A Hierarchical Frailty Model for Familial Testicular Germ-Cell Tumors

Morten Valberg*, Tom Grotmol, Steinar Tretli, Marit B. Veierød, Tron A. Moger, and Odd O. Aalen

* Correspondence to Morten Valberg, Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, P.O. Box 1122, Blindern, N-0317 Oslo, Norway (e-mail: morten.valberg@medisin.uio.no).

Initially submitted June 24, 2013; accepted for publication October 10, 2013.

Using a 2-level hierarchical frailty model, we analyzed population-wide data on testicular germ-cell tumor (TGCT) status in 1,135,320 two-generational Norwegian families to examine the risk of TGCT in family members of patients. Follow-up extended from 1954 (cases) or 1960 (unaffected persons) to 2008. The first-level frailty variable was compound Poisson-distributed. The underlying Poisson parameter was randomized to model the frailty variation between families and was decomposed additively to characterize the correlation structure within a family. The frailty relative risk (FRR) for a son, given a diseased father, was 4.03 (95% confidence interval (CI): 3.12, 5.19), with a borderline significantly higher FRR for nonseminoma than for seminoma ($P = 0.06$). Given 1 affected brother, the lifetime FRR was 5.88 (95% CI: 4.70, 7.36), with no difference between subtypes. Given 2 affected brothers, the FRR was 21.71 (95% CI: 8.93, 52.76). These estimates decreased with the number of additional healthy brothers. The estimated FRRs support previous findings. However, the present hierarchical frailty approach allows for a very precise definition of familial risk. These FRRs, estimated according to numbers of affected/nonaffected family members, provide new insight into familial TGCT. Furthermore, new light is shed on the different familial risks of seminoma and nonseminoma.

cancer, familial; frailty; hierarchical models; nonseminoma; seminoma; survival analysis; susceptibility; testicular germ-cell tumors

Abbreviations: CI, confidence interval; FRR, frailty relative risk; SIR, standardized incidence ratio; TGCT, testicular germ-cell tumor.

Testicular germ-cell tumor (TGCT) is the most common cancer among young Norwegian males aged 15–49 years (1). The incidence rate of TGCT among Caucasian men in Western countries has increased severalfold over the past 50 years (2). This rapid change within only a few generations suggests a major contribution of environmental factors in the development of this type of cancer (1, 3).

There are few established risk factors for TGCT, other than cryptorchidism (undescended testicles) and increasing adult height (4). The increased risk of TGCT among brothers and sons of affected men (5–11), especially for men with an affected monozygotic twin brother (12), points towards a strong heritable component. Several recent studies explored the impact of single-nucleotide polymorphisms on TGCT susceptibility and revealed several genes associated with TGCT risk (13–17).

Calculation of the relative risk of cancer in siblings is not straightforward. One method that is often used is to create a cohort of persons with at least 1 affected sibling and calculate the standardized incidence ratio (SIR) by comparing the incidence rate in this cohort with the incidence rate in the general population (18, 19). However, the SIR approach disregards both the size of each set of siblings and the structure of the family. It is also difficult to incorporate covariates, such as birth year, using the SIR approach. Dividing individuals into different birth cohorts may lead to very small numbers of affected siblings in each cohort. Although it is not intuitively clear how to define a cohort for comparison with the general population, the major advantage of the SIR approach is that the SIR is easy to calculate.

The peak in the age-incidence curve of TGCT has been regarded as a frailty phenomenon (20–22)—that is, a small subgroup of the population is susceptible to the cancer—and the age-incidence curve rises as these frail individuals develop the disease. Over time, a decreasing proportion of the at-risk population is susceptible, and the incidence rate drops.
Hierarchical frailty models have been used to investigate familial risks of diseases (23–25). Susceptibility to the disease varies across families in addition to the variation on the individual level. This forms the basis for the methodology used in this paper, where the model is modified to handle more children in each family. Here, covariates are easily included, and calculating relative risks for any combination of family members versus any other combination is relatively straightforward. The structure of the family and the size of each set of siblings are incorporated in a flexible manner.

We utilized the flexibility of a hierarchical frailty model to analyze high-quality population-based family data, estimating the familial risks of TGCT and its subtypes. To our knowledge, this study was the first of its kind.

