

NORTHERN TERRITORY CANCER REGISTER

DATA QUALITY

1981 – 2001

Contents

1	Summary	3
2	Introduction	4
3	Data Quality Indicators	5
3.1	Indicators of completeness of case ascertainment	5
3.2	Indicators of data accuracy	7
4	Methods	9
4.1	Variation of registration rates over time	9
4.2	Re-screening of notification sources	10
4.3	Comparison with registration rates in a similar population	11
4.4	Mortality: Incidence Ratio	11
4.5	Missing Data	11
4.6	Basis of diagnosis	12
4.7	Accuracy of indigenous status	12
5	Results	13
5.1	Completeness of case ascertainment	13
5.2	Data Accuracy	16
6	Discussion	18
	Appendix One:	
	Confirmation of HRNs	22
	Reference List	23

Suggested citation:

Condon JR, Zhao Y, Armstrong BK, Barnes A. Northern Territory Cancer Register, Data Quality 1981-2001. Department of Health and Community Services, Darwin 2004.

For further information on this publication, or for general enquires about NT cancer statistics, please contact:

Northern Territory Cancer Registry
 Health Gains Planning
 NT Department of Health
 and Community Services
 PO Box 40596
 Casuarina NT 0811
 Phone: 8922 7324
 email: ntghealth.gains@nt.gov.au

Printed by Government Printer
 of the Northern Territory
 ISBN: 0 9750651 4 9
 August 2004



Cancer Series Number Six.
Copyright 2004 John R Condon.

This publication is copyright. The information in this report may be freely copied and distributed for non-profit purposes such as study, research, health service management and public information provided that the source of information is acknowledged.

This is a joint publication of the Northern Territory Cancer Registry, the Menzies School of Health Research, Charles Darwin University and the Cooperative Research Centre for Aboriginal Health.

Acknowledgements

The authors gratefully acknowledge the assistance provided by the staff of Medical Records Departments at all NT hospitals, without whose considerable assistance the project could not have been completed. The assistance provided by the Victorian Cancer Registry and the Clinical Epidemiology Unit of the South Australian Department of Health and Community Services is also greatly appreciated.

This project was funded by postgraduate scholarships from the National Health and Medical Research Council and the Cooperative Research Centre for Aboriginal Health.

1. Summary

Monitoring and reporting on data quality is important to both identify areas of improvement required in the operation of cancer registries and to inform users of the reliability of cancer statistics produced by cancer registries. The Northern Territory Cancer Registry (NTCR) has not previously reported data quality indicators for cancer register data.

Indicators of completeness of case ascertainment and of data accuracy were assessed for the NTCR for cases diagnosed from 1981 (when the Registry commenced operation) to 2001. Data quality indicators were compared to indicators for other Australian cancer registries. Where possible, data quality was assessed separately for cancer registrations for Indigenous and non-Indigenous people. Several notification sources were re-screened to identify cases that had not been previously notified.

Case ascertainment was found to be approximately 40% incomplete for the period 1981–1986 and approximately 10% incomplete for the period 1987–1990. Mandatory notification of cancer cases by pathology laboratories was introduced in 1991. For the period 1991–1999 case ascertainment was estimated to be approximately 5% incomplete. An additional 2.6% of cases were identified during the re-screening of notification sources. After registration of cases identified by re-screening, case ascertainment for 1991–2001 was estimated to be incomplete by 2-3% when

compared to registration rates for other Australian cancer registries.

Data accuracy for cases registered as non-Indigenous was found to be comparable to that for other Australian cancer registries. However, data accuracy was lower for cases registered as Indigenous, possibly because Indigenous people are more likely to be diagnosed with advanced disease and less likely to undergo intensive investigations. There was a small degree of misclassification of indigenous status. When the effects of under-ascertainment of cases and of misclassification of indigenous status are combined it is estimated that Indigenous cancer incidence rates based on NTCR data would under-estimate actual cancer incidence by 15-20%, while non-Indigenous incidence rates would over-estimate actual cancer incidence by only 2-3%, if at all.

Overall, data quality for the NTCR is slightly lower than that achieved by other Australian cancer registries, and steps need to be taken to improve notification of cases, particularly from public hospital laboratories. However, the very large differences between NT Indigenous and total Australian incidence rates for specific cancer sites reported by the NTCR, ranging from 90% lower incidence of melanoma of the skin to 9 times higher incidence of cancer of the liver for NT Indigenous people, are not artefacts of inaccurate or incomplete cancer registration, but real differences in cancer incidence.

2. Introduction

There are considerable gaps in health statistics about Indigenous (Aboriginal and Torres Strait Islander) Australians.¹ Their overall health status is well documented to be much worse than that of other Australians,² but little is known about cancer in Indigenous people. Available information indicates that their cancer mortality is higher than that of other Australians, but that their cancer incidence may be slightly lower, with considerable differences in the incidence of cancers of specific sites.³ However, the accuracy of Indigenous cancer incidence statistics is uncertain, particularly because the accuracy of indigenous identification in cancer registers is unknown.

Accurate classification of indigenous status in cancer registers, as in other health data collections, is difficult. Considerable under-identification of indigenous people has been found in cancer registers in Australia⁴ and other countries.⁵ However, some registries in populations with a high proportion of indigenous people have been able to achieve and maintain a low level of misclassification of indigenous status over many years.^{6,7} Retrospective efforts to improve indigenous identification for previously registered cases have resulted in a considerable increase in the number of cases identified as indigenous for several Native American populations.^{5,8}

Only four of the eight state-based Australian cancer registers have reported cancer incidence statistics for Indigenous Australians. The South Australian Cancer Registry has reported Indigenous cancer incidence rates for 1977 to 1994, after efforts to identify Indigenous cases previously recorded as non-Indigenous, but the degree of remaining misclassification of indigenous status was not reported.⁴ The Western Australian Cancer Registry has reported Indigenous cancer incidence rates, but in 1996 and 1997 17-18% of registrations did not have data on indigenous status and the degree of misclassification of indigenous status when it was recorded has not been reported.⁹⁻¹¹ The Northern Territory Cancer Registry (NTCR) has reported cancer incidence and mortality statistics separately for Indigenous and non-Indigenous people since the 1980s,^{12,13} but has not previously reported the degree of misclassification of

indigenous status. The Queensland Cancer Registry has not reported cancer incidence specifically for Indigenous people, but has reported cancer incidence rates for remote communities where Indigenous people comprise 90% or more of the population.¹⁴

Concerns have been raised that misclassification of indigenous status by cancer registries may have led to under-estimation of incidence rates because of under-identification of Indigenous cases, particularly for cancers where Indigenous incidence rates have been reported to be lower than total Australian rates.^{15,16} No Australian cancer registries have reported a data quality audit of cancer registrations for Indigenous people.

Reporting of cancer statistics for Indigenous Australians is particularly important for the NTCR, since Indigenous people comprise a very high proportion of the NT population. The Northern Territory (NT) comprises approximately one-sixth of the Australian continent but contains only one percent of the Australian population, less than 200,000 people at the time of the 2001 census. Approximately 57,000 are Indigenous Australians, comprising 29% of the NT population and 13% of the total Indigenous Australian population.¹⁷ The majority of NT Indigenous people live in small isolated communities in the tropical north and arid centre of Australia, with restricted access to health services.

The NTCR commenced in 1981. Initially notification of cases by doctors and hospitals was voluntary. In 1991 the Cancer (Registration) Act came into effect, which made it mandatory for pathology laboratories to notify the NTCR of diagnosed cases of cancer and for the Registry of Births, Deaths and Marriages to notify deaths due to cancer; details of the operation of the NTCR have been published elsewhere.¹⁸ Notification is known to be incomplete in the early years of the Registry's operation; during the early 1990s a strenuous effort was made to complete case ascertainment from 1987.

A data quality audit of the NTCR was undertaken during 2001–2003 to assess the reliability of NT cancer statistics reported by the NTCR. The results of the audit are presented in this report.

