Estimation of Risk of Cancers before Occurrence of Acquired Immunodeficiency Syndrome in Persons Infected with Human Immunodeficiency Virus

Yueming Li, Matthew Law, Ann McDonald, Patty Correll, John M. Kaldor, and Andrew E. Grulich

There is methodological debate as to whether cohorts defined by acquired immunodeficiency syndrome (AIDS) diagnosis can be used to estimate risks of cancer in persons with human immunodeficiency virus (HIV) before AIDS. The authors compared risks of non-AIDS-defining cancers before AIDS in persons with HIV using a cohort based on AIDS diagnosis and a second cohort based on HIV diagnosis. National population-based registries of AIDS and HIV diagnoses to August 1999 were matched separately with the National Cancer Registry in Australia. Four analyses were performed. In analysis 1, follow-up was from 5 years before AIDS registration in 8,118 persons with AIDS. Analysis 2 was similar but adjusted expected numbers of cancers for decreased survival. Analysis 3 was based on 7,061 persons registered with HIV, with follow-up from the reported date of diagnosis. Analysis 4 was based on 2,112 AIDS cases previously reported with HIV, with follow-up from 5 years before AIDS diagnosis. In all analyses, follow-up ended at cancer diagnosis, death, 6 months before AIDS, or the end of available cancer data, whichever occurred first. For 10 types of cancer there were at least three cases in any one of the analyses. For these cancers there was no systematic pattern such that one analysis produced consistently higher or lower estimates than the others. These analyses suggest that cancer risk in persons with HIV before AIDS diagnosis may be estimated reliably based on cancer experience 5 years before AIDS. Am J Epidemiol 2002;155:153–8.

The ideal study design to assess associations between human immunodeficiency virus (HIV) infection and cancer occurrence would be to follow persons with HIV infection from seroconversion and to calculate cancer risk in subsequent years. However, non-acquired immunodeficiency syndrome (AIDS)-defining cancer is uncommon in persons with HIV, so large numbers of HIV seroconverters with follow-up over several years would be required to detect an association. To date, linkage of population-based AIDS and cancer registry data has commonly been used to assess the associations between HIV infection and cancers (1–4). In these studies, the period at risk for cancer of a person with AIDS has often been defined as the time from 5 years before AIDS diagnosis to 2 years after AIDS diagnosis. Standardized incidence ratios of cancer have been calculated by comparing the observed number of cancers in the AIDS cohort with the expected number, based on rates among the general population of the same age and sex.

Estimates of cancer risk in persons with HIV, based on cohorts defined by AIDS diagnosis, may be biased for several reasons. First, persons diagnosed with AIDS are likely to have been infected with HIV for several years. Second, inclusion in the analysis requires survival to AIDS and assumes that cancer risk in those persons with HIV infection who survive to AIDS diagnosis is similar to that in persons who die before AIDS diagnosis. Broadly, two biases are possible. First, there is the possibility that some persons with HIV infection may have developed cancer and died before AIDS diagnosis, which would lead to underestimates of cancer incidence. Second, if diagnosis and treatment of cancer resulted in faster progression to AIDS, then cancer incidence would be overestimated. These concerns have led some groups not to use cancer incidence occurring before AIDS in the calculations of relative risk and to use cancer incidence only after AIDS diagnosis in their analyses (2, 5).

Australia is well placed to assess whether the time before AIDS can be used to measure the associations between non-AIDS-defining cancers and HIV infection. National, population-based registrations of both AIDS and HIV diagnoses are available in the National AIDS Registry and the National HIV Database, respectively. The purpose of this study was to compare estimates of relative risk of non-AIDS-defining cancers before AIDS among persons followed up from 5 years before AIDS diagnosis, with estimates based on prospective follow-up of persons with HIV infection. Non-
AIDS-defining cancers are cancers other than Kaposi’s sarcoma, non-Hodgkin’s lymphoma, and cervical cancer.

MATERIALS AND METHODS

Data sources

In Australia, AIDS has been a notifiable condition since 1982 (6). The diagnosing doctor reports AIDS cases to the state/territory health authority, which then forwards the reports to the National Centre in HIV Epidemiology and Clinical Research for entry onto the National AIDS Registry. The information sought on each AIDS case includes the name code (first two letters of first and last name), sex, date of birth, date of AIDS diagnosis in Australia, area of residence at AIDS diagnosis, and date of death following AIDS.

