Completeness of Cancer Registration in Limburg, the Netherlands

LEO J SCHOUTEN,* PAUL HÖPPENER,** PIET A VAN DEN BRANDT,†
J ANDRE KNOTTNERUS** AND JOS J JAGER*.†


The completeness of cancer registration in the IKL (Integraal Kankercentrum Limburg) cancer registry, Limburg, the Netherlands, was evaluated for the years 1988–1990 by means of the independent case ascertainment method. This study was performed in co-operation with the Registration Network of Family Practices (RNFP) of the University of Limburg. The RNFP is a centralized database used by general practitioners (GP), containing their patients' background variables and diagnoses. The contents of the two databases were compared using computerized record linkage. If the information from both databases differed, this was verified using the source forms of the cancer registry and the GP involved. By combining the information from both registries in this way it was determined which malignancies should have been registered by the cancer registry. The IKL cancer registry had recorded 307 of the 319 eligible malignancies (96.2%). Five of the 12 missed registrations could be attributed to systematic shortcomings in the notification procedures. The estimated completeness for all malignancies of the IKL cancer registry is comparable with the results from cancer registries outside the Netherlands which have been established for longer.

Incidence rates calculated from a registry with incomplete case ascertainment can lead to misinterpretation of geographical differences or trends in time. Several methods have been suggested for testing the completeness of a registry.1 Commonly, death certificates are used to measure the completeness of a cancer registry. In this approach, completeness is defined as the proportion of registered cases which were not first identified by death certificate.1 Cases reported by death certificate only and which have apparently not been reported before are an indication of incompleteness of case ascertainment. For non-fatal cancer cases this method does not give a good indication of completeness. Another suggested method is the historic data method in which the expected number of registered cases is calculated based on the basis of known incidence rates of another cancer registry. This method does not take into account the variation in time or in place of cancer incidence. Recently, the capture-recapture technique has been used to estimate the completeness of a registry.2 This technique has been widely used in zoology to estimate the size of animal populations. It is a promising technique, but its usefulness still has to be demonstrated.

A reliable method is independent case ascertainment, which involves comparing the total number of cases registered by the cancer registry with the results of an independent survey in the same area and for the same period. If this estimate can be derived from a linkage of individual records it is perhaps the most reliable method for the estimation of completeness of a registry1 and several cancer registries have used this method.3,4

This paper describes such a study on the completeness of a recently established cancer registry in the southeast of the Netherlands. It is one of nine regional cancer registries in the Netherlands which together comprise the Netherlands national cancer registry.

Regarding the completeness of the Netherlands cancer registry it was originally suggested that 8–15% of the cancer cases would not be recorded.5,6 However, a recent study estimated that its incompleteness was only 1.3%.7

Because these studies had their shortcomings, a new study was performed by the IKL cancer

---

* IKL Cancer Registry, Comprehensive Cancer Centre Limburg, PO Box 2208, 6201 HA Maastricht, the Netherlands.
** Department of General Practice and † Department of Epidemiology, University of Limburg, Maastricht, the Netherlands.
† Radiotherapeutic Institute Limburg, Heerlen, the Netherlands.
registry (IKL = Integraal Kankercentrum Limburg/Comprehensive Cancer Centre Limburg) using data from an independent source.

METHODS
The IKL Cancer Registry
The IKL cancer registry was established in 1984. Since 1986 all nine hospitals and seven pathology laboratories in the region have reported to the registry. Incidence data for the period 1986–1988 have been published recently.8,10

The cancer registry receives lists of newly diagnosed cases on a regular basis from the seven pathology departments in the region. In addition lists of hospitalized cancer patients are obtained from the medical records departments of the nine hospitals and the Radiotherapeutic Institute. Following this notification, the medical records of newly diagnosed patients (and tumours) are collected and the relevant information for the cancer registry is abstracted from the medical records by trained registration clerks. In the cancer registry identifying information (i.e. name, date of birth, gender, address, postal code, place of birth), tumour information (i.e. topography, morphology, staging, initial therapy, basis for diagnosis, date of incidence) and other administrative information (i.e. date of death) is stored. All malignancies, including non-infiltrating malignancies, are registered except basal cell carcinomas of the skin.