MATERIALS AND METHODS

Data material

Since 1960, every individual in Norway has been given a unique personal identification number. We obtained data on familial relationships for sibships from Statistics Norway, and TGCT status was included through linkage with data from the Cancer Registry of Norway. Cancers were registered from 1954 onward, and persons with a cancer diagnosis must have been alive in 1954 or later to be included. Unaffected family members must have been alive in 1960 to be included. Additionally, the father must have been identifiable—that is, alive in 1960 (or 1954 if he had had TGCT) or later—for the sibship to be included.

A family was defined here as brothers who had the same parents and their father. If the father was identified but the mother was not, this was also considered a family. Only the first 5 sons in a family were included; this requirement entailed exclusion of 7 cases that occurred in later-born sons. If a person was a father to more than 1 group of sons, we defined these groups of brothers as independent families. Thus, the same father may have been included in several families. We considered 2 generations only. A person who was a son in one family could have been a father in another. The end of follow-up was TGCT diagnosis, emigration, death, or December 31, 2008, whichever occurred first.

Childhood TGCTs differ from adult TGCTs in that 80%–90% of all prepubertal TGCTs manifest as teratomas or yolk-sac tumors (26), whereas these types of TGCTs are not common in adults. Thus, the age interval 0–13 years was omitted, leading to the exclusion of 73 TGCT cases.

A total of 6,032 cases of testicular cancer were diagnosed after the 13th birthday among (the first 5) sons of an identifiable father, and their fathers, between 1954 and 2008. There were 2,957 seminoma cases, 2,957 nonseminoma cases, and 24 cases classified as both seminoma and nonseminoma. The remaining 94 cases were non-TGCTs, had no clinical diagnosis histologically confirmed, or had a missing histology code and were thus excluded, leaving 5,938 cases. We performed both a total analysis, not distinguishing between seminomas and nonseminomas, and an analysis by histological subtype, where persons with the other subtype were censored. TGCT cases classified as both seminoma and nonseminoma were included as cases in the total analysis but treated as censored in both of the subtype analyses.

The model

Frailty models have been extensively studied in survival analysis (27–30). In the proportional frailty model, the individual hazard rate is given as the product of the frailty variable, $Z_t$, and a specific basic hazard rate, $\lambda(t)$. Given $Z_t$, the individual hazard rate is

$$h(t|Z_t) = Z_t\lambda(t),$$

where $t$ is time. Let $\Lambda(t) = \int_0^t\lambda(s)ds$ be the cumulative basic hazard rate. In the cancer setting, the basic hazard rate is often assumed to be of the Weibull type, due to the Armitage-Doll framework (31); thus,

$$\lambda(t) = kt^{k-1},$$

where the $k$ parameter is to be estimated. The usual scale parameter in the Weibull hazard rate is subsumed into the frailty variable (20).

The population survival function is found by integrating out the frailty variable in the individual survival function $S(t|Z_t) = \exp(-Z_t\Lambda(t))$, such that $S(t) = \mathcal{L}_Z(\Lambda(t))$. $\mathcal{L}_Z$ is the Laplace transform of the distribution function of $Z_t$ and is given by

$$\mathcal{L}_Z(s) = \exp(-z_{t+1}\phi_1(s)),$$

where $\phi_1(s)$ depends on the distribution of $Z_t$. This is a valid Laplace transform for all Lévy-type frailty distributions (29). The parameter $z_{t+1}$ can be randomized by $Z_{t+1}$ to add levels into the model. Here we define a model with 2 levels, and the Laplace transform of the total frailty of an individual is given by

$$\mathcal{L}(s) = \exp(-\phi_2(\phi_1(s))).$$

Details are provided in Web Appendix 1 (available at http://aje.oxfordjournals.org/).

The first level, $Z_1$, represents the variation in frailty between individuals. The second level, $Z_2$, represents the frailty that varies between families. The second-level frailty variable is decomposed additively, such that each son shares half of his parents’ genes. Each son also shares half of the genes with his brothers, such that no brothers share the same half of their parents’ genes (twins are treated as ordinary siblings). On this level, we also add a term accounting for the environment shared by brothers, independently of their father. This can be an environment shared in childhood or a similar in utero environment. A detailed description of second-level frailty is given in Web Appendix 2.