The national surveillance system for newly diagnosed HIV infection was established in 1989 (7). Data have been collected, through the state/territory health authorities, from diagnosing laboratories and doctors on the first occasion of HIV diagnosis in Australia. The information sought on each case includes the name code, sex, date of birth, date of HIV diagnosis, and area of residence at HIV diagnosis. HIV seroconverters, a subset of HIV diagnoses, were defined as cases with a negative or indeterminate test or with a concurrent clinical diagnosis of HIV seroconversion illness within the 12 months before HIV diagnosis; these seroconverters have been registered since 1991 (8).

The National Cancer Registry in Australia has received data from individual state/territory cancer registries on cancer diagnosed in residents of Australia since 1982 (9). The information provided for each case enables record linkage to be performed and the analysis of cancer by type as defined by the International Classification of Diseases, Ninth Revision. Cancer rates among the general population were provided by the Australian Institute of Health and Welfare.

Linkage of AIDS, HIV, and cancer data

Cancer data available for the linkage covered diagnoses between 1982 and 1995–1998, depending on jurisdiction. Accordingly, HIV cases that had sufficient identifiers (name code, date of birth, and sex) and that were diagnosed during the same period and reported to the National HIV Database by the end of August 1999 were used for the matching. AIDS cases used for matching were those diagnosed between 1982 and June 1999 and reported to the National AIDS Registry by the end of August 1999. The National AIDS Registry and the National HIV Database were matched separately with the National Cancer Registry. The linkage procedure and its validation have been described in detail previously (10). Briefly, linkage was performed on the basis of name code, date of birth, and sex, using the computer program Automatch (11). A match was considered definite if there was an exact match on all three fields. Probable matches were manually reviewed and were accepted if supported by the same date of death. If the date of death was unavailable on one or both registries, probable matches were accepted if they had the same area of residence, based on postal codes, in the two matching registries (10).

Calculation of standardized incidence ratios before diagnosis of AIDS

Five analyses were performed according to the AIDS cases in the National AIDS Registry only, the cases of HIV infection in the National HIV Database only, the cases in both the National AIDS Registry and the National HIV Database, and the cases of HIV seroconversion only. Analyses 1 and 2 followed AIDS cases from the beginning of 1982 or 5 years before AIDS diagnosis (whichever occurred later), assuming that HIV infection had occurred before that time. To allow for the possibility that some persons with HIV infection may have developed cancer and died before an AIDS diagnosis, in analysis 2 the expected numbers of cancer were adjusted by assuming that HIV infection and treatment have no effect on the survival rates following cancer diagnosed before AIDS (4) and that South Australian cancer survival rates (12) applied throughout Australia. Analysis 3 followed HIV cases from the date of HIV diagnosis. With the same follow-up as in analysis 1, analysis 4 was restricted to AIDS cases who had previously been reported with HIV. Analysis 5 followed HIV seroconverters from the time of seroconversion. Follow-up in all the analyses ended at the date of cancer diagnosis, date of death, 6 months before AIDS diagnosis, or the end of the period of available cancer data, whichever occurred first. Standardized incidence ratios of cancer were calculated by comparing the observed number of cancers with the expected number, based on age-, sex-, and state-specific annual cancer rates among the general population. Ninety-five percent confidence intervals based on the Poisson distribution were calculated. Standardized incidence ratios and 95 percent confidence intervals of all non-AIDS-defining cancers combined were compared among all five analyses. The risks of specific types of cancer were compared among analyses 1–4 for those types of cancer with at least three cases in any one of these analyses. Specific types of cancer were also assessed in analysis 5 if any cases were diagnosed.

RESULTS

By August 1999, a total of 8,118 AIDS cases had been registered in the National AIDS Registry and were included in analyses 1 and 2. For the period of available cancer data, 7,061 cases of HIV infection, accounting for 46 percent of the cases in the National HIV Database during the study period, had sufficient identifiers in the National HIV Database to enable linkage with the National Cancer Registry and were included in analysis 3. Of the 7,061 HIV cases, 1,081 (15.3 percent) were HIV seroconverters, accounting for 75 percent of the cases in the HIV seroconverter registry, which is a subset of the National HIV Database, and were included in analysis 5. By the end of August 1999, 2,112 (29.9 percent) of the cases of HIV infection, including 220 (20.4 percent) of the HIV seroconvert-
ers, had developed AIDS and were included in analysis 4. There was no significant difference in the median time to AIDS, estimated by the Kaplan-Meier method, between the 1,081 HIV seroconverters and the 5,980 (= 7,061 – 1,081) persons diagnosed with HIV who had no available data for estimating the time of seroconversion (10 years vs. 9 years, \( p = 0.234 \) by log-rank test).