Due to privacy regulations death certificates cannot be used as an additional source of notification of cancer cases in the Netherlands. Tumour data are coded according to the International Classification of Diseases—Oncology (ICD–O).11

On 1 January 1988, the population of the IKL area consisted of 838,199 inhabitants (415,380 males and 422,819 females) (Source: Central Bureau of Statistics). On 1 February 1991, the IKL Cancer registry contained information on 22,941 patients with cancer.

Registration Network of Family Practices (RNFP)
The RNFP was established by the Department of General Practice of the University of Limburg. It is a collaboration of 15 partnerships of general practitioners (GPs) and serves as a sampling frame for research.12 However, two partnerships are located outside the IKL area and their data are not used for this study.

The GP has a central role in health care in the Netherlands. All Dutch citizens are registered with one general practice and treatment from a medical specialist requires referral from a GP. After consultation, specialists have to inform the GP of the diagnosis and treatment. The GPs collaborating in the RNFP have computerized their medical files. Background variables and diagnoses are coded and stored in computer files. Every 3 months the GPs send summaries of these files to the central database of the RNFP.12 In this database identification codes, not names, are included. Identification by name is only possible after linkage with the patient files in the individual practices. Diagnoses are coded according to the International Classification of Primary Care (ICPC).13

RNFP data collection was started in 1988. Medical history is coded and entered into the database for all patients registered with a general practice. The central database keeps separate files containing information on the deceased and patients who have moved away. In February 1991, the database contained information on 30,574 patients living in the IKL area (3.6% of the population). Between 1986 and 1990 375 patients had a diagnosis suggesting a possible malignancy.

Record Linkage and Verification
Record linkage between the 30,574 RNFP and 22,941 IKL records was performed in February 1991, after a pilot study in 1990.14 This study estimates completeness for the incidence years 1988, 1989 and 1990. A protocol developed in 1988 by the IKL cancer registry was used for the record linkage.15 For record linkage date of birth, gender, residential postal code and place of birth were available. All addresses in the Netherlands have been allocated a postal code, which consists of four digits and two letters. In this study only the four digits were available for record linkage. About 3000 to 5000 inhabitants have the same postal code of four digits. There are no data available on the uniqueness of date of birth and gender within a postal code area. The linkage key consisting of date of birth, gender and residential postal code has a sensitivity (i.e. proportion of all correct matches obtained) of 98.4% and a positive predictive value (i.e. proportion of reported matches that represented correct matches) of 91.1%.15 However, due to migration, the value of this key will decrease over time. In 1988, 4.4% of the population in the middle and southern part of Limburg moved to another municipality and thereby a different residential postal code.16 There are no data on migrations within municipalities and hence changes in residential postal code. The identifying information for linkage is presented in Table 1.

After record linkage tumour information (e.g. topography, morphology, date of incidence and treating specialist) and the place of birth on the IKL record was compared with the information in the RNFP record. If this information was in agreement,
Because neither the IKL or RNFP registry can be assumed to be complete a priori, the information from both registries and from the verification phase was combined to determine which malignancies from the populations of the general practices should have been registered by the IKL cancer registry. The RNFP cases that were not recorded by the IKL cancer registry were analysed for possible reasons for failure.

RESULTS

The Linkage

The linkage between the IKL registry and the RNFP database resulted in 315 linked records with regard to identifying information (Figure 1). Of these, 241 patients with 267 malignancies had a cancer diagnosis in both registries. For the other 74 patients with 74 malignancies the identifying information was identical but the RNFP file did not contain a cancer-related diagnosis. These diagnoses could have been missed by the RNFP or were false positive matches.