The first-level frailty variable follows a compound Poisson distribution, which has been successfully used to model incidence rates of both TGCT and other cancers (20, 22, 32–34). This allows for a nonsusceptible subgroup in the population.

The genetic components of second-level frailty are independently identically gamma-distributed, as was previously
done for melanoma (23). The components accounting for the environment shared by brothers are also gamma-distributed. To account for the rise in the TGCT incidence rate, a birth-year covariate is included, as described in Web Appendix 3.

A total of 7 parameters are estimated by numerically maximizing the likelihood function given in Web Appendix 3. An estimate of the covariance matrix of the parameters is obtained by means of the optim routine in R (35). The likelihood contribution of each family is calculated in parallel, using the Rmpi package in R (36). Using 200 processor cores of the Abel cluster (37), the optimization took approximately 2.5–4 hours. Including more than 5 sons in a family would make the model harder to formulate, and it would become more computationally demanding.

The frailty relative risk (FRR) is defined as the probability of individual A’s developing cancer within time $t_A$ after his 13th birthday, given that a family member B develops cancer within time $t_B$ after his 13th birthday, divided by the probability that individual A develops cancer within time $t_A$ after his 13th birthday (22). That is,

$$\text{FRR} = \frac{\Pr(T_A \leq t_A | T_B \leq t_B)}{\Pr(T_A \leq t_A)}, \quad (1)$$

where $T_A$ and $T_B$ denote the times to events, from the 13th birthdays, of the 2 individuals. The expression of the FRR, together with extensions to include more diseased and/or healthy persons, is given in Web Appendix 4.

The results presented here are lifetime FRRs; that is, the $t$’s are set to 70 years (83 years of age), unless otherwise stated. The results are also presented with all birth years set equal to 1950, unless otherwise stated. This is unrealistic, but it was done in order to illustrate that the FRR is reduced when we condition on a greater number of healthy brothers.

### SIR estimates

SIRs are frequently used to estimate familial risks of diseases (8, 10, 18, 38–40). The “cohort” method given by Hemminki et al. (19) is widely cited in these papers. To us, there seems to be a problem with this formula. If it is to be taken literally, families with 2 affected persons make no contribution to the numerator, which is clearly wrong. We have interpreted the approach as follows: We count the cases in a pool of all groups of brothers with at least 2 affected brothers, and we divide by the person-years contributed by persons who have at least 1 affected brother. This gives us an incidence rate in the cohort of brothers of affected individuals, which is compared with the incidence rate in all siblings. The SIR could thus be estimated by

$$\Pr(\text{Individual A develops TGCT} | \text{At least 1 of A’s brothers develops TGCT}) \over \Pr(\text{Individual A develops TGCT}).$$

The variance of the logarithm of the SIR was calculated as $1/(\text{number of cancers in the pooled groups of siblings with at least 2 cancers} – \text{number of families these cases originated from})$, as described by Hemminki et al. (19).

### RESULTS

The mean age at diagnosis for the 5,938 TGCT cases was 33.1 years (range, 13.1–71.5), while the mean ages for seminoma ($n = 2,957$) and nonseminoma ($n = 2,957$) cases were

### Table 1. Distribution of Norwegian Families According to Number of Sons\(^{a}\) in a Study of Testicular Germ-Cell Tumor Risk, 1960–2008

<table>
<thead>
<tr>
<th>No. of Sons</th>
<th>No. of Families</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>719,276</td>
</tr>
<tr>
<td>2</td>
<td>310,911</td>
</tr>
<tr>
<td>3</td>
<td>82,650</td>
</tr>
<tr>
<td>4</td>
<td>17,353</td>
</tr>
<tr>
<td>5</td>
<td>5,130</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,135,320</strong></td>
</tr>
</tbody>
</table>

\(^{a}\) Only the first 5 sons in a family were included.

