

An estimate of limited duration cancer prevalence in New Zealand using 'big' data

Naomi Brewer, June Atkinson, Parry Guilford, Cristin Print, Tony Blakely, Andrea Teng

ABSTRACT

AIMS: Increases in cancer survival may increase cancer prevalence and demand for healthcare. We aimed to estimate cancer prevalence in the New Zealand population.

METHODS: We used national linked health, social and census datasets from the Stats NZ Integrated Data Infrastructure to identify the number of New Zealand residents who had at least one cancer diagnosis in New Zealand. We included all primary cancers recorded on the New Zealand Cancer Registry from January 1995 to June 2013, and used the 2013 census for demographic and socioeconomic data.

RESULTS: On 30 June 2013, 140,600 of 4,438,900 (3.2%) New Zealand residents had been diagnosed with cancer in the last 18.5 years. In ≥ 15 year olds, the age-standardised prevalence of cancer diagnosed 0 to ≤ 1 year, and >1 to ≤ 5 years, prior to 30 June 2013 was 0.4% and 1.1% in men and 0.3% and 0.9% in women, respectively. Over the 18.5-year period prevalence was greatest in the oldest ages, European/Other, highest qualified, highest income, least deprived, ex-smokers, and Canterbury, Bay of Plenty and Nelson/Marlborough District Health Boards (age-standardised).

CONCLUSIONS: Groups with the highest survival and the greatest access to healthcare had the highest cancer prevalences.

A larger number of people are being diagnosed with cancer in New Zealand,¹ while cancer mortality is decreasing,² meaning more people living beyond a cancer diagnosis. It is important to understand how many of these people there are since cancer prevalence estimates are a direct measure of the demand for healthcare, psychosocial support, practical assistance and social services.

Cancer prevalence can be defined as the number of people in a specified population who are still alive after at least one cancer diagnosis.³ This measure of prevalence, termed total prevalence, includes a heterogeneous group comprised of those who were diagnosed many years ago (and are therefore likely to be completely cured) and people who were diagnosed only recently (and therefore may still be undergoing treatment).³ Limited duration prevalence is an indicator of the number of people

who have been diagnosed over a defined (usually relatively recent) period of time. The meaning of limited duration prevalence depends on the specific cancer site, but limited duration prevalence within one year may indicate the number of people in initial treatment, greater than five years is often used as an indicator of the number of people in complete remission, and 10 years tends to indicate the number of cured patients (although statistical cure time varies by cancer site⁴).³ Cancer prevalence is influenced by both the number of people who are diagnosed with cancer and the number of people who die from cancer; thus a high incidence rate coupled with a low mortality rate will result in a high prevalence.

There is a scarcity of estimates of the prevalence of cancer in New Zealand. A Ministry of Health (MoH) report on urban/rural health comparisons using data from the 2002/03 New Zealand Health Survey (which

involved interviews with approximately 12,900 people aged ≥ 15 years) indicated that the age-standardised self-reported lifetime prevalence of cancer in urban females was 6.8%, the proportion in rural females was 5.8%, and the corresponding proportions in males were 4.8% and 5.2%.⁵ Globocan 2018 estimates of the number of five-year prevalent cases of all cancers (International Classification of Diseases 10th revision (ICD-10) C00–97) in males of all ages in New Zealand were 77,878 (3.3%), and 55,838 (2.3%) for females of all ages.⁶ When excluding non-melanoma skin cancer (C44) the corresponding estimates were 36,814 (1.6%) for males, and 37,171 (1.5%) for females.⁶ However, the Globocan prevalence estimates are based upon projections and are not available by ethnicity or other demographic factors.^{6,7}

We aimed to provide estimates of limited duration cancer prevalence in the New Zealand population on 30 June 2013 using national linked health, social and census data. We aimed to examine how prevalence differed by age, sex, ethnicity, socioeconomic status (SES) and region, and how the prevalence differed by the major cancer sites.

Methods

We used the Stats NZ Integrated Data Infrastructure (IDI; <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/>). We included only individuals who were in the central IDI spine (data refreshed in March 2017; see Appendix Figure 1). The spine is an ever resident population that includes individuals with data on New Zealand tax, births or long-term immigration visa records; and is a centralised list of people to which health and census datasets were each probabilistically linked. The probabilistic data linkages to create the spine and link the datasets were undertaken by Stats NZ using IBM QualityStage version 8.5.⁸ The proportion of unique individuals within the MoH dataset who were linked to the spine (ie, the link rate) was 85%, with a false-positive rate of 0.8%.

We abstracted details of people with primary cancers (ICD-10 C00–96, excluding basal cell epithelioma and squamous cell carcinoma of the skin (see below)) recorded on the New Zealand Cancer Registry (NZCR)

from 1 January 1995 up to 30 June 2013. This latter date was the closest date to the most recent available New Zealand Population Census (5 March 2013) for which the IDI estimated resident population (IDI-ERP; see below) was available. We used the census as the source of most of the demographic data (see below). Thus we estimated 18.5 year prevalence, as well as one-, five- and 10-year prevalence. The 18.5 year limited duration prevalence is an estimate of the percentage of the population diagnosed with cancer at some point in the 18.5 years prior to 30 June 2013 and who were still alive at 30 June 2013.

The NZCR was established in 1948, and the Cancer Registry Act came into effect in 1994, making cancer registration mandatory.^{9,10} Information was not available for cancers diagnosed outside New Zealand, eg, for recent migrants. The NZCR does not record basal cell epithelioma and squamous cell carcinoma of the skin except when of the skin of genital organs.