In all 136 patients were known with cancer in the RNFP database and had an incidence year from 1988 onwards but could not be linked to a record in the cancer registry. Linkage with a part of the linkage key (excluding the postal code, or year of birth, or month of birth, or day of birth or gender one by one) resulted in the linking of an additional 44 patients with 44 malignancies. In 38 of these 44 cases, the postal code was missing in one of the records or differed due to a migration. The remaining 92 RNFP records with cancer (that were linked to an IKL record) were verified with the original source of the information.

The 241 patients with 267 malignancies that were linked between the IKL and the RNFP databases and that contained a cancer-related diagnosis in the RNFP record, were compared 'manually' for agreement. In addition, the 44 patients with 44 malignancies that were matched with a part of the linkage key were also extensively checked. Of the total of 311 malignancies 24 had discordant information and had to be checked with the original source of information.

As a control group 288 patients were added to the enquiry as a sample of the unmatched records to check the validity of this part of the linkage. In total, 478 forms were sent to the GPs to obtain information with respect to the (presence of) cancer. In all 464 forms were returned, leaving 14 forms for which the GPs could not find necessary information. The procedure is summarized in Figure 1.

Results of the Verification

The results of the verification are summarized in Table 2, where the verified records are separated according

---

**TABLE 1 Available identifiers for linkage in the Registration Network of Family Practices (RNFP) file and the Comprehensive Cancer Centre Limburg (IKL) cancer registry file**

<table>
<thead>
<tr>
<th>Identifier</th>
<th>RNFP file</th>
<th>IKL-registry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of birth</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Place of birth</td>
<td>95.2</td>
<td>41.2a</td>
</tr>
<tr>
<td>Gender</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Residential postal code</td>
<td>99.1</td>
<td>95.0</td>
</tr>
</tbody>
</table>

a Percentages for all registrations with incidence years 1986–1990.

b Because of hospital registration practices in the Netherlands.

the matched records were considered to be identical. If a RNFP record with tumour information and an incidence year since 1988 could not be linked to an IKL record, a second linkage was performed with a part of the linkage key (e.g. year of birth excluded), to eliminate the possibility of a negative match due to an error in the identifying information. If after this second linkage the record still could not be linked to an IKL record, the information was verified with the GP. In addition, if an IKL record (with tumour information) had been linked with a RNFP record without a cancer diagnosis, the information was verified with the GP and the IKL (check of original cancer registry forms and if necessary, consultation with hospital administration).

Finally, a number of additional RNFP records were added to the records that had to be verified as an extra check for the validity of the linkage procedure. These records did not contain any cancer-related diagnosis and were not linked to any IKL record. If the linkage procedure has been performed well, then no cancer should be found in these records. These records were selected for a related study on epilepsy, with all age categories represented. In Figure 1 the linkage procedure is shown schematically.

Verification of disagreements regarding tumour information was performed by contacting the GP involved. The GP was asked to provide explicit information with regard to the presence and type of cancer. In this way we were able to determine whether the case matched the criteria for cancer registration. Malignancies (including non-infiltrating), diagnosed between 1988 and 1990 and living, at the time of diagnosis, within the IKL area were included. Borderline lesions or basal cell carcinomas of the skin were excluded.
FIGURE 1  Scheme for record linkage of RNFP and IKL files

T = Tumour; P = Patient.
to the outcome of the verification and the eligibility for cancer registration in this study. Reasons for non-eligibility are a diagnosis of basal cell carcinoma of the skin or a year of incidence before 1988 or after 1990. For the former eligibility could only be determined afterwards, because the ICPC code for skin cancer does not discriminate between melanoma and basal cell carcinoma. In the verification phase it also appeared, that in the RNFP records, recurrence was sometimes coded as a primary cancer, and that the true primary cancer was diagnosed before 1988.

In the sample of 288 control records no cancer was known to the GP for 276 cases. Only one patient in this group was known to have cancer by his GP and this had been diagnosed many years before 1988 and could not have been recorded by the IKL cancer registry. For 11 patients the GP could not find the necessary information.