3. Data Quality Indicators

Data quality is an essential element of the effective functioning of cancer and other disease registries. Poor data quality reduces or eliminates the ability of the registry to report accurate cancer statistics. Even if data quality is high, regular monitoring and reporting of data quality is necessary to maintain confidence in the accuracy of registry data. The overall quality assurance program for a cancer registry should be based on a specific legislative mandate or other authority under which the registry operates, with clearly defined functions and administrative responsibility for the operation of the registry.¹⁹ Details of these aspects of the operation of the NTCR have been published elsewhere.¹⁸

This report focuses on data quality – the completeness of registration of cases and the accuracy of information recorded for each case. A variety of indicators of case ascertainment and data accuracy are recommended for cancer registries by the International Agency for Research on Cancer.²⁰ Not all of these indicators could be used for the NTCR because necessary data was not recorded at the time of registration and resources were not available to undertake the more intensive investigations required for some indicators. Potential indicators of data quality for cancer registers, and their availability for the NTCR, are summarised below.

3.1 Indicators of completeness of case ascertainment

3.1.1 Stability of incidence rates over time

Cancer incidence rarely changes suddenly in large populations. A sudden increase or decrease in the registration rate probably indicates a change in case ascertainment rather than a change in incidence. Comparison of registration rate for several years combined enables identification of particular years in which there may have been a deficiency in the number of registrations. The 95% confidence interval of the average annual rate can be used to indicate the ‘expected’ range of the registration rate for individual years. However, this indicator is less precise in small populations (such as the NT) because there is

relatively greater random fluctuation in the number of cases diagnosed each year, and consequently a greater expected range for the annual registration rate, than in large populations. This method will also not detect a consistent under-ascertainment of cases over a long period of time. This data quality indicator could be used for the NTCR.

3.1.2 Comparison of registration rates in different populations

If the incidence of cancer in the NT population is the same as that in the rest of the Australian population, the registration rate should also be the same if case ascertainment by the NTCR is comparable to that of other Australian cancer registries. However, only the registration rate for the NT non-Indigenous population is comparable to that of the total Australia population because there is evidence that cancer incidence is different for Indigenous than other Australians.³ The population of the NT is also younger and has a higher proportion of males than the total Australian population, particularly in the population aged 50 and over in which the majority of cancers occur; in 2001 18% of the NT non-Indigenous population was aged 50 years and over, of which 58% were male, compared with 28% and 47% respectively of the total Australian population.^{21,22} Registration rates have to be age-sex-standardised to enable a valid comparison. This indicator could be used for the NTCR but its validity depends on the degree of misclassification of indigenous status in NTCR registrations (see Section 4.7).

3.1.3 Age-specific incidence curves

Cancer incidence usually follows a pattern of constant exponential increase over the human age range up to the last couple of decades; an unusual pattern of age-specific incidence rates may indicate a defect in case ascertainment.²³ This indicator could be used for the NTCR, although random variation in the small number of cases diagnosed in younger age-groups in the small NT population reduced the reliability of this indicator for younger age groups.

3.1.4 Childhood cancers

Childhood cancers are mostly a quite distinct group of cancers, different to the typical epithelial cancers of adulthood. The registration rate of cancer in childhood (0-14 yrs) is used as an indicator of case ascertainment because the incidence of childhood cancer varies less between populations than that of adult cancers.²³ However, because the incidence of childhood cancers is very low and there would be a large degree of random variation in the number of childhood cases occurring in the small NT population, this indicator was not used for the NTCR.

3.1.5 Mortality: Incidence ratio

The Mortality:Incidence ratio (M:I ratio) is the ratio of the number of cancer deaths compared to the number of cancer registrations in the same period. If cancer survival is similar in two populations and death registration is complete and cause of death is accurately recorded, then a difference between the M:I ratio of the two populations indicates a difference in the level of cancer registration. In the NT death registrations are complete and cause of death is coded in a reliable and consistent way. However, there is evidence that cancer survival for Indigenous people is considerably lower than that for other Australians,³ so the M:I ratio for the total Australian population could only be compared to that for the NT non-Indigenous population. Even for the NT non-Indigenous population there is evidence that cancer survival is slightly lower than total Australian survival rates (NTCR unpublished), so that the NT non-Indigenous M:I ratio may not be directly comparable to the total Australian M:I ratio.

3.1.6 Re-screening of usual notification sources

Re-examination of records at one or more of a registry's usual notification sources to identify cases that have not been previously notified and registered provides a direct indicator of under-ascertainment of cases. Re-screening assesses the effectiveness of notification and registration from usual sources, but does not assess the extent of under-ascertainment of cases that are not known

to the usual sources. Re-screening can also provide an indication of case ascertainment separately for Indigenous and non-Indigenous cases. For the NTCR re-screening of death registrations, pathology reports and hospital separations data from public hospitals and laboratories was possible, but re-screening for private laboratories and hospitals was not performed because of privacy and practical restrictions (see Section 6).

3.1.7 Average number of notification sources per case

A high average number of notification sources per registration indicates that several sources are actively notifying cases. This indicator could not be used for the NTCR because, although multiple notifications were received and processed for cancer registrations, prior to 2002 the source of each notification was recorded in the NTCR database only for the first pathology report received.

3.1.8 Proportion of cases with morphological verification of diagnosis (MV%)

Morphological examination (by histology, cytology or haematology) of a tissue specimen is the most accurate confirmation of the diagnosis of cancer, but not all cases are diagnosed in this way. A proportion of cases are diagnosed on radiological or clinical evidence without morphological investigation. If the MV% is very high this indicates that the registry is not being notified of cases from sources other than pathology laboratories. This indicator could be used for the NTCR, but only for cases diagnosed since 1994 because the basis of diagnosis was not recorded for prior years, and only for cases registered as non-Indigenous. The MV% for cases registered as Indigenous cannot be used as an indicator of completeness of case ascertainment in comparison to other Australian cancer registries because factors such as late diagnosis and less intensive investigation may cause differences in the proportion of Indigenous people with cancer who have morphological confirmation of their diagnosis.

3.1.9 Proportion of cases first notified to the registry at the time of death (Death Certificate Notifications, DCN%)

A low DCN% indicates that other sources of notification are working well, in that the registry has been previously notified of the diagnosis of cancer for almost all people who later die from cancer. This indicator could not be used for the NTCR because the source of the first notification received for each case was not recorded. For cases where the notification of death due to cancer was the first notification that the NTCR received, NTCR staff obtained information about the original diagnosis of cancer from the person's treating doctor or hospital, so that it cannot be determined which cases were first notified via death registration. The NTCR commenced recording the first notification source in 2002.

3.1.10 Independent case ascertainment

This is the most reliable indicator of completeness of case ascertainment, but also the most difficult to perform. It requires a specific search of data sources which contain all cases of cancer diagnosed in a defined area or for a specific type of cancer (such as the case register of a specialist cancer treatment centre) to identify all cases and determine what proportion have not been previously notified to the cancer registry. This indicator could not be used for the NTCR because no comprehensive source of cancer cases exists in the NT.

3.1.11 Capture-recapture study

This approach requires two or more lists of cancer patients. If there are two lists, the cases in each list are matched to determine the number of cases which appear in both lists and in only one or other of the two lists. From this, an estimate of the number of cases that do not appear in either list can be calculated. Three or more lists are used in a similar, but more complex, way. The various notification sources for the registry can be used for this purpose, if all notification sources for each case have been reliably recorded. The cases notified from each source are regarded as a 'virtual list' of cancer

cases. This method has several requirements that must be met relating to:

- accuracy of diagnosis
- ability to match all cases which appear on more than one list
- an equal likelihood for all cancer cases to appear on any particular list
- no movement of cancer patients in or out of the population between the time different lists are compiled
- independence between lists

This indicator was not used for the NTCR. Although three lists of people diagnosed with cancer were available after the re-screening of NTCR notification sources, these lists would not have satisfied the above criteria (particularly the third and fourth) without further work, the time and resources for which were not available.

3.2 Indicators of data accuracy

3.2.1 Basis of diagnosis

Cases diagnosed with histological, cytological or haematological confirmation of the neoplastic morphology of the tumour are more likely to have a correct diagnosis of cancer and of the primary site and morphological type of the cancer than cases diagnosed on only clinical or radiological evidence. A high proportion of cases registered with histological verification (HV%) or any morphological verification (MV%) of the diagnosis indicates that data accuracy is high for variables relating to the diagnosis of cancer. Morphological verification includes cytology and (for lymphoma and leukaemia only) haematology as well as histology. Conversely, a high proportion of cases registered only on the basis of a death certificate stating that the cause of death was cancer (DCO%), with no other verification of the diagnosis, indicates that some cases may not have an accurate diagnosis of cancer, the primary site of the cancer, the date of diagnosis or the morphological type of cancer. These three indicators were available for the NTCR for cases diagnosed since 1994; basis of diagnosis was not recorded for earlier years.