We estimated person prevalence, and hence considered only the first (according to diagnosis date) primary cancer diagnosed in each person, therefore only including each person once in our analysis results for all cancers. The exception to this was when we estimated the prevalence by specific cancer types. We identified the first case of one or more different cause-specific cancers where a person may have had more than one primary tumour. We included people diagnosed in New Zealand, at all ages, in the New Zealand resident population. We used the IDI-ERP dataset for 30 June 2013 produced by Stats NZ (at the time of data collation the IDI-ERP was only estimated annually for 30 June (ie, it was not available for March 2013)).¹¹ The methodology is based on that used by Stats NZ, and selects people into the population if they have had activity in a variety of administrative datasets in the IDI in the 12 months before the reference date (eg, 30 June 2013). It then removes people who died or moved overseas (defined as being out of New Zealand for ≥ 6 months). Sex was based on census data but if that was missing data from MoH datasets was used, and if that was also missing then data from Stats NZ IDI was used. Age was calculated at 30 June 2013 based on the Stats NZ date of birth information or from the NZCR data

if the Stats NZ data was missing. All other demographic variables were derived from the census (approximately 25% of records in the IDI-ERP dataset were not linked to the census and therefore were 'missing' demographic variables (see Appendix Figure 1)).

Statistical analysis

Analysis involved categorising prevalent cancers by cancer site, time since diagnosis ($0 \leq 1$, $1 < 5$, $5 \leq 10$, $10 < 18.5$ and $0-18.5$ years) and the following demographic variables: sex, age group (≤ 14 , 15–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84, 85+ years), partnership status (partnered, non-partnered), standardised highest qualification (no qualification, school, post-school), urban or rural area of usual residence, smoking status (never smoked regularly, ex-smoker, regular smoker), equivalised household income (tertiles based on national income pattern), New Zealand Deprivation Index 2013 (NZDep2013) (small area index of SES;¹² quintiles), prioritised ethnicity (Māori, Pacific, Asian, European/Other [non-Māori, non-Pacific, non-Asian]),¹³ DHB of residence and Cancer Network of residence.

For each of these demographic variables, we present the age-standardised (to the WHO World Standard Population; <https://seer.cancer.gov/stdpopulations/world.who.html>) prevalence proportions up to 18.5 years since diagnosis. Age-standardised prevalence rate ratios were calculated comparing demographic variable categories in Excel using methods described by Rothman et al (2008).¹⁴ Analysis was done using Stata version 14.0 MP (Stata Corp LLC, College Station, Texas, US).

To comply with the confidentiality requirements associated with using IDI data the count results have been random rounded to base 3. Count data produced from cancer registrations in the IDI and proportions based on those data were suppressed if the underlying unrounded count was less than six.

Results

There were approximately 4,438,900 people in the residential population at 30 June 2013 with age and sex recorded in the dataset. Among the 4,438,900 people there were 140,600 diagnosed with cancer in New Zealand between 1 January 1995 and

30 June 2013, giving a crude prevalence of 3.17% ($140,600/4,438,900 \times 100$) and an age-standardised prevalence of 2.15% for men and 2.10% for women (Table 1). The mean age of the New Zealand population (in the IDI-ERP) was 39 years (standard deviation (SD) 24 years) in 2013, while the mean age of people living with a cancer diagnosis was 66 (SD 12 years).

Age-standardised prevalence was greatest in the oldest age groups, European/Other and Māori women, those who were the most qualified, had the highest income, lived in the least deprived areas, ex-smokers, and those in the Canterbury, Bay of Plenty and Nelson/Marlborough DHB regions (Table 1). The lowest prevalences were in Asian and Pacific people, and in the Counties Manukau, Capital & Coast and Hutt Valley DHBs. Prostate cancer did not appear to explain the differences by DHB (Appendix Table 1 and Appendix Figure 2). Ethnicity and socioeconomic differences in prevalence were generally more extreme in men than in women. Māori men had 0.83 (0.80–0.86) times the cancer prevalence of European/Other men, whereas women were similar at 1.00 (0.97–1.02) times the cancer prevalence. The most deprived quintile of men had 0.82 (0.80–0.84) times the cancer prevalence of the least deprived, while for women this figure was 0.95 (0.93–0.98) times.

Table 2 compares age-standardised cancer prevalence diagnosed in adults ≥ 15 years old in the last 18.5 years with shorter-term limited duration prevalence, eg, cancers that were diagnosed in the last year, or the last two to five years. The age-standardised prevalence of cancer diagnosed in any of the 18.5 years was 2.88% of men and 2.82% of women. For diagnoses in the last year these figures were 0.37% in men and 0.34% of women; and for diagnoses in the last five years prevalence was 1.42% of men and 1.29% of women (adding one-year and two-to-five-year prevalences together).

The most prevalent recorded cancers (diagnosed in the past 18.5 years and age-standardised) in men ≥ 15 years, were prostate cancer (1.14%), melanoma (0.51%) and colorectal cancer (0.40%). In women the most prevalent cancers were breast cancer (1.15%), melanoma (0.51%) and colorectal cancer (0.33%).

Table 1: The total population and the 18.5 year prevalence of ICD-10 C code (C00–96) cancer diagnoses made in New Zealand as at 30 June 2013, by demographic group.