### Table 2. Distribution of Norwegian Families According to the Types of Family Members Affected by a Testicular Germ-Cell Tumor,\(^{a}\) 1960–2008

<table>
<thead>
<tr>
<th>Affected Family Member(s)</th>
<th>All TGCTs</th>
<th>Seminoma</th>
<th>Nonseminoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>No affected family members</td>
<td>1,127,796</td>
<td>1,131,374</td>
<td>1,131,731</td>
</tr>
<tr>
<td>Father only</td>
<td>2,268</td>
<td>1,367</td>
<td>909</td>
</tr>
<tr>
<td>First son only</td>
<td>3,487</td>
<td>1,738</td>
<td>1,762</td>
</tr>
<tr>
<td>Second son only</td>
<td>1,272</td>
<td>629</td>
<td>663</td>
</tr>
<tr>
<td>Third son only</td>
<td>328</td>
<td>151</td>
<td>182</td>
</tr>
<tr>
<td>Fourth son only</td>
<td>65</td>
<td>31</td>
<td>37</td>
</tr>
<tr>
<td>Fifth son only</td>
<td>18</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Father and first son</td>
<td>19</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Father and second son</td>
<td>10</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Father and first and second sons</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>First and second sons</td>
<td>33</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>First and third sons</td>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>First and fourth sons</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>First and fifth sons</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Second and third sons</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Second and fourth sons</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Second and fifth sons</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Third and fourth sons</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Third and fifth sons</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>First, second, and third sons</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>First, second, and fourth sons</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**No. of Families**

**Total** 1,135,320 1,135,320 1,135,320

Abbreviation: TGCT, testicular germ-cell tumor.

\(^{a}\) If a given combination is not listed, no families had that particular combination.
36.8 years (range, 13.3–71.5) and 29.5 years (range, 13.1–71.3), respectively. As described above, we considered the first 5 sons, and Table 1 shows the numbers of families having particular numbers of sons. There were 29 TGCT-affected father-son pairs; 2 affected fathers had 2 affected sons, 52 families had 2 affected brothers, and 3 families had 3 affected brothers (Table 2).

For seminoma, there were 7 affected father-son pairs, 10 families had 2 affected brothers, and 2 families had 3 affected brothers (Table 2). For nonseminoma, there were 6 affected father-son pairs, 1 father had 2 affected sons, and 21 families had 2 affected brothers (Table 2).

Parameter estimates found through maximum likelihood estimation, with 95% confidence intervals, are given in Table 3. Note especially that the \( k \) parameter, usually interpreted as the number of transitions the cell has to go through to become malignant, is estimated to be approximately 3. The seminoma \( k \) is above 4, and the nonseminoma \( k \) is below 3.

The estimated lifetime FRR of developing TGCT for a son born in 1980, given an affected father born in 1960, was 4.03 (95% confidence interval (CI): 3.12, 5.19). The FRR was borderline significantly higher for nonseminoma than for seminoma (\( P = 0.06 \)): The FRR was 2.63 (95% CI: 1.39, 5.02) for seminoma and 5.70 (95% CI: 3.52, 9.24) for nonseminoma. The father-son estimates were similar across birth years (results not shown), which were arbitrarily chosen here.

The FRR for developing TGCT by age 15 years, given that a brother developed TGCT during his lifetime (by age 83 years), was 6.12 (95% CI: 4.65, 8.06). If the brother developed cancer by age 15 years, the FRR was slightly higher, at 6.41 (95% CI: 4.48, 9.18) (Table 4). Similar results were found for the histological subtypes, with no significant difference between them. The lifetime FRR for an individual, given a diseased brother, was 5.88 (95% CI: 4.70, 7.36) (Table 4). If, in addition, there were 3 healthy brothers in the family, the FRR was reduced to 5.07 (95% CI: 4.11, 502 Valberg et al.