3.2.2 The proportion of cases with missing or unknown data

This indicator was calculated separately for the variables primary site, age/date of birth, sex, date of diagnosis, histological type and place of residence.

3.2.3 Re-abstracting and re-coding data from notification sources

This indicator is a direct measure of the accuracy of data abstraction and coding from the data sources that were used for the original registration. Although it is the most rigorous indicator of data accuracy, re-abstraction and re-coding is very labor-intensive and was not used for the NTCR because resources were not available.

3.2.4 Internal consistency

Checks for consistency between variables such as primary site and gender or age to identify biologically impossible combinations (such as cervical cancer in a male) or rare combinations which should be re-checked before data entry (such as prostate cancer in a child) can be used to identify errors in registry data. Edit checks were done for the NTCR at the time of data entry and during data management; they were not repeated for this project.

4. Methods

The following data quality indicators were measured for the NTCR for cases diagnosed in the period 1991–1999:

Completeness of case ascertainment

- variation in annual registration rates
- comparison of the age-sex-adjusted registration rate for NTCR cases registered as non-Indigenous with the total Australian registration rate for the period 1991–1999
- comparison of sex-adjusted age-specific registration rates for NTCR cases registered as non-Indigenous with total Australian registration rates for the period 1991–1999
- comparison of the age-sex-adjusted M:I ratio for NT non-Indigenous cases to that for the total Australian population for the period 1991–1999.
- re-screening of three notification sources (public hospital pathology laboratories, public hospital separations data and death registrations) for the period 1991–2001.

Data accuracy

- basis of diagnosis (HV%, MV% and DCO%) for cases diagnosed in 1994–2001
- the proportion of records with missing or unknown data for several individual data items

Case ascertainment was initially assessed for cases diagnosed in 1981–1999, in the NTCR dataset current at the end of February 2001. After completion of the re-screening of notification sources and registration of new cases identified during the re-screening, case ascertainment indicators were repeated for the period 1991–2001 in the dataset current at the end of July 2003. Data accuracy indicators were also assessed on the July 2003 dataset. Data quality indicators were compared to published statistics for the same indicators in total Australian cancer registrations or to published statistics for individual state cancer registries if total Australian comparisons were not available. NT registrations were not excluded from the comparison group because published statistics which excluded NT registrations were not

available and because excluding NT registrations, which comprise less than one percent of total Australian registrations, would have made no appreciable difference to the comparison group. All analyses included all invasive cancers except non-melanotic skin cancers which were not registered by the NTCR (ICD10 codes C00 to C97, excluding C44).

Total Australian registration rates and M:I ratios used for comparison to NTCR indicators were calculated using cancer registration data from the national cancer registrations database maintained by the National Cancer Statistics Clearing House²⁴ and cancer deaths data from the national mortality workbooks published by the Australian Institute of Health and Welfare.²⁵ Data on NT non-Indigenous deaths are obtained from the national deaths data unit record file (Australian Bureau of Statistics (ABS) unpublished dataset).

4.1 Variation of registration rates over time

Variation of registration rates over time was assessed by comparison of the age-sex-standardised registration rate (by year of diagnosis) for all-cancers combined for each year from 1981 to 1999. The age and sex distribution of the Australian 1991 population was used as the standard distribution for age-sex-standardisation. The mean annual age-sex-standardised rate for the years 1990–1999 combined, with 95% confidence intervals,²⁶ was used to indicate the expected range within which the annual registration rates should lie. The 95% confidence interval was calculated from the annual average number of registered cases in each age-group. The years before 1990 were not included when calculating the expected range because of prior concerns that case ascertainment may not have been complete during those years.

Case ascertainment appears to have been incomplete between 1981 and 1990 (Figure 1); cases diagnosed prior to 1991 were excluded from further examination of completeness of case ascertainment.

4.2 Re-screening of notification sources

Three notification sources were re-screened – histology and cytology reports from public pathology laboratories, hospital separations data from public hospitals and death registrations. Re-screening of the hospital separations dataset was undertaken before re-screening of pathology reports because of delays in obtaining pathology report data. Cases confirmed as eligible for registration with the NTCR were registered after re-screening each notification source.

Hospital separations dataset

Coded diagnosis data for all separations from the five NT public hospitals from January 1991 to May 2002 were searched, to identify all separations with any diagnosis coded as cancer (ICD9 codes 140-208 excluding code 173, or ICD10 codes C00 to C99, excluding C44). All separations with a cancer diagnosis code were matched by Hospital Registration Number (HRN) to the NTCR dataset to identify cases already registered. NT public hospitals have a shared client administration system (called 'Caresys'); each person has a single HRN for all public hospitals. The accuracy of HRNs recorded in the NTCR had been previously checked by matching the NTCR with the Caresys Client Master Index (CMI) (see Appendix). After electronic matching by HRN, unmatched separations were manually matched to the NTCR by name and date of birth. For separations which did not match to the NTCR, further information was obtained from the relevant hospital to confirm the diagnosis of cancer and obtain demographic and diagnostic information.

Pathology reports

An electronic key-word search was performed on histology and cytology reports from the pathology laboratories at the Royal Darwin Hospital (RDH) and the Alice Springs Hospital (ASH); the other three NT public hospital laboratories do not perform histology or cytology. All histology and cytology reports produced by the RDH laboratory for the years 1991–2001 were re-screened. At the ASH laboratory, histology reports were re-screened for the period

December 1994 to December 1996. ASH laboratory ceased performing histology in January 1997; from that time histology specimens received by the ASH laboratory were referred to the RDH laboratory for processing and were included in the re-screening of RDH laboratory reports. Histology reports produced by the ASH laboratory before December 1994 were not available electronically for re-screening; only paper reports were produced before this time. Paper histology reports at the ASH laboratory had been previously re-screened by NTCR staff in 1994. ASH laboratory has never performed cytology; all cytology specimens received by the ASH laboratory are referred to an interstate laboratory. Cytology reports were not available in electronic form for re-screening; only printed reports were returned to ASH laboratory from the interstate laboratory. Resources were not available to manually re-screen printed cytology reports.

The key-word search was performed by scanning the text of pathology reports for 45 text strings which may have indicated a diagnosis of cancer. The key-word list was modified from a list developed by the Victorian Cancer Registry.²⁷ All pathology reports which were positive to the key-word search were matched by HRN to the NTCR dataset to identify reports for cases which were already registered. All key-word positive reports which did not match to the NTCR dataset were read by one of the authors (JRC) to confirm the diagnosis of a cancer. For confirmed cancer reports the NTCR was then manually searched (by name and date of birth) to identify any cases which were already registered but which had not matched by HRN. For potentially un-notified cases remaining after the manual matching, demographic and diagnostic information was obtained from treating hospitals and doctors.

The key-word search of cytology reports was refined to identify reports containing any of three phrases that indicated the report did not diagnose cancer ('no evidence of malignancy', 'no malignant cells detected', 'negative Pap smear') or variants of these phrases. All cytology reports

from RDH laboratory for the years 1990 and 1991 (9,106 reports) were examined to test the specificity of these phrases; 1,008 of these reports were positive for a key-word indicating a diagnosis of cancer. Of these, 301 reports for people already registered on the NTCR (identified by matching on HRN) were excluded. Of the remaining 603 cytology reports, 571 reports were positive for one of the 'non-cancer' phrases. All 603 reports were read manually. Only one of the 571 reports which were positive for the 'non-cancer' phrases was confirmed as a new cancer registration. Thereafter cytology reports for later years that were positive for one of the 'non-cancer' phrases on key-word search were assumed to be negative for cancer without manual confirmation.