	Total population n (%)	Cancer cases n (%)	Age-standardised prevalence %		Age-standardised rate ratio (95% CI)	
			Male	Female	Male	Female
All (crude prevalence = 3.17%)	4,438,914 (100)	140,616 (100)	2.15	2.10	-	-
≥15 years	3,546,288 (100)	139,791 (100)	2.88	2.82	-	-
≥25 years	2,915,025 (100)	138,540 (100)	3.66	3.58	-	-
5-year age groups at 30/6/2013						
0–14	892,626 (20.11)	825 (0.59)	0.10	0.09	-	-
15–24	631,263 (14.22)	1,251 (0.89)	0.20	0.19	-	-
25–34	559,713 (12.61)	2,367 (1.68)	0.39	0.46	-	-
35–44	590,163 (13.30)	6,702 (4.77)	0.84	1.42	-	-
45–54	624,945 (14.08)	17,013 (12.10)	1.85	3.55	-	-
55–64	503,949 (11.35)	29,469 (20.96)	5.28	6.39	-	-
65–74	363,762 (8.19)	40,182 (28.58)	12.62	9.59	-	-
75–84	196,809 (4.43)	30,276 (21.53)	19.87	11.74	-	-
85+	75,684 (1.71)	12,531 (8.91)	23.09	13.04	-	-
Ethnicity						
Māori	553,494 (12.47)	9,666 (6.87)	2.02	2.33	0.83 (0.80–0.86)	1.00 (0.97–1.02)
Pacific	270,966 (6.10)	3,003 (2.14)	1.37	1.75	0.56 (0.53–0.60)	0.75 (0.72–0.78)
Asian	430,056 (9.69)	3,975 (2.83)	0.92	1.19	0.38 (0.36–0.40)	0.51 (0.49–0.53)
European/Other (non-Māori, non-Pacific, non-Asian)	2,562,774 (57.73)	113,406 (80.65)	2.44	2.34	1.00	1.00
Missing	683,037 (15.39)	10,896 (7.75)	-	-	-	-
Partnership status, for those ≥25 years only						
Partnered	1,700,343 (58.33)	84,453 (60.96)	3.95	3.77	1.00	1.00
Non-partnered	725,271 (24.88)	39,690 (28.65)	3.64	3.91	0.92 (0.91–0.93)	1.04 (1.03–1.05)
Missing	489,423 (16.79)	14,397 (10.39)	-	-	-	-
Standardised highest qualification, for those ≥25 years only						
No qualification	497,523 (17.07)	35,262 (25.45)	3.77	3.75	1.00	1.00
School qualification	1,063,761 (36.49)	51,009 (36.82)	3.94	3.84	1.05 (1.03–1.06)	1.02 (1.01–1.04)
Post-school qualification	767,685 (26.34)	29,949 (21.62)	3.98	3.86	1.06 (1.04–1.07)	1.03 (1.02–1.04)
Missing	586,068 (20.10)	22,323 (16.11)	-	-	-	-
Urban/rural residency (usual residence)						
Urban area	3,247,062 (73.15)	112,353 (79.90)	2.27	2.23	1.00	1.00
Rural area	514,419 (11.59)	18,252 (12.98)	2.29	2.27	1.01 (0.99–1.03)	1.02 (1.00–1.04)
Missing	677,448 (15.26)	10,017 (7.12)	-	-	-	-
Smoking status, for those ≥15 years only						
Regular smoker	423,801 (11.95)	10,170 (7.27)	2.38	2.76	0.78 (0.76–0.80)	0.95 (0.93–0.97)
Ex-smoker	661,410 (18.65)	43,827 (31.35)	3.24	3.33	1.07 (1.05–1.08)	1.15 (1.13–1.17)
Never smoked regularly	1,771,212 (49.95)	69,657 (49.83)	3.04	2.89	1.00	1.00
Missing	689,880 (19.45)	16,143 (11.55)	-	-	-	-
Equivalised household income, tertile						
Lowest income	1,041,402 (23.46)	34,860 (24.79)	2.19	2.22	0.91 (0.89–0.93)	0.98 (0.96–1.00)
Middle income	1,075,152 (24.22)	37,320 (26.54)	2.26	2.24	0.94 (0.92–0.96)	0.99 (0.97–1.01)
Highest income	1,073,895 (24.19)	38,934 (27.69)	2.41	2.27	1.00	1.00
Missing	1,248,477 (28.13)	29,511 (20.99)	-	-	-	-

Table 1: The total population and the 18.5 year prevalence of ICD-10 C code (C00–96) cancer diagnoses made in New Zealand as at 30 June 2013, by demographic group (continued).

NZDep2013, quintile						
1 (least deprived)	762,231 (17.17)	28,614 (20.35)	2.45	2.28	1.00	1.00
2	742,416 (16.73)	27,027 (19.22)	2.34	2.26	0.95 (0.93–0.98)	0.99 (0.97–1.02)
3	719,949 (16.22)	25,263 (17.97)	2.26	2.22	0.92 (0.90–0.94)	0.97 (0.95–1.00)
4	700,866 (15.79)	23,958 (17.04)	2.19	2.20	0.89 (0.87–0.92)	0.97 (0.94–0.99)
5 (most deprived)	674,118 (15.19)	18,567 (13.20)	2.00	2.17	0.82 (0.80–0.84)	0.95 (0.93–0.98)
Missing	839,349 (18.91)	17,193 (12.23)	-	-	-	-
District Health Board						
Northland	121,200 (2.73)	5,208 (3.70)	2.32	2.27	1.00 (0.96–1.05)	1.01 (0.96–1.06)
Waitemata	452,913 (10.20)	15,003 (10.67)	2.32	2.25	1.00	1.00
Auckland	360,969 (8.13)	10,068 (7.16)	2.20	2.19	0.95 (0.92–0.98)	0.97 (0.94–1.01)
Counties Manukau	389,757 (8.78)	10,197 (7.25)	2.04	2.09	0.88 (0.85–0.91)	0.93 (0.90–0.97)
Waikato	304,683 (6.86)	10,728 (7.63)	2.30	2.22	0.99 (0.96–1.03)	0.99 (0.95–1.02)
Lakes	82,893 (1.87)	3,015 (2.14)	2.32	2.21	1.00 (0.95–1.06)	0.98 (0.93–1.04)
Bay of Plenty	173,085 (3.90)	7,452 (5.30)	2.41	2.39	1.04 (1.00–1.08)	1.06 (1.02–1.11)
Tairāwhiti	35,550 (0.80)	1,224 (0.87)	2.28	2.19	0.98 (0.91–1.07)	0.98 (0.90–1.06)
Taranaki	96,486 (2.17)	3,735 (2.66)	2.26	2.29	0.97 (0.93–1.03)	1.02 (0.97–1.07)
Hawke's Bay	129,531 (2.92)	5,064 (3.60)	2.23	2.24	0.96 (0.92–1.01)	1.00 (0.95–1.05)
Whanganui	50,517 (1.14)	2,229 (1.59)	2.33	2.29	1.01 (0.94–1.07)	1.02 (0.96–1.09)
Midcentral	141,189 (3.18)	5,265 (3.74)	2.21	2.21	0.95 (0.91–1.00)	0.99 (0.94–1.03)
Hutt Valley	121,440 (2.74)	3,888 (2.76)	2.23	2.05	0.96 (0.92–1.01)	0.91 (0.87–0.96)
Capital & Coast	244,845 (5.52)	7,134 (5.07)	2.13	2.12	0.92 (0.88–0.96)	0.95 (0.91–0.98)
Wairarapa	35,820 (0.81)	1,578 (1.12)	2.19	2.20	0.94 (0.88–1.01)	0.98 (0.91–1.06)
Nelson/Marlborough	118,425 (2.67)	5,088 (3.62)	2.36	2.26	1.02 (0.97–1.07)	1.01 (0.96–1.06)
West Coast	25,863 (0.58)	1,059 (0.75)	2.20	2.27	0.95 (0.87–1.04)	1.01 (0.92–1.11)
Canterbury	416,562 (9.38)	15,690 (11.16)	2.44	2.32	1.05 (1.02–1.09)	1.03 (1.00–1.07)
South Canterbury	49,317 (1.11)	2,223 (1.58)	2.21	2.18	0.95 (0.89–1.02)	0.97 (0.91–1.04)
Southern	258,579 (5.83)	9,126 (6.49)	2.10	2.18	0.91 (0.87–0.94)	0.97 (0.94–1.01)
Missing	829,299 (18.68)	15,648 (11.13)	-	-	-	-
Cancer Network						
Northern	1,324,842 (29.85)	40,473 (28.78)	2.22	2.19	1.00	1.00
Midland	596,208 (13.43)	22,419 (15.94)	2.36	2.27	1.06 (1.04–1.09)	1.04 (1.01–1.06)
Central	819,828 (18.47)	28,896 (20.55)	2.24	2.21	1.01 (0.99–1.03)	1.01 (0.99–1.03)
Southern	868,746 (19.57)	33,186 (23.60)	2.32	2.27	1.04 (1.02–1.06)	1.04 (1.01–1.06)
Missing	829,299 (18.68)	15,648 (11.13)	-	-	-	-