The observed total number of non-AIDS-defining cancers before AIDS diagnosis and the standardized incidence ratios in analysis 1 were 66 and 1.26 (95 percent confidence interval [CI]: 0.99, 1.60), respectively; 66 and 2.46 (95 percent CI: 1.90, 3.13) in analysis 2; 57 and 1.04 (95 percent CI: 0.80, 1.34) in analysis 3; 16 and 1.27 (95 percent CI: 0.78, 2.07) in analysis 4; and 5 and 1.12 (95 percent CI: 0.46, 2.68) in analysis 5.

For 10 types of cancer, there were at least three cases in any one of analyses 1–4. They were cancer of the anus (International Classification of Diseases, Ninth Revision, codes 154.2–154.3), bladder (code 188), colon (code 153), Hodgkin’s disease (code 201), leukemia (codes 204–208), lip (code 140), lung (code 162), melanoma (code 172), prostate (code 185), and testis (code 186). The number of observed cancer cases, standardized incidence ratios, and 95 percent confidence intervals calculated from analyses 1–4 are presented in figure 1. There was no systematic pattern such that one analysis produced consistently higher or lower estimates than the others. Standardized incidence ratios from the four analyses were in the same direction for five types of cancer (anus, bladder, Hodgkin’s disease, leukemia, and lip) and were within 95 percent confidence intervals of each other for the others, except for testicular cancer where standardized incidence ratios from analyses 1 and 2 were not included in the 95 percent confidence interval of the standardized incidence ratio from analysis 3. However, 95 percent confidence intervals from the four analyses did overlap.

Four types of cancer (lung, melanoma of the skin, testicular, and leukemia) occurred in the 1,081 HIV seroconverters (table 1). Based on the small observed number of cancer cases, all standardized incidence ratios but one were not significantly different from one. The estimated 95 percent confidence intervals were wider and contained all the corresponding standardized incidence ratios from the other four analyses except those for testicular cancer from analyses 1, 2, and 4 and that for lung cancer from analysis 1.

**DISCUSSION**

We have validated, for the first time, that estimates of relative risk of non-AIDS-defining cancers before AIDS, among AIDS cases followed up from 5 years before AIDS diagnosis, were quantitatively similar to estimates based on prospective follow-up of persons with HIV. Five separate analyses, based on the time period before AIDS, the time from HIV diagnosis, and the time from HIV seroconversion, gave estimates of relative risk of non-AIDS-defining cancers that were broadly consistent, both overall (except analysis 2) and for specific types of cancer. These analyses suggest that the relative risk of non-AIDS-defining cancers in persons with HIV before AIDS may be estimated reliably based on cancer experience 5 years before AIDS diagnosis.

The accuracy of the validation depends on the completeness of AIDS and cancer registrations and on the accuracy of the linkage procedure. The validation of linking cancer and AIDS register data (10) revealed that AIDS registration was 98 percent complete. The linkage was 99 percent sensitive and 100 percent specific in identifying the cases of non-Hodgkin’s lymphoma that occurred in persons who had been reported to both registers. A linkage study on the risk of cancer in persons with AIDS in New South Wales, where 60 percent of Australian AIDS diagnoses are made, found that 32 percent of Kaposi’s sarcoma cases that occurred as initial AIDS-defining illness were not reported to the cancer register and that Kaposi’s sarcoma was the only cancer with substantial underreporting (1). Underregistration of Kaposi’s sarcoma may be attributable to the fact that some other skin cancers such as squamous and basal cell cancers of the skin are not registrable in Australia.

Our study included about half the persons registered with HIV infection on the National HIV Database for the period of available cancer data. This proportion was limited by insufficient identifiers for HIV diagnoses on the National HIV Database, particularly in the early years of the registry. However, this is not crucial in a study where the follow-up was prospective from HIV diagnosis. For the purpose of estimating cancer risk before AIDS diagnosis in persons with HIV infection, it is important that the date of HIV diagnosis is close to the time of HIV seroconversion to ensure a full and unbiased follow-up. A good approximation of HIV diagnoses to the time of HIV seroconversion was suggested by the similarity in AIDS-free survival between the HIV seroconverters and the HIV diagnoses with no available data for estimating the time of seroconversion (median time to AIDS, 10 years vs. 9 years).