Death registrations

An electronic key-word search (using the same text strings as for the pathology report key-word search) was performed by the Registry of Births, Deaths and Marriages for all deaths from January 1991 to December 2002. All deaths identified by the key-word search were matched to the NTCR dataset by name, sex and dob; all deaths which did not match were manually checked against the NTCR database. For deaths for which the case was not registered further information was obtained from the relevant hospital and/or doctor to confirm the diagnosis of cancer and obtain demographic and diagnostic information for registration.

4.3 Comparison with registration rates in a similar population

To compare the NTCR registration rate with that of other Australian cancer registries the NT non-Indigenous cancer registration rate was compared with the total Australian rate for the period 1991–1999 in the February 2001 NTCR dataset. The age-sex-standardised registration rate and age-specific (sex-standardised) registration rates for all-cancers combined for cases recorded as non-Indigenous in the NTCR were compared with the equivalent total Australian registration rates. Age-specific rates were sex-standardised to a 50% male and 50% female sex distribution.

Prior to calculation of non-Indigenous registration rates, 241 registrations (8% of total registrations for 1991–1999) were found with data for indigenous status either missing or coded as 'unknown'; indigenous status for these records was obtained from the Client Master Index (CMI) of the 'Caresys' patient administration system in NT public hospitals (see section 4.6.3). Indigenous status could not be obtained from the Caresys CMI for 21 registrations (0.7% of total registrations); there was no indication from available information that these cases were Indigenous people. These 21 registrations were coded as unknown indigenous status and were included with non-Indigenous cases for calculation of registration rates.

Population estimates for the NT non-Indigenous population were derived by subtracting the estimated NT Indigenous population from the total NT Estimated Resident Population within each category of five-year age-group, sex and year. The mid-year NT Estimated Resident Population as published by the Australian Bureau of Statistics was used for the total NT population. The Indigenous population for each year from 1991–1999 was estimated by a back-casting method described elsewhere.²⁸

4.4 Mortality: Incidence Ratio

The M:I ratios for NT non-Indigenous people and the total Australian population were calculated as the ratio of the number of deaths registered with cancer as the underlying cause of death in 1991–1999 to the number of cancer cases registered with a date of diagnosis in the same period. The M:I ratios were age-sex-standardised to the age and sex distribution of the 1991 Australian population distribution.

4.5 Missing Data

The proportion of registrations with missing (or unknown) data for individual variables was calculated for cases diagnosed in the period 1991–2001, using the August 2003 NTCR dataset, separately for cases identified as Indigenous and as non-Indigenous. The

proportion missing was calculated for each of the variables sex, date of birth, indigenous status, community of residence, postcode, primary site and histological type. The proportion missing for NTCR variables was compared to published reports of the same indicators from other Australian cancer registries, for those variables for which published reports were available.

4.6 Basis of diagnosis

The proportion of cases registered on the basis of histological verification (HV%), any morphological verification (histology, cytology or haematology) (MV%) or death certificate only (DCO%) was calculated for cases diagnosed in the period 1994–2001, for the NTCR dataset current at August 2003. Data on the basis of diagnosis of each registration were not recorded for cases diagnosed in 1991–1993. Published reports of the same indicators from other Australian cancer registries were used for comparison.

4.7 Accuracy of indigenous status

Several indicators of completeness of case ascertainment involve comparison of cases registered as non-Indigenous by the NTCR with total Australian registration rates. The reliability of these indicators depends on the accuracy of indigenous status data in the NTCR. Information on indigenous status in the NTCR is obtained primarily from the Caresys CMI, and also from death certificates, hospital medical records and contact with treating doctors. Most pathology reports received by the NTCR do not include information on indigenous status.

The accuracy of indigenous status as recorded in the Caresys CMI had previously been evaluated by comparison with self-identified indigenous status obtained by direct interview of hospital inpatients.²⁹ The indigenous status

data in the CMI agreed with self-identified indigenous status for 94% of people. Most of those for whom CMI data did not agree with self-identification were Indigenous people who were recorded as non-Indigenous in the CMI. Indigenous status as recorded in the CMI had a Predictive Value Positive (PVP) of 97.5% (95% CI 94.4-99.2), and Predictive Value Negative (PVN) of 91.7% (86.8-95.2), compared with direct interview of hospital inpatients. Separate results for males and females and for particular age groups were not reported.

To estimate the degree of misclassification of indigenous status in the NTCR a random sample of 500 NTCR records was extracted with a year of diagnosis in the period 1991–1999. Indigenous status as recorded in the CMI was obtained for each person. Indigenous status recorded in the NTCR was not identical to that in the CMI because other information sources were also used by the NTCR to determine indigenous status. A CMI record could not be found for 32 cases (6.4%). The number of people who would self-identify as Indigenous and non-Indigenous in the remaining 468 cancer cases was then estimated using the previously estimated PVP and PVN for indigenous status data in the Caresys CMI. The ratio of the estimated to actual number of Indigenous and non-Indigenous cases in the sample of 468 was calculated to estimate the accuracy of identification of indigenous status in the NTCR.

Misclassification of indigenous status could not be estimated for specific cancer sites, or for specific age-groups, because the PVP and PVN for CMI indigenous status data were not available for individual age-groups or cancer sites.²⁹

Stata version 8 was used for statistical analysis.

5. Results

5.1 Completeness of case ascertainment

5.1.1 Stability of registration rates over time

NT cancer registrations were approximately 40% incomplete for the years 1981–1986, and approximately 10% incomplete for 1987–1990, compared to the registration rate for the years 1991–1999 (Figure 1a). After re-screening notification sources (see Section 5.1.2), the annual registration rates for the years 1991–2001, based on the NTCR dataset current at August 2003, were relatively stable and within the expected range for all years (Figure 1b).

5.1.2 Re-screening of notification sources

Most potentially un-registered cancer cases identified by electronic re-screening of notification sources were not found, after further investigation, to be registrable cases; most were either not cancer, already registered or were not eligible for registration by the NTCR (Table 1). As a result of the re-screening of notification sources, an additional 87 cases were registered

with a date of diagnosis in the period 1991–1999. This was an increase of 2.6% on the number of cases registered as diagnosed in the same period at the end of February 2001. 15 of these were Indigenous, an increase of 2.4% in the number of cases registered as Indigenous, compared to an increase of 2.8% for non-Indigenous cases. However, other data quality checks found that several existing registrations were interstate residents or were diagnosed earlier than 1991; the net increase in registrations diagnosed in the period 1991–1999 was 38 (3190 c/w 3152).

For 72 potential un-notified cases found during the re-screening of notification sources no further information could be found from hospitals or treating doctors to determine whether the diagnosis was correct and the cancer registrable in the NT (Table 1). If the final outcome for these cases were the same as for other potentially un-notified cases, only a small proportion would have been confirmed as registrable cases.

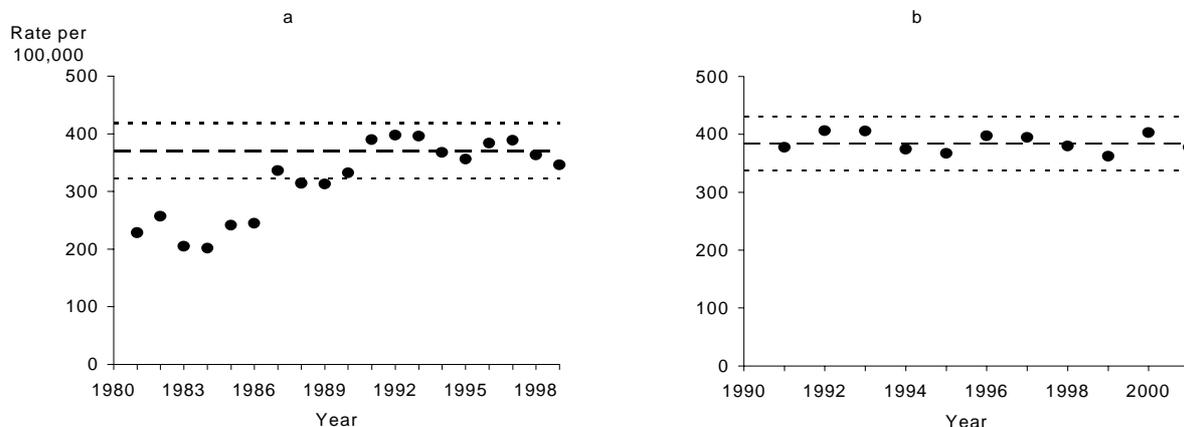


Figure 1: Annual registration rate^a and expected range^b, all-cancers combined, (a) before re-screening, 1981–1999 and (b) after re-screening, 1991–2001

- a. standardised to the Australian 1991 age-sex-population distribution
 b. expected range based on age-sex-standardised rate and 95%CI for 1990–1999 in figure (a) and 1991–2001 in figure (b).