Age-standardised to the WHO World population. Cancer cases are an individual's first primary cancer only. All counts were random rounded to base 3, in accordance with Stats NZ confidentiality requirements. Note some of the comparisons are for different age groups (eg, exclude children) and therefore the overall crude and age standardised prevalences are higher.

Table 2: ICD-10 C code cancer diagnoses, by cancer type and different prevalence times.

	ICD codes	Males			Females		
		Cancer cases n	Crude prevalence percentage	Age-standardised prevalence per 100 (95% CI)	Cancer cases n	Crude prevalence percentage	Age-standardised prevalence per 100 (95% CI)
Adults (≥15 years)							
All cancers (total)	C00-96						
0–≤1 year prev period		8,757	0.506	0.373 (0.365–0.381)	8,307	0.458	0.339 (0.331–0.346)
>1–≤5 years prev period		24,780	1.431	1.054 (1.041–1.067)	23,280	1.283	0.945 (0.933–0.958)
>5–≤10 years prev period		19,917	1.150	0.829 (0.818–0.841)	20,256	1.116	0.793 (0.782–0.805)
>10–≤18.5 years prev period		18,531	1.070	0.756 (0.745–0.767)	21,972	1.211	0.827 (0.815–0.838)
0–18.5 years prev period		68,508	3.957	2.878 (2.858–2.899)	71,364	3.932	2.819 (2.798–2.840)
Prostate	C61						
0–≤1 year prev period		2,862	0.165	0.118 (0.114–0.123)	N/A	N/A	N/A
>1–≤5 years prev period		10,362	0.598	0.422 (0.414–0.430)	N/A	N/A	N/A
>5–≤10 years prev period		8,535	0.493	0.335 (0.328–0.342)	N/A	N/A	N/A
>10–≤18.5 years prev period		7,290	0.421	0.270 (0.264–0.276)	N/A	N/A	N/A
0–18.5 years prev period		29,049	1.678	1.145 (1.132–1.157)	N/A	N/A	N/A
Colorectal	C18-21						
0–≤1 year prev period		1,344	0.078	0.055 (0.052–0.058)	1,194	0.066	0.042 (0.040–0.045)
>1–≤5 years prev period		3,588	0.207	0.146 (0.141–0.150)	3,276	0.181	0.112 (0.108–0.116)
>5–≤10 years prev period		2,742	0.158	0.108 (0.103–0.112)	2,814	0.155	0.091 (0.088–0.095)
>10–≤18.5 years prev period		2,595	0.150	0.099 (0.095–0.103)	2,829	0.156	0.086 (0.082–0.089)
0–18.5 years prev period		10,194	0.589	0.404 (0.396–0.412)	10,059	0.554	0.330 (0.323–0.336)
Breast	C50						
0–≤1 year prev period		24	0.001	0.001 (0.001–0.001)	2,832	0.156	0.121 (0.116–0.126)
>1–≤5 years prev period		90	0.005	0.004 (0.003–0.004)	9,282	0.511	0.385 (0.377–0.393)
>5–≤10 years prev period		48	0.003	0.002 (0.001–0.003)	8,307	0.458	0.327 (0.320–0.334)
>10–≤18.5 years prev period		39	0.002	0.002 (0.001–0.002)	8,568	0.472	0.315 (0.308–0.321)
0–18.5 years prev period		204	0.012	0.008 (0.007–0.009)	28,983	1.597	1.147 (1.134–1.160)
Melanoma	C43						
0–≤1 year prev period		1,137	0.066	0.049 (0.046–0.052)	1,047	0.058	0.044 (0.041–0.046)
>1–≤5 years prev period		3,879	0.224	0.168 (0.163–0.173)	3,615	0.199	0.148 (0.143–0.153)
>5–≤10 years prev period		3,429	0.198	0.147 (0.142–0.152)	3,543	0.195	0.144 (0.139–0.149)
>10–≤18.5 years prev period		3,438	0.199	0.146 (0.141–0.151)	4,284	0.236	0.170 (0.165–0.175)
0–18.5 years prev period		11,874	0.686	0.510 (0.501–0.520)	12,483	0.688	0.505 (0.496–0.515)
Lung	C33-34						
0–≤1 year prev period		582	0.034	0.024 (0.022–0.025)	555	0.031	0.021 (0.019–0.023)
>1–≤5 years prev period		516	0.030	0.021 (0.019–0.023)	666	0.037	0.025 (0.023–0.027)
>5–≤10 years prev period		219	0.013	0.009 (0.008–0.010)	255	0.014	0.009 (0.008–0.011)
>10–≤18.5 years prev period		183	0.011	0.007 (0.006–0.008)	183	0.010	0.006 (0.005–0.007)
0–18.5 years prev period		1,494	0.086	0.060 (0.057–0.063)	1,653	0.091	0.062 (0.058–0.065)
Non-Hodgkin's Lymphoma	C82-85, C96						
0–≤1 year prev period		333	0.019	0.015 (0.013–0.016)	246	0.014	0.010 (0.009–0.011)
>1–≤5 years prev period		1,068	0.062	0.047 (0.044–0.050)	831	0.046	0.033 (0.031–0.035)
>5–≤10 years prev period		816	0.047	0.036 (0.034–0.039)	714	0.039	0.027 (0.025–0.030)
>10–≤18.5 years prev period		576	0.033	0.026 (0.023–0.028)	606	0.033	0.023 (0.021–0.025)
0–18.5 years prev period		2,790	0.161	0.123 (0.119–0.128)	2,394	0.132	0.093 (0.089–0.097)