Of the other 190 patients that were verified by GPs, 66 tumours were not registered by IKL. Of these 66, 54 were not eligible for this study. Some 40 cases had not been recorded by the RNFP (of which 37 were eligible). For another 23 cases the difference in coding could be attributed to coding errors in the RNFP (four eligible for this study). In seven cases the correct match was found after the verification (three are eligible). A few cases (N = 15) were notified to the cancer registry after 1 February 1991. For the majority of these cases the year of incidence was 1990. The explanation is the delay in the delivery of the information from medical records departments. For 14 cases the GP could not find the necessary information, of which 11 came from the control group. In total, of the 478 records that were verified by GPs 69 were eligible for cancer registration in this study.

The IKL cancer registry had missed 12 malignancies that were eligible for this study. Two were diagnosed by a medical specialist, but without clinical admission or a pathological examination. One patient was only treated by the GP and was never referred to a medical specialist. Two patients were treated outside the IKL area (one abroad and one in a Dutch hospital outside the IKL area). Seven patients were clinically admitted and/or the malignancy was pathologically verified. These malignancies should have been notified to the IKL cancer registry. This suggests some problems in the notification procedures.

Characteristics of Completeness
After the linkage and the verification it was determined which malignancies should have been registered by the IKL cancer registry. The information from both databases and the verification procedure were combined and the tumour data were coded according to ICD-O. Of the 274 patients with 287 malignancies with identical information in the IKL and the RNFP records 250 malignancies were eligible for this study. The other 37 malignancies could be excluded because of the incidence year (before 1988) or the morphology (basal cell carcinoma of the skin). Of the 478 patients

<table>
<thead>
<tr>
<th>Result of verification</th>
<th>Eligible* for cancer registration in the period 1988-1990</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Case missed by RNFP</td>
<td>37</td>
</tr>
<tr>
<td>Coding error by RNFP</td>
<td>4</td>
</tr>
<tr>
<td>Correct match found after verification</td>
<td>3</td>
</tr>
<tr>
<td>No cancer</td>
<td>-</td>
</tr>
<tr>
<td>False positive match&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Notified to cancer registry after 1 February 1991</td>
<td>13</td>
</tr>
<tr>
<td>Case not recorded by IKL</td>
<td>12</td>
</tr>
<tr>
<td>Not found by GP</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
</tr>
</tbody>
</table>

* i.e. Basal cell carcinoma of the skin, incidence year before 1988 or patient not known with cancer (see text).
<sup>b</sup> Cases from the sample of the RNFP records that were not linked to an IKL record and did not contain a cancer-related diagnosis.
<sup>c</sup> People with identical record linkage key (date of birth, gender, residential postal code), but after verification not considered to be a correct match.
verified by GPs 69 had a malignancy eligible for this study. It was therefore concluded that 319 malignancies should have been registered by the cancer registry. Of these 319 malignancies 270 (84.6%) were registered by both registries, 37 only by the IKL cancer registry and 12 only by the RNFP.

Overall, the IKL cancer registry had recorded 307 of the 319 eligible malignancies (96.2%) in the period 1988–1990 (95% confidence interval (CI) : 93.5–98.0%). In Tables 3, 4 and 5 the characteristics of completeness are summarized. Completeness in 1988 is lower (92.6%) than in 1989 and 1990 (97.3 and 97.6% respectively). Completeness varies by groups of tumour site (Table 4). Completeness is low for haematological malignancies (60%, N = 5), male genital cancer (88.8%, N = 18) and melanoma of the skin (83.3%, N = 6). Completeness for e.g. malignancies of the brain and unknown primary malignancies is 100%. It should be stressed however, that the numbers are small.

Completeness is similar for infiltrating malignancies (96.2%) and non-infiltrating (in situ) malignancies (96.2%). The completeness for non-microscopically confirmed malignancies is rather low (three out of six).