Table 1: Outcome of re-screening of notification sources

	Hospital Separations	Pathology Reports	Death Registrations
Total records searched	560,980	99,685	9929
Positive on keyword search	10,522	13,600	2186
No. remaining after match to NTCR	658	7,747	269
Result of follow-up			
already registered	197	244	66
not a registrable cancer	83	7221	54
interstate resident	167	138	76
diagnosed before 1991	26	6	7
insufficient information	25	29	18
second pathology report ^a	na	50	na
new registrations 1991-1999 ^b	28	49	10
new registrations 2000-2001 ^b	83	9	12
new registrations 2002 ^b	49	1	26

a. pathology reports for which a new registration had been confirmed from a previous report

b. year of diagnosis

na. not applicable

5.1.3 Comparison to total Australian registration rates

The NT non-Indigenous registration rate for 1991–1999, based on the NTCR dataset current at February 2001, was 1.0% lower than the total Australia registration rate for the same period (Table 2). For the dataset current at August 2003 (which included new registrations found during the re-screening of notification sources), the NT non-Indigenous registration rate was 1.6% higher than the total Australian rate. Neither difference was statistically significant. NT non-Indigenous age-specific registration rates were similar to or higher than total Australian rates for all age-groups except 85+ (Figure 2).

The NT non-Indigenous M:I ratio, based on the February 2001 dataset, was considerably higher than the total Australian M:I ratio, indicating that there was under-registration of non-Indigenous cases by the NTCR. The NT non-Indigenous ratio based on the August 2003 dataset was only slightly lower than that based on the earlier dataset.

5.1.4 Proportion of cases with morphological verification (MV%)

The MV% for NTCR cases registered as non-Indigenous was comparable to that of other Australian cancer registries (Table 3).

Table 2: Indicators of completeness of case ascertainment, NTCR non-Indigenous cases compared with total Australian cases

Indicator	NTCR dataset at:			Total Australia
	Feb-01	Aug-03	Aug-03	
	Cases diagnosed:			
	1991-1999	1991-1999	1991-2001	
Registration rate (95% CI)	398 (379-417)	410 (390-429)	408 (391-425)	402 (401-403) ^a
M:I ratio ^b	0.53	0.52	na	0.44 ^a
MV% ^c	93 ^d	93 ^e	93 ^f	88-94 ^g

a. Australia 1991–1999

b. the ratio of the number of cancer deaths to the number of cancer cases

c. proportion of registrations with any morphological (histology, cytology or haematology) verification of diagnosis

d. for cases diagnosed 1997–1999

e. for cases diagnosed 1994–1999

f. for cases diagnosed 1994–2001

g. Australian state cancer registries, 1993–1997³⁰

na. not available: cancer mortality data for 2001 was incomplete at the time of analysis.

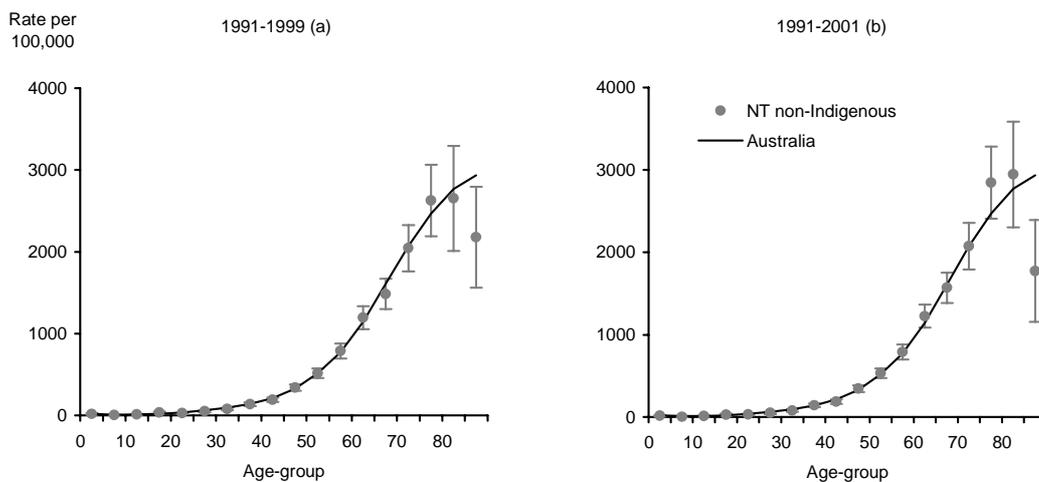


Figure 2: Age-specific registration rates, NT non-Indigenous and total Australia

a. NT non-Indigenous registration rates based on the February 2001 NTCR dataset

b. NT non-Indigenous registration rates based on the August 2003 NTCR dataset;

Australian rates for 1991–1999

5.2 Data Accuracy

5.2.1 Basis of diagnosis

For cases registered as non-Indigenous, the proportion of cases with histological verification (HV%) was slightly lower than the HV% for the NSW Cancer Registry in 1998, while the proportion registered with any morphological verification (MV%) was at the upper end of the range of MV% for all other Australian cancer registries in the period 1993–1997 (Table 3). The proportion of cases registered with only a death certificate to verify the diagnosis of cancer (DCO%) was within the range for other Australian cancer registries.

Data quality was not as high for cases registered as Indigenous; both the HV% and MV% were lower than for cases registered as non-Indigenous, although the DCO% was within the range for other Australian cancer registries (Table 3).

Table 3: Basis of diagnosis, NTCR cases diagnosed 1994–2001

Basis of diagnosis	Indigenous	Non-Indigenous	Comparison
HV%	71	85	87-88 ^a
MV%	85	94	88-94 ^b
DCO%	0.8	0.6	0-1 ^b

a. NSW Cancer Registry, male and female, 1998³¹

b. range for other Australian state cancer registries, 1993-1997³⁰

5.2.2 Missing or unknown data

No registrations had missing data on sex, date of birth, postcode, primary site or histological type. 0.26% of registrations had missing or unknown data for suburb/community of residence. Indigenous status was recorded as ‘unknown’ for 0.9% of registrations; in comparison, 17-18% of Western Australian Cancer Registry registrations diagnosed in 1996 and 1997 had missing data on indigenous status;^{10,32} comparable information has not been reported by other Australian cancer registries.

For cases registered as non-Indigenous, the proportion of cases with unknown primary site was within the range for other Australian cancer registries (Table 4). The proportion with unknown morphology was higher than the proportion with unknown primary site, but no comparable figures have been published for other Australian cancer registries.

For cases registered as Indigenous, both the proportion of cases registered with unknown primary site and with unknown morphology was approximately double that for cases registered as non-Indigenous. A very high proportion of cases registered as Indigenous had only an approximate year of birth rather than an exact date of birth.

Table 4: Proportion of registrations with missing or unknown data, NTCR cases diagnosed 1991–2001

Data quality indicator	Indigenous (%)	Non-Indigenous (%)	Comparison (%)
Exact date of birth	57	99	np
Unknown primary site	10.2	4.5	3.4-4.7 ^a
Unknown morphology	14.8	7.2	np

a. Australian state cancer registries 1993–1997³⁰

np: not published (ie, no comparable statistics are available for other Australian cancer registries)

5.2.3 Indigenous status

500 NTCR registrations with a date of diagnosis in the period 1991–1999 were randomly selected for comparison of indigenous status as recorded in the NTCR with that recorded in the Caresys CMI (Table 5). 32 registrations could not be matched to the Caresys CMI. Of the 468 remaining, 79 registrations were recorded on the CMI as Indigenous and 389 as non-Indigenous. After adjustment by the previously estimated PVP and PVN for indigenous status data in the

CMI,²⁹ it was estimated that 109.5 of these people would have self-identified as being Indigenous (compared to 93 recorded as Indigenous in the NTCR), and 358.5 would have self-identified as non-Indigenous (compared to 375 in the NTCR). It is thus estimated that the NTCR over-estimates the number of non-Indigenous cancer registrations by a net 4.6% and under-estimates the number of Indigenous registrations by a net 15%.