Table 3. Estimated\(^a\) Values for Statistical Parameters Used in a Study of Testicular Germ-Cell Tumor Risk, Norway, 1960–2008

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All TGCTs</th>
<th>Seminoma</th>
<th>Nonseminoma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>95% CI</td>
<td>Estimate</td>
</tr>
<tr>
<td>( \nu )</td>
<td>2.48 \times 10^7</td>
<td>1.70 \times 10^7, 3.68 \times 10^7</td>
<td>4.59 \times 10^{10}</td>
</tr>
<tr>
<td>( \eta )</td>
<td>0.97</td>
<td>0.67, 1.39</td>
<td>0.29</td>
</tr>
<tr>
<td>( \delta )</td>
<td>8.46 \times 10^{-3}</td>
<td>5.78 \times 10^{-3}, 1.24 \times 10^{-2}</td>
<td>1.67 \times 10^{-2}</td>
</tr>
<tr>
<td>( \delta_2 )</td>
<td>1.62 \times 10^{-3}</td>
<td>5.24 \times 10^{-4}, 5.00 \times 10^{-3}</td>
<td>6.29 \times 10^{-4}</td>
</tr>
<tr>
<td>( k )</td>
<td>2.93</td>
<td>2.82, 3.04</td>
<td>4.38</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>-9.95</td>
<td>-10.15, -9.76</td>
<td>-10.39</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.52</td>
<td>0.50, 0.54</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; TGCT, testicular germ-cell tumor.

\(^a\) Maximum likelihood estimation.

Table 4. Frailty Relative Risk\(^{ab}\) of an Individual’s (Brother A) Developing a Testicular Germ-Cell Tumor by a Specified Age, Given That Up to 4 of His Brothers (Brothers B, C, D, and E) Do or Do Not Develop a Testicular Germ-Cell Tumor by a Specified Age, Norway, 1960–2008

<table>
<thead>
<tr>
<th>Age Limit for Brother A, years</th>
<th>Affected Brother(s)</th>
<th>Unaffected Brother(s)</th>
<th>All TGCTs</th>
<th>Seminoma</th>
<th>Nonseminoma</th>
<th>( P ) Value(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brother(s)</td>
<td>Age Limit, years</td>
<td>Brother(s)</td>
<td>Age Limit, years</td>
<td>FRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>15</td>
<td>B</td>
<td>15</td>
<td></td>
<td></td>
<td>6.41</td>
<td>4.48, 9.18</td>
</tr>
<tr>
<td>15</td>
<td>B</td>
<td>83</td>
<td></td>
<td></td>
<td>6.12</td>
<td>4.65, 8.06</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td></td>
<td></td>
<td>5.88</td>
<td>4.70, 7.36</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td>C</td>
<td>83</td>
<td>5.50</td>
<td>4.56, 6.64</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td>C, D</td>
<td>83</td>
<td>5.22</td>
<td>4.30, 6.33</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td>C, D, E</td>
<td>83</td>
<td>5.07</td>
<td>4.11, 6.27</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td>C, D, E</td>
<td>83</td>
<td>21.71</td>
<td>8.93, 52.76</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td>D</td>
<td>83</td>
<td>18.10</td>
<td>9.13, 35.90</td>
</tr>
<tr>
<td>83</td>
<td>B</td>
<td>83</td>
<td>D, E</td>
<td>83</td>
<td>15.80</td>
<td>9.56, 26.11</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; FRR, frailty relative risk; TGCT, testicular germ-cell tumor.

\(^a\) Defined in Web Appendix 4.

\(^b\) All birth years were set to 1950.

\(^c\) Test for no difference between seminoma FRR and nonseminoma FRR (Web Appendix 4).
The corresponding estimates for seminoma were 6.35 (95% CI: 4.07, 9.92) and 4.15 (95% CI: 2.61, 6.60). For nonseminoma, the estimates were 7.53 (95% CI: 4.72, 12.02) and 7.15 (95% CI: 5.02, 10.21), respectively (Table 4). No significant difference between subtypes was found for brothers, although the difference increased with the number of healthy brothers (P decreased from 0.61 to 0.07). The lifetime FRR for an individual, given 2 diseased brothers, was 21.71 (95% CI: 8.93, 52.76) (Table 4). If, in addition, there were 2 healthy brothers in the family, the FRR was reduced to 15.80 (95% CI: 9.56, 26.11). The uncertainties in the subgroup analyses, in this setting, were too large for the results to provide any meaningful insight. Thus, these estimates are not presented in Table 4.