Table 2: ICD-10 C code cancer diagnoses, by cancer type and different prevalence times (continued).

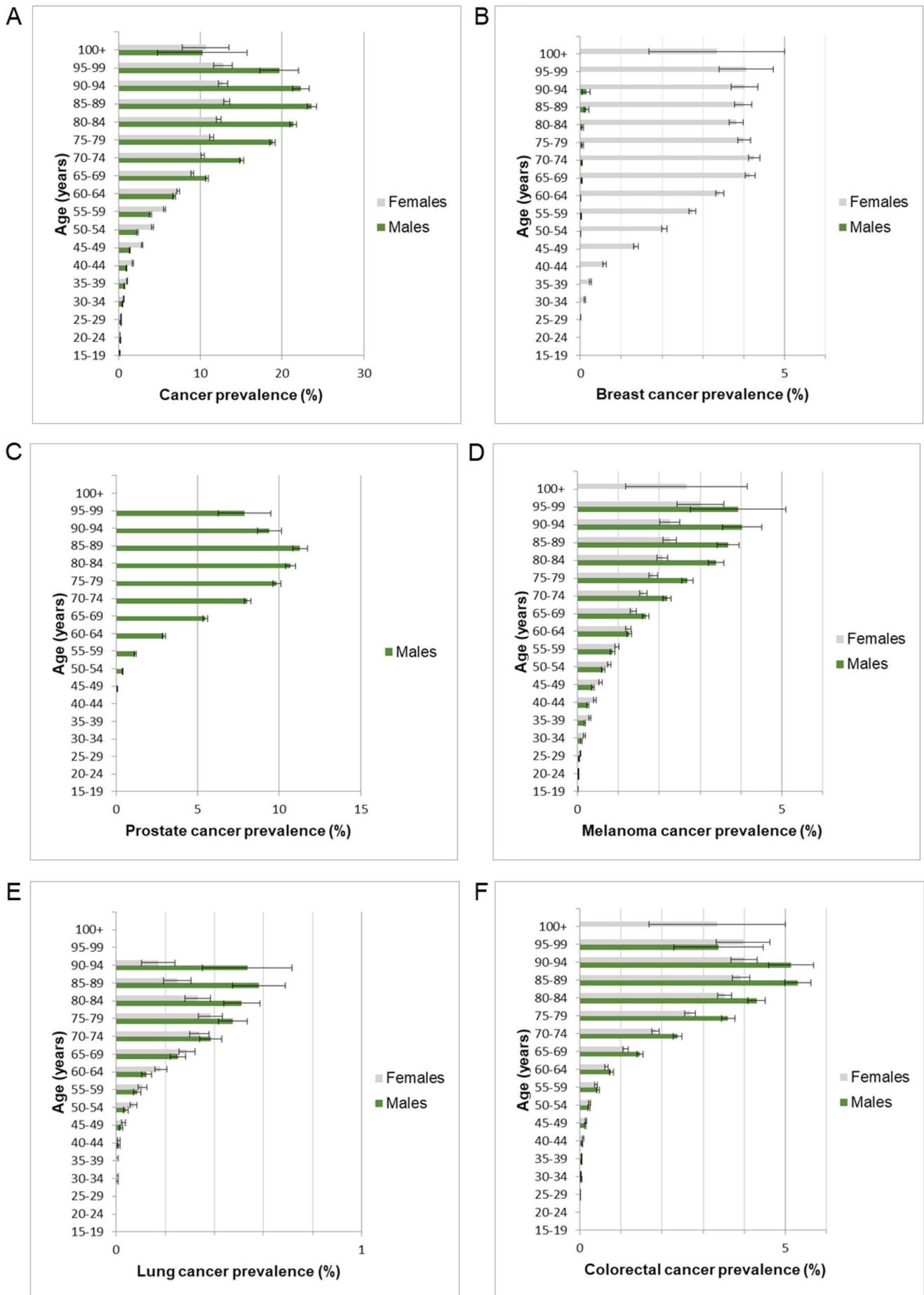
Leukaemia	C91-95						
0–≤1 year prev period		261	0.015	0.012 (0.010–0.013)	165	0.009	0.007 (0.006–0.008)
>1–≤5 years prev period		669	0.039	0.031 (0.028–0.033)	432	0.024	0.018 (0.016–0.020)
>5–≤10 years prev period		621	0.036	0.028 (0.026–0.030)	474	0.026	0.019 (0.018–0.021)
>10–≤18.5 years prev period		669	0.039	0.033 (0.030–0.035)	474	0.026	0.021 (0.019–0.023)
0–18.5 years prev period		2,220	0.128	0.102 (0.098–0.107)	1,545	0.085	0.064 (0.061–0.068)
Endometrial	C54-55						
0–≤1 year prev period		N/A	N/A	N/A	468	0.026	0.020 (0.018–0.022)
>1–≤5 years prev period		N/A	N/A	N/A	1,344	0.074	0.055 (0.052–0.058)
>5–≤10 years prev period		N/A	N/A	N/A	1,098	0.061	0.041 (0.039–0.044)
>10–≤18.5 years prev period		N/A	N/A	N/A	1,101	0.061	0.038 (0.035–0.040)
0–18.5 years prev period		N/A	N/A	N/A	4,008	0.221	0.153 (0.149–0.158)
Renal	C64-66, C68						
0–≤1 year prev period		303	0.017	0.013 (0.012–0.015)	159	0.009	0.006 (0.005–0.007)
>1–≤5 years prev period		843	0.049	0.036 (0.034–0.039)	456	0.025	0.018 (0.016–0.020)
>5–≤10 years prev period		606	0.035	0.026 (0.023–0.028)	369	0.020	0.014 (0.012–0.015)
>10–≤18.5 years prev period		453	0.026	0.019 (0.017–0.020)	342	0.019	0.013 (0.012–0.014)
0–18.5 years prev period		2,205	0.127	0.093 (0.089–0.097)	1,323	0.073	0.051 (0.048–0.053)
Pancreas	C25						
0–≤1 year prev period		99	0.006	0.004 (0.003–0.005)	108	0.006	0.004 (0.003–0.005)
>1–≤5 years prev period		57	0.003	0.003 (0.002–0.003)	66	0.004	0.003 (0.002–0.003)
>5–≤10 years prev period		18	0.001	0.001 (0.001–0.001)	27	0.001	0.001 (0.001–0.001)
>10–≤18.5 years prev period		24	0.001	0.001 (0.001–0.001)	27	0.001	0.001 (0.001–0.001)
0–18.5 years prev period		201	0.012	0.009 (0.007–0.010)	225	0.012	0.009 (0.008–0.010)
Childhood (ages 0–14 years)	C00-96						
0–≤1 year prev period		69	0.015	0.015 (0.012–0.019)	54	0.012	0.012 (0.009–0.015)
>1–≤5 years prev period		207	0.045	0.046 (0.040–0.052)	147	0.034	0.034 (0.028–0.039)
>5–≤10 years prev period		114	0.025	0.025 (0.020–0.029)	123	0.028	0.029 (0.024–0.034)
>10–≤18.5 years prev period		57	0.012	0.012 (0.009–0.016)	57	0.013	0.013 (0.010–0.017)
0–18.5 years prev period		447	0.098	0.098 (0.089–0.107)	381	0.087	0.088 (0.079–0.096)