Table 5: Indigenous status, NTCR compared with CMI

	CMI	Adjusted for PVP and PVN	
		Indigenous	Non-Indigenous
Indigenous	79	77.1	1.9
non-Indigenous	389	32.4	356.6
Total	468	109.5	358.5
NTCR	468	93	375
Ratio NTCR to estimated self-identification		84.90%	104.60%

6. Discussion

For cases diagnosed since 1991, case ascertainment by the NTCR appears to have been slightly lower than for other Australian cancer registries. Before the re-screening of notification sources the NTCR registration rate for non-Indigenous cases was similar to the total Australian rate (Table 2), but if adjusted for misclassification of indigenous status the non-Indigenous registration rate would have been approximately 5.6% lower than the total Australian rate. Re-screening of notification sources found an additional 2.6% of cases diagnosed in 1991–1999, confirming that there had been a small under-ascertainment of cases before re-screening. Private health services (the two private pathology laboratories and the Darwin Private Hospital) were not able to be included in the re-screening of notification sources; a small number of un-notified cases would probably be found if these notification sources could be re-screened. After registering the additional cases found during re-screening, the NT non-Indigenous registration rate was slightly higher than the total Australian rate, but still 3% lower if adjusted for misclassification of indigenous status.

Before 1991, case ascertainment was obviously incomplete, particularly before 1987 (Figure 1). In 1993 the NTCR attempted to retrospectively improve case ascertainment for the years 1987 to 1992; it appears that this effort was successful for cases diagnosed in 1991 and 1992, but not so for those diagnosed in 1987–1990.

Under-ascertainment does not appear to be restricted to any particular age group, with the possible exception of age-group 85+ (Figure 2). In 2001 only 0.3% of the NT population were aged 85²¹ and over and in 1991–2001 only 1.5% of registered cancer cases were in this age-group; under-ascertainment in this age-group would have had little effect on overall case ascertainment.

The proportion of un-notified Indigenous cases identified by re-screening (2.4%) was similar to that for non-Indigenous cases (2.8%), indicating

that the degree of under-ascertainment was similar for Indigenous and non-Indigenous cases. No other indicators of case ascertainment were available specifically for Indigenous cases.

Comparison of the NT non-Indigenous and total Australian M:I ratios suggested that there was a greater degree of under-ascertainment in the NT than was indicated by other measures of case ascertainment. However, cancer mortality rates for the NT non-Indigenous population are slightly higher than for the total Australian population³³ and cancer survival rates slightly lower (NTCR unpublished); the NT non-Indigenous M:I ratio would be expected to be slightly higher than the total Australian ratio even if NTCR case ascertainment was comparable to that of other Australian cancer registries. The M:I ratio is also somewhat ‘volatile’; when standardised to the age-sex distribution of the NT non-Indigenous 1991 population instead of to the Australian 1991 population the M:I ratios were much closer (NT non-Indigenous 0.39, total Australian 0.36), indicating only a small degree of under-ascertainment of cases. The M:I ratio does not provide strong evidence of a high degree of under-ascertainment of cases, particularly given inconsistency of the M:I ratio with other indicators of case ascertainment.

For NTCR cases registered as non-Indigenous, data accuracy appears to be similar to that of other Australian cancer registries. However, data accuracy is not as high for cases registered as Indigenous. A recent study in the NT has found that, for four of five specific sites studied, Indigenous people are more likely to be diagnosed with advanced cancer than non-Indigenous people;³⁴ an earlier study in South Australia found similar results.⁴ Primary site may be harder to determine and investigations less intensive for people diagnosed with advanced disease. Lower accuracy of diagnostic data may also be due to the reluctance of many Indigenous people from remote communities to undergo intensive investigations and treatment in hospitals away from their home communities.

The high proportion of Indigenous people with only an approximate year of birth is a feature of the NT Indigenous population, many older members of which were born in remote locations before the advent of European record-keepers.²⁸ Over 40% of cases registered as Indigenous did not have an exact date of birth, with their year of birth obtained from their estimated age at first regular contact with government services or church missions. These estimates were made between 1900 and 1960 during childhood or early adult life for many NT Indigenous people currently aged over 40 years. Their estimated year of birth is probably accurate to within a few of years, but there is no direct way to assess this.

Limitations

Cancer incidence in the NT non-Indigenous and total Australian populations

Comparison of NT non-Indigenous and total Australian registration rates is a valid indicator of completeness of case ascertainment only if cancer incidence is the same, or very similar, in the two populations. Many measures of health status and risk factors for the NT non-Indigenous population are similar to the total Australian population (after adjustment for the younger age distribution of the NT population), with one exception relevant to cancer incidence: smoking prevalence is higher in the NT non-Indigenous than total Australian populations. In a telephone survey in 2000, 29% of NT non-Indigenous respondents reported that they were current smokers compared to 24% of respondents sampled from the total Australian population in the 2001 National Health Survey.^{35,36} The incidence of lung cancer and other smoking-related cancers would be expected to be higher in the NT non-Indigenous population than the total Australian population.

On the other hand, cancer incidence may be slightly lower for the NT non-Indigenous than total Australian populations because of a 'healthy migrant worker' effect. The NT non-Indigenous population is, in part, a migrant-worker population with very high mobility from and to other Australian states; in 1996 34% of NT non-

Indigenous residents reported that they had been resident in another state or overseas five years earlier, compared to 22% of Australian Capital Territory residents and 6-13% of residents in the six states (ABS unpublished data, 1996 Census). Cancer incidence may be slightly less in the NT population than in the total Australian population if the health effects of cancer risk factors or the pre-diagnosis symptoms of cancer reduce the likelihood of people moving to the NT for employment. No evidence is available of the potential effect of high population mobility on the incidence or prevalence of cancer, or other diseases, in the NT.

Re-screening

Not all notification sources could be re-screened. In particular, the two private laboratories and hospital separations data from the Darwin Private Hospital were not re-screened. Hospital separations data and pathology laboratory reports from public hospitals and laboratories were able to be accessed because the public hospital and laboratories were operated by the same organization as the NTCR (the NT Department of Health and Community Services). The hospital separations dataset for the Darwin Private Hospital was not screened for un-notified cases because hospital separations data were not available for the DPH before 1998 and because of privacy limitations on access to data from the private hospital; notification from hospitals is not required under the NT Cancer (Registration) Act. Re-screening of pathology reports from the two private laboratories operating in the NT was not undertaken because of practical, privacy and resource issues; the information management services of both laboratories were located in other states, there were privacy restrictions on providing NTCR staff with a dataset of all pathology reports to undertake the key-word search and no funds were available to provide resources to the pathology laboratories to undertake the key-word search and manual confirmation of cases themselves. Re-screening of private notification sources may have identified a small additional number of unregistered cases.

Implications for NT cancer statistics

Cancer incidence rates reported by the NTCR for 1991–2001 will only slightly under-estimate, if at all, cancer incidence for the total NT population. For the non-Indigenous population, misclassification of Indigenous cases as non-Indigenous will cause over-estimation of non-Indigenous incidence rates by 4-5%, while under-ascertainment may cause under-estimation by perhaps 2-3%; non-Indigenous incidence rates will probably be within a few percent of actual cancer incidence.

For the Indigenous population, misclassification of indigenous status will cause under-estimation of Indigenous incidence rates by approximately 15%, while under-ascertainment may also cause under-estimation by perhaps 2-3%; Indigenous incidence rates reported by the NTCR will probably under-estimate actual Indigenous cancer incidence by 15-20%.

The higher proportion of Indigenous cases with unknown primary site will cause site-specific cancer incidence rates to be lower than they would have been had primary site been known for more cases; Indigenous site-specific incidence rates would be approximately 6% higher if the proportion of NT Indigenous registrations with unknown primary site were similar to that for other Australian cancer registries (and the likelihood of a case being registered with unknown primary site was the same for all specific cancer sites). However, the differences in incidence of cancer at specific cancer sites between the NT Indigenous and total Australian populations range from 90% lower incidence for melanoma of skin to 9 times higher for cancer of the liver,³⁷ these differences are not artefacts produced by incomplete or inaccurate cancer registration data.