Earlier birth years will produce a higher FRR. A lower FRR is found for later birth years. Thus, an older brother with TGCT will produce a higher FRR for an individual, while a younger brother will produce a lower FRR. To illustrate the impact of birth year, Table 5 shows the lifetime FRR for an individual who develops TGCT, given an affected brother, for selected birth years. The FRR ranges from 5.26 (95% CI: 4.38, 6.32) to 5.88 (95% CI: 4.70, 7.36).

The SIR estimate was 6.38 (95% CI: 4.95, 8.21). Furthermore, the estimated SIR was 5.44 (95% CI: 3.22, 9.19) for seminoma and 9.59 (95% CI: 6.31, 14.55) for nonseminoma. These were slightly higher estimates, but with wider 95% confidence intervals than the FRRs above.

The variance for second-level frailty may be decomposed into 1 genetic part and 1 part due to an environment shared by brothers (Web Appendix 3). The proportion of this variance accounted for by the genetic terms was 75%. For seminoma and nonseminoma, the proportions were 37.6% and 84.4%, respectively.

**DISCUSSION**

We have used a hierarchical frailty model to assess the familial risk of TGCT and its subtypes. The approximately 4-fold increased lifetime risk of TGCT in sons of affected fathers supports previous findings (6–11). We found a borderline significantly higher lifetime FRR in nonseminoma cases than in seminoma cases. Earlier studies found higher point estimates for nonseminomas, but the 95% confidence intervals were almost entirely overlapping (8, 10).

The lifetime FRRs of a brother’s developing TGCT, given 1 affected brother, varied from approximately 5.3 to 5.9, depending on birth year. Somewhat higher FRR estimates were found for developing TGCT by age 15 years. The point estimates were slightly lower than those in previous studies but were not unreasonable given the relatively large confidence intervals usually associated with such estimates (6–11). Although results were not significant, the tendency toward a higher lifetime FRR for nonseminomas than for seminomas was also seen for brothers, especially when there were several healthy brothers in the family. Earlier studies found insignificant results in both directions (8, 10).

In an earlier Norwegian hierarchical frailty modeling study of TGCT, Moger et al. (22) estimated the lifetime FRR for an individual, given 1 affected brother, as 7.4 (95% CI: 4.3, 13), based on hospital registry data. Moger et al. did not decompose the second-level frailty and used parameter estimates from an earlier study to account for the rise in susceptibility to TGCT (20). Although we found a lower point estimate, our 95% confidence interval was included in their confidence interval.

The lifetime FRRs for a brother, given 2 of his brothers’ both developing cancer, varied from 16 to 22, depending on the number of healthy brothers in the family. We have not been able to find any other study that tried to estimate this type of relative risk for TGCT, but in previous studies of breast cancer and thyroid cancer, relative risks were large when more than 1 family member was affected (40–42).

The higher familial risk of TGCT among brothers than among fathers and sons has been attributed to factors such as a possible recessive mode of inheritance and the possible importance of maternal lineage of inheritance (10). In addition, environmental factors acting early in life are thought to play a role. In a family study of all major cancers, TGCT had the highest proportion of childhood shared environmental effects (43).

Our data indicate that there is a stronger heritable component pertaining to nonseminoma than to seminoma. Although it is well established that the 2 histological subtypes have different gene expression patterns, as well as quite distinct epigenomes (44), the difference between their respective heritable components has not previously been estimated or tested formally, to our knowledge. The average 7-year-later age at diagnosis of seminoma than of nonseminoma also supports the
idea of a stronger heritable component of nonseminoma (45). In a previous frailty analysis of Norwegian TGCT data, the estimate corresponding to our $k$ parameter, interpreted as the number of (epi)genetic events necessary to reach malignancy, was approximately 4 for seminoma and 3 for nonseminoma, suggesting a weaker heritable component pertaining to the former than to the latter (20).

We have defined an FRR that produces a relative risk similar to the effect estimates usually considered when studying familial risk of diseases. An alternative is to substitute the denominator in equation 1 with a conditional probability—for example, the probability of individual A's developing TGCT given that the family members do not develop cancer (23, 25). This produces an effect estimate similar to a traditional relative risk, comparing 2 exclusive groups. Such an estimate will always produce a higher FRR than the one presented in this paper. Considering 2 members of a family, this alternative definition will produce different estimates of the FRR depending on which of the two is considered the proband. This leads to, for example, different FRR estimates for a father given a diseased son and a son given a diseased father. With the current definition of the FRR, these 2 estimates are equal.

