td>
<td>25</td>
<td>40</td>
<td>406</td>
</tr>
<tr>
<td>Total</td>
<td>153</td>
<td>140</td>
<td>157</td>
<td>188</td>
<td>190</td>
<td>143</td>
<td>150</td>
<td>174</td>
<td>179</td>
<td>154</td>
<td>128</td>
<td>171</td>
<td>1927</td>
</tr>
<tr>
<td>Seminomas (W,M+N+W)</td>
<td>43</td>
<td>43</td>
<td>46</td>
<td>59</td>
<td>57</td>
<td>43</td>
<td>50</td>
<td>44</td>
<td>50</td>
<td>49</td>
<td>38</td>
<td>38</td>
<td>560</td>
</tr>
<tr>
<td>Teratomas (W,M+N+W)</td>
<td>30</td>
<td>43</td>
<td>34</td>
<td>42</td>
<td>47</td>
<td>31</td>
<td>38</td>
<td>34</td>
<td>45</td>
<td>34</td>
<td>22</td>
<td>39</td>
<td>439</td>
</tr>
<tr>
<td>&quot;Expected&quot; total</td>
<td>164-9</td>
<td>158-7</td>
<td>175-7</td>
<td>167-3</td>
<td>170-7</td>
<td>162-1</td>
<td>157-9</td>
<td>156-8</td>
<td>155-7</td>
<td>145-9</td>
<td>152-9</td>
<td>1927-0</td>
<td></td>
</tr>
<tr>
<td>O/E ratio</td>
<td>0-928</td>
<td>0-882</td>
<td>0-894</td>
<td>1-124</td>
<td>1-113</td>
<td>0-882</td>
<td>0-950</td>
<td>1-110</td>
<td>1-128</td>
<td>0-989</td>
<td>0-877</td>
<td>1-119</td>
<td>1-0</td>
</tr>
</tbody>
</table>

**(b) 1/3-Annual cycle**

<table>
<thead>
<tr>
<th></th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>O/E ratio</td>
<td>1-078</td>
<td>0-917</td>
<td>0-907</td>
<td>1-118</td>
<td>1-0</td>
</tr>
</tbody>
</table>

χ² (11) [Annual cycle] = 21·31 : p<0·05
χ² (3) [1/3-Annual cycle] = 15·55 : p<0·001
April and May, a prominent notch in November—these providing the chief basis of the yearly cycle—but with secondary peaks in August/September and December. The fluctuations were evident in each data set, although the April/May peak was most prominent in the regional data sets, and the August/September peak was more evident in the national material. The overall pattern was equally evident among the seminomas and among the teratomas.

These findings are difficult to interpret, partly because of the small numbers in the regional material and partly because the pattern does not correspond with a preformulated hypothesis. The task is made no easier by the fact that the birth rate varied, month by month, in the years when these men were being born. A series of “expected” values was therefore calculated, using national birth data for 1950 and correcting the material for the length of the month, with February averaged over leap years and non leap years. In addition, the annual data for 1950 were corrected for the secular trend in births taking place at that time, and thus influencing the distribution within a single reference year. The expected values are given in the table, together with observed/expected (O/E) ratios, and these form the basis of a chi-square test (with 11 degrees of freedom). The fluctuations are probably to be interpreted as statistically significant.

The material was redistributed according to the week of the year but without further clarification. It was also regrouped in the form of a one-third annual cycle, as shown in the second panel of table 3. The purpose here is to display the suggested three-peak pattern more explicitly and to display the extent of the deviation between the observations and the random expectations. The presentation must be regarded strictly as a hypothesis formulating rather than a hypothesis testing exercise, but the extent of the deviation suggests a quite powerful effect.

The material was also examined for 2-year, 3-year, 4-year, and 5-year cyclical patterns without detecting anything or suggesting any parallels with measles, mumps, or chicken-pox: or any other infectious disease with a multi-year cyclical epidemic pattern. The absence of a gross excess of “affected” births in late winter, in the manner exhibited by measles (for example), confirms these exclusions. Figure 2 displays the 24-month arrangement. The regular cycle is clearly demonstrated.

Discussion

Our examination has supplied one firm negative result, one questionable result, and one apparently positive result. The negative result is that cancer of the testis shows no evidence whatsoever that its origins might be related to any stepwise increase of events related to pregnancy or the perinatal period, of a kind analogous with the stilboestrol episode. Radiation fallout effects are less certainly excluded because of the more diffuse time relations between exposure and putative result, but there was certainly nothing positive to suggest any such sequence.

The questionable result related to irregular temporal clusters. Although there was a suggestion of such clusters in one data set, the significance levels were no better than suggestive and showed no consistent pattern between the several sets of data. In view of the harmonic phenomena discussed below, they should in any case probably be regarded as no more than peaks in a more extended oscillation.

Cyclical phenomena were specifically sought, and strong evidence of their existence was elicited. The period of the cycle, however, was entirely surprising and therefore demands the assembly and examination of additional data. This is in hand. In the meantime, we must recognise the apparent strength of the oscillation, and note that we cannot explain it in terms of seasonal patterns of dates of birth in the population as a whole. A more detailed examination of the data than that presented in table 2 suggests that the pattern is present every year, but that the timings of the peaks may vary slightly from year to year, that is, the aggregation displayed in the lower part of table 3 may have blurred the true pattern. In figure 2 we represent the numbers of births, month by month, when the data are (“overlaid”) in successive two-year periods. To these investigators’ eyes, at least, the pattern looks convincing.

Fig 2 Month of birth of men with cancer of testis.
Birth dates of men with cancer of the testis

Pending confirmation of this finding through reference to additional data, we speculate upon possible mechanisms. Two possibilities are suggested. The first is a toxic exposure of the fetus or germ cells associated with work and holiday patterns of the parents; alcohol would be an example. The second is an oscillation in the transmission of an infectious agent other than the ones we have already excluded. There is some prior evidence that this is possible.

Analyses by Fine and Clarkson relating to measles epidemiology confirmed the validity of theoretical models of the rise and fall of epidemics of this disease. The numbers of cases in one time period are determined as the product of three parameters in the previous time period, namely, the number of cases, the number of susceptibles, and the “force of infection”. The “force of infection” is in effect the rate of child contact within a population. Their analysis separated the three components and demonstrated that the “force of infection” followed a one-third annual pattern related to the school terms. Presumably this applies also to infections other than measles. A prenatal infectious hypothesis for the aetiology of testicular cancer seems for the time being to offer the most plausible explanation of the finding.

This work was supported by the Medical Research Council. We thank the West Midlands Cancer Registry (Dr Waterhouse and Miss J Powell), the North West Regional Cancer Registry (Mr T Benn), and OPCS for supplying data.

References

22 Ehrengut W, Schwartau M, Hubmann R. Testicular underlying pathology, especially mumps orchitis in patients developing testicular tumors.