The high proportion of Indigenous registrations with estimated year of birth may alter age-standardised incidence rates if there is a bias towards under-or over-estimation of year of birth.

There is no evidence available about either possibility. The effect of misclassification of age on cancer incidence rates cannot be estimated.

Survival rates calculated from NTCR data may be affected by under-ascertainment of cases if un-registered cases have different survival than registered cases. This may have occurred if people who died from cancer were more likely to have been registered than people who survived, which could have occurred if notification of cancer deaths by the Registry of Births, Deaths and Marriages was more complete than notification of cases from pathology laboratories and the hospital separations dataset. However, unregistered cases were not exclusively people who have survived (since unregistered cases were found during re-screening of death data), and the estimated under-ascertainment of cases was small; inclusion of an additional 2-3% of cases would make little difference to survival rates.

Misclassification of indigenous status would also affect survival rates reported separately for Indigenous and non-Indigenous people. Analysis of cancer survival cannot be as readily adjusted for incorrect indigenous identification (predominantly Indigenous people misclassified as non-Indigenous) as can analysis of cancer incidence. If survival rates were considerably lower for Indigenous than non-Indigenous cases, misclassification of Indigenous cases as non-Indigenous would cause under-estimation of non-Indigenous survival rates. However, it is estimated that only 4-5% of cases registered as non-Indigenous were actually Indigenous people, so non-Indigenous survival rates would be under-estimated by only a very small amount. In addition, Indigenous people mis-classified as non-Indigenous may have better cancer survival than Indigenous people correctly recorded as Indigenous, which may reduce any under-estimation of non-Indigenous survival rates. If this is the case, Indigenous survival rates would also be under-estimated to a small extent.

Conclusion

Data quality should be regularly monitored and reported by all cancer registries, including the NTCR. Data quality indicators reported here demonstrate that cancer statistics for the NT population for 1991–2001 are reliable, though not perfect; cancer incidence rates for the non-Indigenous population are close to correct and rates for the Indigenous population are under-estimated to a moderate degree, mainly because of misclassification of some Indigenous people with cancer as non-Indigenous.

Re-screening of notification sources demonstrated that reporting by public hospital pathology laboratories was incomplete in 1991–2001, and possibly remains so. This confirms prior concerns by NTCR staff about completeness of notification from public hospital laboratories; notification procedures from these laboratories need to be reviewed and improved. Completeness of notifications from private laboratories could not be assessed. Re-screening of reports at these laboratories should be done in future if possible; if not, the NTCR should record data about notification sources to enable monitoring of indirect indicators for private laboratories such as the number of notifications received from individual laboratories and identifying which laboratories were found to have failed to notify cases during follow-up of notifications from hospital separations and deaths data.

Misclassification of indigenous status is a crucial issue for the NTCR. Assessment of misclassification of indigenous status in this study was indirect, through the previously estimated accuracy of indigenous status recorded in the Caresys CMI. The accuracy of indigenous status recorded in the CMI has not been assessed since 1997 and should be repeated at regular intervals, not only for the NTCR but also for analysis of hospital statistics and other NT health data which relies on the CMI for demographic data. Direct assessment of indigenous status should also be considered by the NTCR, as has been done previously for the CMI.²⁹

Even more important than reporting data quality is improving data quality through better case ascertainment and accurate indigenous identification. In the small NT population, with well-developed primary and secondary health care services and only a handful of hospitals and pathology laboratories to deal with, this should be easier to achieve than elsewhere is Australia.

Although there are aspects of data quality which need to be improved by the NTCR, data quality is high for cases registered since 1991. This assessment of NTCR data quality provides assurance that cancer statistics for the NT population, both Indigenous and non-Indigenous, can be used with confidence to inform the general public, health services and government policy-makers about cancer incidence, mortality and survival for the NT population.

Appendix One: Confirmation of HRNs

The NTCR records the HRN of each person registered in the NTCR who has a record in the Caresys CMI. For some cases, duplicate entries exist in the CMI – all duplicate HRNs are recorded in the NTCR. For cases where information is obtained from the Darwin Private Hospital, the DPH Unit Registration Number (URN) is also recorded.

Initial assessment of HRN data recorded in the NTCR found that there was a high proportion of cases with missing or incorrect data. A matching process was undertaken to match the NTCR database with the CMI to identify the correct HRNs for all NTCR cases where a CMI record could be found. A six-step matching process was undertaken with manual confirmation of positive matches between each step. In 1996 the CMI was culled to remove records of people who were deceased or had not attended an NT public hospital for several years; culled records were archived to a separate database. After completion of matching for the currently active CMI, the

process was repeated with the archived database for NTCR registrations which had not matched to the current CMI. Only the first five steps were performed for the archived data. After matching with the archived database, a manual search of the CMI was performed for NTCR registrations which remained unmatched. 89.3% of registrations were matched by electronic matching, and an additional 2.1% by manual matching (Table 6).

The matching criteria used in each step were:

1. surname, given name, sex and date of birth
2. surname (first 2 characters), given name (first 2 characters), sex and date of birth
3. surname and given name
4. surname, sex and date of birth
5. given name, sex and date of birth
6. surname (first character) & sex & year of birth, or
NTCR HRN matched with CMI HRN

Table 6: Matching of NTCR to Caresys CMI, results of each matching step

CMI source and matching criteria	Initial registrations	Potential match	Confirmed match	Unmatched
<i>Current CMI</i>				
1	4998	2927	2927	2071
2	2071	291	271	1800
3	1800	544	237	1563
4	1563	96	84	1479
5	1479	208	107	1372
6	1372	119	114	1258
<i>Archived CMI</i>				
1	1258	272	264	994
2	994	147	143	851
3	851	629	151	700
4	700	157	150	550
5	550	61	17	533
Manual matching	533		104	429

Reference List

- (1) Aboriginal and Torres Strait Islander Health and Welfare Information Unit. The Aboriginal and Torres Strait Islander health information plan. 1997. Canberra, Aboriginal and Torres Strait Islander Health and Welfare Information Unit, for Australian Health Ministers Advisory Council and Australian Institute of Health and Welfare.
- (2) Australian Bureau of Statistics. The health and welfare of Australia's Aboriginal and Torres Strait Islander peoples, 1999. Cat. No. 4704.0. 1999. Canberra, Australian Bureau of Statistics and Australian Institute of Health and Welfare.
- (3) Condon JR, Armstrong BK, Barnes A, Cunningham J. Cancer in Indigenous Australians: a review. *Cancer Causes Control* 2003; 14(2):109-121.
- (4) South Australian Cancer Register. Cancer incidence, mortality and case survival in the South Australian Aboriginal population. *Epidemiology of cancer in South Australia, 1977 to 1996*. Adelaide: South Australian Cancer Registry, South Australian Health Commission, 1997: 11-19.
- (5) Sugarman JR, Holliday M, Ross A, Castorina J, Hui Y. Improving American Indian Cancer Data in the Washington State Cancer Registry using linkages with the Indian Health Service and Tribal Records. *Cancer* 1996; 78(7 Suppl):1564-1568.
- (6) Lanier AP. Epidemiology of cancer in Alaska Natives. *Alaska Med* 1993; 35(4):245-247.
- (7) Key CR. Cancer incidence and mortality in New Mexico, 1973-77. *Natl Cancer Inst Monogr* 1981;(57):489-595.
- (8) Partin MR, Rith-Najarian SJ, Slater JS, Korn JE, Cobb N, Soler JT. Improving cancer incidence estimates for American Indians in Minnesota. *Am J Public Health* 1999; 89(11):1673-1677.
- (9) Threlfall TJ, Whitfort M, Thompson JR. Cancer in Aboriginals. *Cancer Incidence and Mortality in Western Australia, 1992-1994*. Perth: Health Department of Western Australian, 1996.
- (10) Threlfall TJ, Thompson JR. Cancer incidence and mortality in Western Australia, 1997. 1999. Perth, Health Department of Western Australia.
- (11) Threlfall TJ, Thompson JR. Cancer incidence and mortality in Western Australia, 1996. 1998. Perth, Health Department of Western Australia.
- (12) d'Espaignet ET, Measey M.L., Condon JR, Jelfs P, Dempsey KE. Cancer in the Northern Territory 1987-1993. 1996. Darwin, Territory Health Services.
- (13) Selva-Nayagam S, Pearce M, Croft C. Cancer. In: Condon JR, Warman G, Arnold L, editors. *The Health and Welfare of Territorians*. Darwin: Territory Health Services, 2001: 105-112.
- (14) Coory M, Thompson A, Ganguly I. Cancer among people living in rural and remote Indigenous communities in Queensland. *Med J Aust* 2000; 173(6):301-304.
- (15) Carrick S, Clapham K, Paul C, Plant AJ, Redman S. Breast cancer and Aboriginal and Torres Strait Islander women. 1996. Sydney, National Breast Cancer Centre.
- (16) Australian Bureau of Statistics and Australian Institute of Health and Welfare. The health and welfare of Australia's Aboriginal and Torres Strait Islander peoples, 2003. Cat. No. 4704.0. 2003. Canberra, Australian Bureau of Statistics and Australian Institute of Health and Welfare.
- (17) Australian Bureau of Statistics. Census of population and housing: Aboriginal and Torres Strait Islander people, Australia. Cat. No. 2034.0. 1998. Canberra, Australian Bureau of Statistics.
- (18) Condon JR, Zhao Y. Northern Territory Cancer Registry: data collection, analysis and reporting procedures. 2004. Darwin, Northern Territory Department of Health and Community Services.
- (19) Hilsenbeck SG. Quality control. In: Menck H, Smart C, editors. *Central cancer registries: design, management and use*. Chur: Harwood Academic Publishers, 1991: 131-178.
- (20) Parkin DM, Chen V, Ferlay J, Galceran J, Storm H, Whelan S. Comparability and quality control in cancer registration. 19. 1994. Lyon, International Agency for Research on Cancer. IARC Technical Report.