Although other methods for calculating familial risks are intuitive and should produce reasonable results (6, 11, 46), a hierarchical frailty model allows us to be very detailed as to what kind of familial risk we are actually estimating. We may calculate the risk of developing the disease within a specified time, given various numbers of family members that have or have not developed the disease at given points in their lives (although families were limited to 5 sons in this paper). This is different from the SIR approach, where the estimate is averaged over families of various sizes consisting of persons with different survival times (times at risk) born within a large range of birth years. The increased risk among family members of cancer patients should decrease with the number of healthy persons in the family, and this is handled by the frailty model. Since a hierarchical frailty model takes the structure of the family into account and since we present FRRs for chosen birth years and disease statuses for the family members, the FRRs and the SIR estimate should not necessarily coincide. Although the uncertainty of the FRR estimate may be large for rare combinations, it may allow for counseling of cancer patients and their families according to their specific individual histories. We have not found any other studies that have calculated the relative risk of TGCT in brothers when various healthy or diseased family members are taken into account.

Families consist of individuals who share different parts of their genes and who might also share different degrees of environmental influences. Additive (correlated) frailty models are constructed to take this into account (47–50). The second-level frailty variable in the hierarchical model is decomposed additively, consisting of a genetic part and an environmental part that is shared by brothers independently of their fathers. This ensures that the second-level frailty correlation is larger than 0.5 for brothers, while it is less than 0.5 for father-son pairs. By having 2 levels, we allow for more flexibility in the model. Different members of a family may have, for example, different probabilities of being nonsusceptible. For melanoma, it was found that a hierarchical frailty model achieved a better fit than an additive model (23).

In additive frailty models, heritability is usually defined as the proportion of the frailty variance associated with the genetic frailty (e.g., see Jonker and Boomsma (50)). In the present study, this proportion was found for second-level frailty alone, which is decomposed additively, and was estimated to be 75%. In the model, the probability of being susceptible varies across families. Genetics and environment shared by brothers account for 75% and 25% of this variance, respectively. Since the covariate does not affect the correlation structure between the family members, the proportion is constant across birth years. Note that the variation between families comes on top of the individual variation in susceptibility and does not mean that 75% of TGCTs are caused by genetic factors. Nevertheless, the lower proportion for seminoma than for nonseminoma further underscores the weaker heritable component in this subtype.

Although frailty models are well-described in the statistical literature (27–30), hierarchical frailty modeling has been developed in recent years (22–25). Applications seen so far have been for TGCT and infant mortality data (22, 24, 25), where no decomposition of second-level frailty was used, and melanoma (23). The model in the melanoma study differed slightly from the one used here (23), and a case-cohort sample was used. That should provide valid estimates (51), but the present study is nevertheless the first to have analyzed data from a complete population-based register with this methodology.

In conclusion, we have estimated familial risk of TGCT utilizing the flexibility of a hierarchical frailty model. The results are largely consistent with those of earlier studies, although the increased risk among brothers of TGCT patients was slightly lower than the risks estimated by other investigators. The FRR may be used to find relative risks for different constellations of family members compared with any other, and to our knowledge this study was the first to estimate the risk of TGCT for persons with 2 affected brothers and to take additional healthy brothers into account. Furthermore, the present results shed light on the difference in familial risk between seminoma and nonseminoma. As disease registries become older and computing power increases, the hierarchical frailty modeling framework has the potential to become a more important tool in studies of familial diseases.

ACKNOWLEDGMENTS

Author affiliations: Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway (Morten Valberg, Tron A. Moger, Marit B. Veierød, Odd O. Aalen); Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway (Tom Grotmol, Steinar Tretli); and Department of Health Management and Health Economics, Institute of Health and Society, University of Oslo, Oslo, Norway (Tron A. Moger).

This work was supported by the Norwegian Cancer Society (grant 171851). Prof. Odd O. Aalen was partly supported by the Research Council of Norway (grant 191460/V50).

Conflict of interest: none declared.
REFERENCES