Age-standardised to the WHO World population. Includes multiple primary cancers for individuals (not just their first primary cancer). The smaller year breakdowns of cancer prevalence do not equal the 0–18.5 year breakdown because individuals can appear more than once in the smaller groups but only once in the bigger group. That is, a person may have been diagnosed with cancer 1 year ago and 5 years ago, but in the 0–18.5 year count they’ll only appear once (so the number of cancers is shown in the smaller breakdowns but the number of people in larger breakdown). The total population of adult males was 1,731,510 and the total population of male children was 456,972. The total population of adult females was 1,814,790 and the total population of female children was 435,654. All counts were random rounded to base 3, in accordance with Stats NZ confidentiality requirements. Prev period: prevalence period.

Figure 1 shows age-specific prevalences for all cancer and some of the most prevalent cancers. Cancer prevalence peaks in the 85–89 year old age group at 23.65% in men and 13.22% in women. Most cancers

show this pattern; however, breast cancer prevalence is similar in women from 65–99 years old at around 4%. Men have higher age-specific prevalences of colorectal, melanoma and lung cancer than women.

Figure 1: Cancer prevalence in adults by age and sex.



A: Total cancer prevalence; B: Breast cancer prevalence; C: Prostate cancer prevalence; D: Melanoma prevalence; E: Lung cancer prevalence; F: Colorectal cancer prevalence. Each figure includes adults (aged ≥ 15 years) of all ethnic groups (including missing) for 18.5-years prevalence. The figures include more than one primary cancer for each person if they were diagnosed with more than one cancer type.

Discussion

We estimated that 3.2% of New Zealand residents (still alive at 30 June 2013) in 2013 had been diagnosed with cancer in New Zealand since 1995, and that crude prevalence was 15.7% in ≥ 75 year olds (42,807/272,493*100). Cancer prevalence was greatest in European/Other people, those with higher SES, and ex-smokers. Prevalence differed slightly by DHB after adjustment for age.

Our results were higher than those of Sharp et al¹⁵ who estimated 17-year limited-duration cancer prevalence in Ireland. After excluding non-melanoma skin cancer they estimated that there were 112,610 people diagnosed with cancer in 1994–2011 and still alive at 31 December 2011¹⁵—a crude prevalence of 2.5% (using the Central Statistics Office population of Ireland in April 2011 of 4,588,252¹⁶), which is lower than our crude prevalence of 3.2%.

Maxwell et al¹⁷ estimated complete (using a 29-year look-back period), as well as limited-duration (for up to 1, 1–5, and 5–10 years) cancer prevalence for Western Australia in 2011. They estimated complete prevalence of 3.8% in males and 3.7% in females of all ages.¹⁷ Like Sharp et al,¹⁵ Maxwell et al¹⁷ did not age-standardise their results, and their results do not include non-melanoma skin cancer.¹⁷ Our findings (crude prevalence of 3.2%) are slightly less than those from Western Australia possibly due to our (shorter) 18.5 year look-back period.