NORTHERN TERRITORY CANCER REGISTER

- (21) Australian Bureau of Statistics. Population by age and sex, Australian states and territories, June 2001. Cat. No. 3201.0. 2001. Canberra, Australian Bureau of Statistics.
- (22) Australian Bureau of Statistics. Australian Demographic Statistics, September Quarter 2002. Cat. No. 3101.0. 2003. Canberra, Australian Bureau of Statistics.
- (23) Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG. Cancer registration: principles and methods. Lyon: International Agency for Research on Cancer, 1991.
- (24) National Cancer Statistics Clearing House. Interactive Cancer Data. 2003. Canberra, Australian Institute of Health and Welfare.
- (25) Population Health Unit. Australian longterm mortality trends workbooks (data file). 2003. Canberra, Australian Institute of Health and Welfare.
- (26) StataCorp. Stata base reference manual, volume 1 A-F: release 8.0 (pg 299). College Station, TX: Stata Corporation, 2003.
- (27) Victorian Cancer Registry. Victorian Cancer Registry: Information kit for pathology laboratories. 2003. Melbourne, Cancer Council Victoria.
- (28) Condon JR, Barnes A, Cunningham J, Smith LR. Indigenous mortality and demographic characteristics in the Northern Territory: changes over four decades. 2004. Darwin, Cooperative Research Centre for Aboriginal Health. Occasional Papers.
- (29) Condon JR, Williams DJ, Pearse MC, Moss E. Northern Territory hospital morbidity dataset: validation of demographic data 1997. 1999. Darwin, Territory Health Services.
- (30) Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J, editors. Cancer Incidence in Five Continents, Volume VII. Lyon: International Agency for Research on Cancer, 1997.
- (31) Coates M, Tracey E. Cancer in New South Wales, incidence and mortality 1998 and incidence for selected sites 1999. 2001. Sydney, NSW Cancer Council.
- (32) Threlfall T, Thompson J. Cancer incidence and mortality in Western Australia, 1996. 1998. Perth, Western Australian Cancer Registry.
- (33) Dempsey KE, Condon JR. Mortality in the Northern Territory 1979-1997. 2000. Darwin, Territory Health Services.
- (34) Condon JR, Armstrong B, Barnes A, Selva-Nayagam S, Elwood M. Stage at diagnosis and cancer survival of Indigenous and non-Indigenous people in the Northern Territory, 1991-2000 (in press). 2004. Melbourne National Cancer Control Initiative.
- (35) d'Espaignet ET, Measey ML, Dal Grande E. Northern Territory health and wellbeing survey, non-Indigenous population, December 2000. 2002. Darwin, NT Department of Health and Community Services.
- (36) Australian Bureau of Statistics. National health survey 2001: summary of results. 4364.0. 2002. Canberra, Australian Bureau of Statistics.
- (37) Zhao Y, Condon JR, Garling L. Cancer incidence and mortality, Northern Territory 1991-2001. 2004. Darwin, Northern Territory Department of Health and Community Services.

Health Gains Planning - Selected Publications

Cancer

Kinmonth TJN, **Cancer in the Northern Territory 1981**, NT Department of Health.

Kinmonth TJN, **Cancer in the Northern Territory 1982**, NT Department of Health.

Kinmonth TJN, **Cancer in the Northern Territory 1981-83**, NT Department of Health, 1985.

Honari M, Saint-Ives I, **The Northern Territory Cancer Report 1981-1985, An Analysis of Ethnic, Age & Sex Specific Rates**, Epidemiology and Assessment Branch, Northern Territory Department of Health, Darwin, 1987.

d'Espaignet ET, Measey ML, Condon JR, Jelfs P, Dempsey KE, **Cancer in the Northern Territory 1987-1993**, Epidemiology Branch, Territory Health Services, 1996.

Zhao Y, Condon JR, Garling LS, **Cancer Incidence and Mortality Northern Territory 1991-2001**, NT Cancer Registry, Department of Health and Community Services, Darwin, 2004.

Condon JR, Zhao Y, Northern Territory Cancer Registry, **Northern Territory Cancer Registry Data Collection, Analysis and Reporting Procedures**, Department of Health and Community Services, Darwin, 2004.

Mortality

Dempsey KE, Condon JR, **Mortality in the Northern Territory 1979-1997**, Territory Health Services, Darwin, 1999.

Condon JR, Williams DJ, Pearce MC, Moss E, **Northern Territory Hospital Morbidity Dataset, Hospital Morbidity – Validation of Demographic Data 1997**, Territory Health Services, Darwin 1998.

Measey ML, d'Espaignet ET, Cunningham J, **Mortality and Morbidity Attributable to Smoking – Northern Territory 1986-1995**, Territory Health Services, 1998.

Other Publications

Markey PG, d'Espaignet ET, Condon JR, Woods M, **Northern Territory Midwives Collection- Trends in the Health of Mothers and Babies Northern Territory 1986-1995**, Territory Health Services, Darwin, 1998.

Northern Territory Midwives Collection, Mothers and Babies 1999, NT Perinatal Information Management Group, Department of Health and Community Services, Darwin, 2002.

Condon JR, Warman G, Arnold L, **The Health and Welfare of Territorians**, Epidemiology Branch, Territory Health Services, Darwin, 2001.

Measey ML, d'Espaignet ET, Dal Grande E 2002, **Northern Territory Health and Wellbeing Survey, Non -Indigenous Population, Northern Territory**, Department of Health and Community Services, Darwin, 2000.

Beaver C, Zhao Y, **Investment Analysis of the Aboriginal and Torres Strait Islander Primary Health Care Program in the Northern Territory**, Department of Health and Ageing, Australian Government, Canberra, 2004.

Zhao Y, Guthridge S, Magnus A, Vos T, **Burden of Disease and Injury in Aboriginal and Non-Aboriginal Populations in the Northern Territory**, MJA 2004: 180:498-502, 2004

Upcoming Publications

Li SQ and Guthridge SL, **Mortality in the Northern Territory 1981-2000, Part 1 All-Cause Mortality**, Department of Health and Community Services, Darwin, (2004).

Villeseche A, Li SQ, Measey MA, **Northern Territory Physical Activity Survey 2003, Non-Indigenous Population**, Department of Health and Community Services, Darwin, (2004).

You J, Guthridge S, **NT Injury Profile 1991-2001**, Department of Health and Community Services, (2004).