The method of adjustment for decreased survival after cancer diagnosis in analysis 2 was approximate and could theoretically have resulted in either under- or overestimates of the expected numbers of cancers. If HIV infection and treatment had, in fact, resulted in poorer survival after diagnosis of a cancer, then the method would underestimate standardized incidence ratios. If diagnosis of cancer and treatment for cancer resulted in faster progression to AIDS, then the standardized incidence ratios would be overestimated. The observed higher standardized incidence ratio from analysis 2 for all non-AIDS cancers combined and the standardized incidence ratios similar to those from other analyses for specific types of cancer suggest that this may be an overadjustment. We would recommend that in publications both adjusted and unadjusted analyses are presented.

The main limitation of this study was that, although it was based on national data, it had relatively limited statistical power, as the observed number of cancer cases was small for most of the specific types of cancer. However, there was no tendency toward a systematic difference among the five analyses. The standardized incidence ratios from analyses 1,
3, and 4 for all non-AIDS-defining cancers combined were very similar and were included in each other’s 95 percent confidence intervals. For specific types of cancer, there was no systematic pattern in difference of estimations of standardized incidence ratios and 95 percent confidence intervals among the five analyses. Nevertheless, the availability
Uncertainty about the validity of calculated cancer risk before AIDS in AIDS and cancer linkage studies has led to a variety of methodological approaches. A recent overview excluded cancers that occurred before AIDS (5); some have

of national registrations of HIV, AIDS, and cancer diagnoses in Australia provided us a unique opportunity to validate estimation of risks of non-AIDS-defining cancers before AIDS.

FIGURE 1. Observed cases and standardized incidence ratios of non-acquired immunodeficiency syndrome (AIDS)-defining cancer in persons with human immunodeficiency virus before AIDS diagnosis in Australia from 1982 to 1999 for cancers diagnosed between 1982 and 1998. SIR, standardized incidence ratio; CI, confidence interval; A1, analysis 1; A2, analysis 2; A3, analysis 3; A4, analysis 4.
calculated only the period prevalence ratios before AIDS (2), whereas others have presented the relative risk adjusted for decreased survival after cancer diagnosis (1, 4). Our analyses suggest that the calculation of the relative risk of cancers among AIDS cases followed up from 5 years before AIDS gives a reliable estimate of cancer risk in persons with HIV.

ACKNOWLEDGMENTS

The National Centre in HIV Epidemiology and Clinical Research is funded by the Commonwealth Department of Health and Aged Care.

The authors acknowledge the assistance of the Australian Institute of Health and Welfare and the South Australian Cancer Register, which undertook the record linkage component of this work, and the National HIV Surveillance Committee. The National HIV Surveillance Committee comprises the following state/territory and national representatives: Irene Passaris (Australian Capital Territory), Robert Menzies (New South Wales), Dr. Jan Savage (Northern Territory), Dr. Hugo Ree (Queensland), Therese Davey (South Australia), Neil Cremasco (Tasmania), Cathy Keenan (Victoria), Dr. Gary Dowse (Western Australia), and John Kaldor and Ann McDonald (National Centre in HIV Epidemiology and Clinical Research).

REFERENCES


TABLE 1. Observed and expected cases and standardized incidence ratios of non-AIDS*-defining cancer in HIV* seroconverters prior to AIDS, Australia, 1982–1999

<table>
<thead>
<tr>
<th>ICD-9* code</th>
<th>Type of cancer</th>
<th>Analysis 5 (1,081 HIV seroconverters)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Observed cases (no.)</td>
<td>Expected cases (no.)</td>
<td>SIR*</td>
<td>95% CI*</td>
</tr>
<tr>
<td>162</td>
<td>Lung</td>
<td>1</td>
<td>0.32</td>
<td>3.16</td>
<td>0.45–22.45</td>
</tr>
<tr>
<td>172</td>
<td>Melanoma skin</td>
<td>1</td>
<td>1.04</td>
<td>0.96</td>
<td>0.13–6.79</td>
</tr>
<tr>
<td>186</td>
<td>Testis</td>
<td>2</td>
<td>0.36</td>
<td>5.59</td>
<td>1.40–22.33</td>
</tr>
<tr>
<td>204–208</td>
<td>Leukemia</td>
<td>1</td>
<td>0.16</td>
<td>6.16</td>
<td>0.87–43.73</td>
</tr>
<tr>
<td></td>
<td>All non-AIDS defining</td>
<td>5</td>
<td>4.48</td>
<td>1.12</td>
<td>0.46–2.68</td>
</tr>
</tbody>
</table>

* AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus; ICD-9, International Classification of Diseases, Ninth Revision; SIR, standardized incidence ratio; CI, confidence interval.