There is no clear relation between the patients’ age and the completeness of the cancer registry (Table 5). The only patient with cancer below the age of 10 years was not recorded by the IKL cancer registry. Above the age of 80 years completeness did not decline. Between males and females there is only a small difference in completeness of the cancer registry.

**DISCUSSION**

In this study we evaluated the completeness of the case ascertainment of the IKL cancer registry by record linkage with a registry of GPs. This method is probably the best method for measuring the completeness. However its success is largely dependent upon

---

**TABLE 3** Completeness of the cancer registry according to year of incidence

<table>
<thead>
<tr>
<th>Year of incidence</th>
<th>Number of eligible malignancies</th>
<th>Malignancies registered by the IKL registry (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>81</td>
<td>75 92.6</td>
</tr>
<tr>
<td>1989</td>
<td>112</td>
<td>109 97.3</td>
</tr>
<tr>
<td>1990</td>
<td>126</td>
<td>123 97.6</td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>307 96.2</td>
</tr>
</tbody>
</table>

**TABLE 4** Completeness of the cancer registry according to primary site

<table>
<thead>
<tr>
<th>Primary site</th>
<th>Number of eligible malignancies</th>
<th>Malignancies registered by the IKL registry (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head, neck cancers</td>
<td>8</td>
<td>8 100.0</td>
</tr>
<tr>
<td>Digestive tract</td>
<td>72</td>
<td>71 98.6</td>
</tr>
<tr>
<td>Respiratory tract</td>
<td>46</td>
<td>44 95.6</td>
</tr>
<tr>
<td>Squamous cell skin cancer</td>
<td>15</td>
<td>14 93.3</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>6</td>
<td>5 83.3</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>56</td>
<td>55 98.2</td>
</tr>
<tr>
<td>Gynaecological cancer</td>
<td>23</td>
<td>22 95.7</td>
</tr>
<tr>
<td>Male genital cancer</td>
<td>18</td>
<td>16 88.8</td>
</tr>
<tr>
<td>Urinary tract</td>
<td>35</td>
<td>34 97.1</td>
</tr>
<tr>
<td>Brain</td>
<td>6</td>
<td>6 100.0</td>
</tr>
<tr>
<td>Endocrine glands</td>
<td>6</td>
<td>6 100.0</td>
</tr>
<tr>
<td>Malignant lymphoma</td>
<td>14</td>
<td>14 100.0</td>
</tr>
<tr>
<td>Haematological malignancies</td>
<td>5</td>
<td>3 60.0</td>
</tr>
<tr>
<td>Unknown primary and other sites</td>
<td>9</td>
<td>9 100.0</td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>307 96.2</td>
</tr>
</tbody>
</table>
the reliability of the record linkage. The RNFP provided an attractive opportunity to compare its data with the IKL cancer registry, because the data are collected completely independently from the IKL registry.

In this study we used a record linkage protocol developed for use by the Dutch cancer registry for epidemiological research. Date of birth, gender and postal code were used as the key for record linkage. According to the earlier study, this key has a sensitivity of 98.4% and a positive predictive value of 91.1%. According to the protocol we increased the sensitivity by linking with parts of the key. The positive predictive value of the linkage key in the current study was only slightly lower (88.3%).

Verification of the matches, with the help of the GP and the original cancer registry forms, enabled us to eliminate false positive matches. There is a possibility that a GP overlooked the presence of a malignancy and a correct match was incorrectly judged as false positive. But if this error has occurred, it would only have led to a lower number of registrations in both the numerator and the denominator of the completeness percentage. The effect of these overlooked positive matches would be a lower estimate of the completeness than in reality.

In contrast, the possibility that a false negative match has been overlooked is small. All patients with a cancer code in the RNFP database, who were not linked to a record in the cancer registry, were considered as potentially missed cases by the cancer registry. This was only changed, if for example, the verification made it clear that the case did not meet the criteria for cancer registration.

Finally, in the control group no cancer was detected that should have been recorded by the IKL cancer registry. We therefore assume that the outcome of the record linkage is reliable.