The Globocan 2018^{6,7} 5-year prevalence estimates excluding non-melanoma skin cancer were slightly lower than our findings, which could be a result of their different methods. The self-reported lifetime prevalence proportions (age-standardised to the WHO World Standard Population) reported in the 2002/03 Health Survey⁵ were considerably higher than our findings, which is probably due to the longer look-back period in the Health Survey (lifetime, rather than 18.5 years) and the different method (self-report compared with administrative data, eg, non-melanoma skin cancers may have been recalled).

We found higher cancer prevalence proportions in those with higher SES (compared to lower SES), which is in agreement with our previous work that

has shown that the highest SES groups have higher incidence of cancers with high survival rates, such as prostate, breast and melanoma, (possibly due to higher screening and healthcare access), but that the lowest SES groups are more likely to die from cancer (eg, lung, colorectal, breast, prostate and melanoma), and that socio-economic inequalities in cancer mortality have increased over time.¹⁸ Furthermore, between 1981–2011 the largest increases in cancer incidence were among the highest-income people, widening the cancer incidence inequalities.¹⁸ It is likely that cancer is diagnosed more frequently in people with greater access to healthcare (if people do not visit a doctor cancer cannot be diagnosed even if it exists), meaning that cancer prevalence will be (incorrectly) lower in people with less access to healthcare (their extant cancer will only be diagnosed after death). To our knowledge there are no data available on how access to healthcare (eg, ethnicity and SES) affects death certificate-only cancer diagnoses in the NZCR and we are therefore unable to estimate the potential effect on our results. Cumulative relative cancer survival in New Zealand has also been shown to be higher in the least deprived two quintiles than in the most deprived quintile.¹⁹ Thus, more high-income people are diagnosed with cancer compared to low-income people, and more of them survive the cancer, resulting in higher cancer prevalence. We found greater SES differences in men compared to women, which may be due to larger differences in the incidence of prostate cancer (greater than breast cancer) and melanoma by SES in men, and a bigger SES gap in mortality for colorectal cancer in men than women.¹⁸

We found lower cancer prevalence proportions in Asian and Pacific people compared to European/Other people. Asian people in New Zealand have a lower incidence rate of cancer and a lower mortality rate compared to European/Other people,²⁰ reflected in our findings. We were only able to include people diagnosed with cancer in New Zealand and will therefore have missed people who were diagnosed overseas. This is likely to have a relatively larger impact on Asian and Pacific people who are more likely to be recent migrants, and we have therefore likely underestimated true prevalence for these groups. DHBs with high

levels of inward migration are also likely to have an underestimated cancer prevalence, eg, Counties Manukau DHB. In contrast, we found higher prevalence proportions in Māori compared to European/Other people. Māori have higher cancer incidence and mortality rates than European/Other people,²⁰ with the higher incidence reflected in our results. The linkage rates of health data in the IDI for Pacific and Asian people and older Māori are lower than those for European/Other people (Stats NZ, personal communication). Cancer incidence (and therefore prevalence) could be undercounted if there are more missed links in these groups. Further research into linkage error and missing data in this dataset is warranted to explore these differences.

The strengths of our study include the use of the Stats NZ IDI, including high-quality cancer registry data, providing a large national dataset to examine cancer prevalence and its association with demographic and socioeconomic variables from the 2013 census (although there were a large number (approximately 25%) of records that were not linked to the census). We have age-standardised our results, allowing for comparison with other studies.

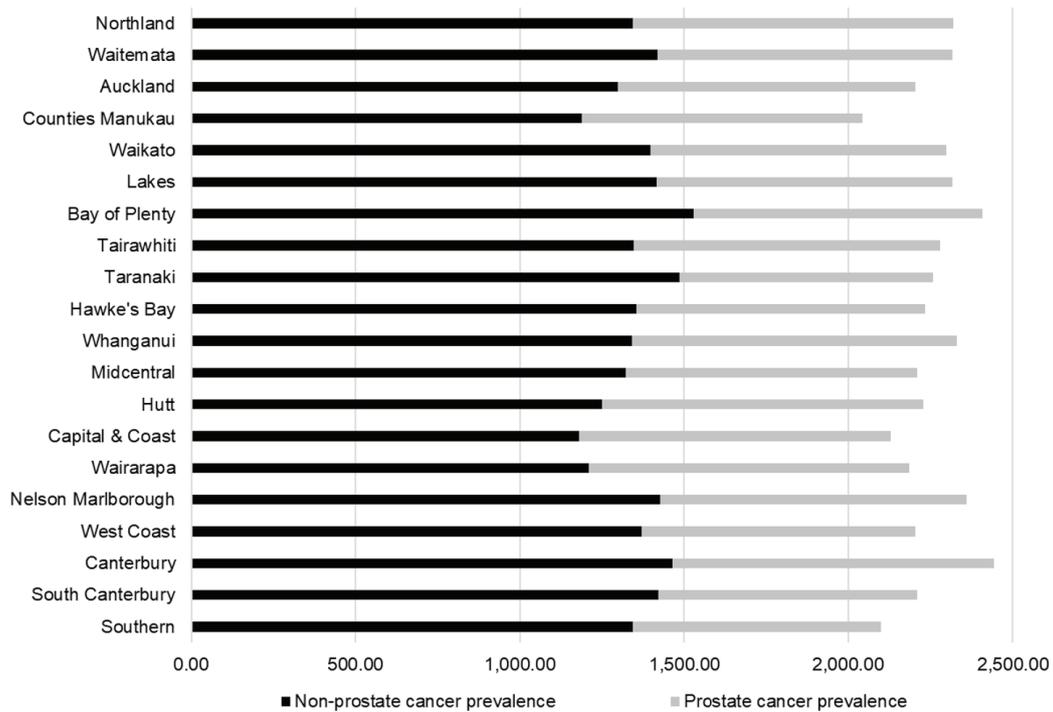
The limitations of our study include that our data were limited in duration (starting in 1995) and we will therefore have missed some prevalent cases; that is, our prevalence estimate falls short of the true prevalence because it excludes long-term survivors (more than 18.5 years). The number of cases that we missed will be larger for the cancers that have an average longer survival time, such as childhood cancers. The authors of an Italian study estimated that their limited duration analysis (15 years) missed 64% of the prevalent cases (that is, 64% of the childhood cancer survivors were diagnosed before the start of the observation period).²¹ Furthermore, since we calculated age at 30 June 2013 people who were diagnosed with a cancer when they were a child will be recorded as being older in our analyses