The GP performs a key role in health care in the Netherlands. The GP will be informed by the medical specialist about his patients with respect to diagnosis and treatment. The GPs collaborating in the RNFP will code this information and enter it into the computer. However, this study revealed that the RNFP missed several cases. This could be attributed to several coding errors. Also, cancer diagnosed at autopsy was sometimes missed by the RNFP.

Linkage was conducted at patient level, however the cancer registry is concerned with malignancies and some patients had more than one (e.g. stomach and prostate cancer). In fact there were 307 patients with 319 eligible malignancies.

For all cancers (excluding basal cell cancer of the skin) the observed completeness of the cancer registry was 96.2%.

The methods of cancer registration in the Netherlands will lead to some systematic incompleteness. Patients treated by the GP only, patients treated ambulatory without pathological diagnosis and patients treated abroad will generally not be recorded by the cancer registry. However, this systematic underregistration seems to be limited. Only five out of the 12 missed cases could be attributed to these reasons (1.6% of all eligible cases).

Earlier warnings from GPs that the Netherlands cancer registry will have a high incompleteness because of the systematic deficiencies in the notification procedures, seem to be unjustified. This study confirms the results of a study of another regional cancer registry in the Netherlands which estimated systematic incompleteness at 1.3%. The seven missed cases should have been notified to the IKL cancer registry. It is conceivable, that these cases are not registered due to occasional errors in notification procedures. Most cases have multiple notifications and over time some of the missed notifications will be detected.

There seems to be a relation between completeness and the year of incidence. Completeness increased from 92.6% in 1988 to 97.6% in 1990. The first year that the IKL cancer registry was presumed to be complete was 1986 and errors in notification procedures
can most often be expected to occur at the start. The general incidence rates of the IKL cancer registry have been fairly constant since 1986. Also, some correct matches may be missed because of migration and hence changes in postal codes. This may have occurred in the first year especially.

Completeness is not associated with gender or age, with the exception of very young patients. The IKL cancer registry area does not have any specialised centres for childhood cancer. Such children would often be referred directly to such a centre after the first suspicion of malignancy. It was therefore anticipated that the IKL cancer registry might be incomplete for childhood malignancies. However, since 1989 all regional cancer registries in the Netherlands are fully operational and data will be submitted to the Netherlands cancer registry. From 1989 onwards, it has been possible to exchange information on patients who are treated outside the area of their own regional cancer registry. At the time of this study this exchange had not yet been effected. However, because of the low numbers a further study is warranted to allow more definitive conclusions.

The estimate of completeness for the IKL cancer registry (96.2%) is comparable to that of other registries outside the Netherlands which have not been established longer. Using the same method in northwest England completeness of cancer registration was estimated at 94%. Completeness of the registration of cervical cancers in Denmark was estimated at 97.8%. For all malignancies the completeness in the Danish cancer registry was estimated at 96.2%. In the Ontario Cancer Registry completeness was estimated around 95% after using the capture-recapture technique.

It should be noted that the IKL cancer registry is the only one of those mentioned that does not have access to death certificates. Malignancies notified by death certificates are an important means for many registries to improve completeness. Despite the lack of this notification source, the infrastructure of the Netherlands health care and the notification procedures used have made it possible to establish a cancer registry with high completeness in a short time.

ACKNOWLEDGEMENTS
The authors thank the GPs for their assistance in the conduct of this study.

REFERENCES
5 NIVEL. Continue Morbiditeits Registratie. Peilstatistieken Nederland: Jaarverslag 1984, 1985
6 Schaid E, Uenemann G. Over de volledigheid van de voorgestelde landelijke kankerregistratie. Huisarts en Wetenschap 1987; 30: 394-95
16 Centraal Bureau voor de Statistiek. Loop van de Bevolking per Gemeente 1988. 's Gravenhage, SDU-Uitgeverij, 1989
(Revised version received November 1992)