if the cancer diagnosis was not near in time to 30 June 2013 (a person diagnosed at age five in 1995 will have been 23 in 2013). There is evidence that cancer survivors are at high-risk for late effects months or years after treatment has ceased, and this may be particularly important for survivors of cancer diagnosed as a child or young adult since they (potentially) have many years of life left.^{22,23}

Finally, there is some evidence for differences in linkage rates by sex, age, ethnicity, and age-by-ethnicity.^{24,25} For this reason we may have underestimated cancer prevalence and the differences by sex, age and ethnicity, particularly in the oldest age groups. Some 8% (12,837) of cancers registered prior to 30 June 2013 (with no record of death before this date) were not linked to our study population, perhaps because people emigrated, they were only visiting New Zealand or their health data were mistakenly not linked to the spine. The latter would only slightly underestimate the prevalence reported here (the truer crude prevalence could not be more than 3.4%).

In summary, cancer prevalence is affected by cancer incidence and survival rates. Prevalence provides an outline for policy makers and health planners about the extent of the burden of cancer in different populations, whether historically or within the last one, five or 10 years. Although the extent of need for healthcare will differ among cancer survivors depending on factors such as cancer type^{22,26} and comorbidities,²⁷ nevertheless estimates of cancer prevalence can help decision-makers, planners and service providers to develop follow-up and after-care services to more effectively meet survivors' needs as they change over time (a person diagnosed one year ago is likely to have quite different needs to someone diagnosed 10 years ago), and to understand the differences that they might expect to find in prevalence by age, sex, ethnicity, socioeconomic status and smoking status of cancer survivors.

Appendix Figure 2: Non-prostate cancer prevalence & prostate cancer prevalence by District Health Board.



Appendix Table 2: Prevalence of ICD-10 D code cancer diagnoses.

	Cancer cases n (%)	Total population n	Crude prevalence %
Age (years) in 2013, mean (SD)	69 (15.0)		
Sex			
Male	756 (54.4)	2,188,482	0.03
Female	633 (45.6)	2,250,444	0.03
5-year age groups at 30/6/2013			
0–14	9 (0.7)	892,626	0.00
15–24	6 (0.4)	631,263	0.00
25–34	21 (1.5)	559,713	0.00
35–44	57 (4.1)	590,166	0.01
45–54	138 (10.0)	624,945	0.02
55–64	237 (17.2)	503,949	0.05
65–74	351 (25.4)	363,765	0.10
75–84	387 (28.0)	196,806	0.20
85+	174 (12.6)	75,687	0.23
Ethnicity			
Māori	129 (9.3)	553,494	0.02
Pacific	57 (4.1)	270,966	0.02
Asian	45 (3.2)	430,056	0.01
European/Other	1,041 (75.0)	2,562,774	0.04
Missing	117 (8.4)	683,037	0.02

The D code diagnoses included in this table were D45, D460, D461, D462, D464, D465, D466, D467, D469, D471, D473, D474, and D475, these are neoplasms of uncertain or unknown behaviour. Because of the

random rounding the different groups do not add to the same total. Only the first diagnosis for an individual is included (not multiple primary cancers).

Competing interests:

Dr Brewer has worked through a university on an unrelated project funded by Janssen-Cilag Pty Limited. Dr Brewer reports grants from New Zealand Government, through Ministry of Business, Innovation & Employment during the conduct of the study; grants from Janssen-Cilag Pty Limited outside the submitted work. Dr Teng reports grants from Ministry for Business Innovation and Employment during the conduct of the study.

Acknowledgements:

The results in this paper are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI), managed by Statistics New Zealand. The opinions, findings, recommendations, and conclusions expressed in this paper are those of the author(s), not Statistics NZ. Access to the anonymised data used in this study was provided by Statistics NZ under the security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this paper have been confidentialised to protect these groups from identification and to keep their data safe. Careful consideration has been given to the privacy, security, and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz. This study was funded through the Healthier Lives National Science Challenge by the Ministry of Business, Innovation and Employment (UOOX1513).

Author information:

Naomi Brewer, Research Officer, Department of Public Health, University Of Otago, Wellington, and Centre for Public Health Research, Massey University, Wellington; Andrea Teng, Senior Research Fellow, Department of Public Health, University Of Otago, Wellington; June Atkinson, Senior Analyst, Department of Public Health, University Of Otago, Wellington; Parry Guilford, Research Professor, Department of Biochemistry, University Of Otago, Dunedin; Cristin Print, Professor, Faculty of Medical and Health Sciences, The University of Auckland, Auckland; Tony Blakely, Research Professor, Department of Public Health, University Of Otago, Wellington, and Melbourne School of Population and Global Health, The University of Melbourne, Australia.

Corresponding author:

Dr Andrea Teng, Department of Public Health, University Of Otago, PO Box 7343, Newtown, Wellington 6242.
andrea.teng@otago.ac.nz

URL:

www.nzma.org.nz/journal-articles/an-estimate-of-limited-duration-cancer-prevalence-in-new-zealand-using-big-data

REFERENCES:

1. Ministry of Health. Cancer: New registrations and deaths 2013. Wellington, New Zealand: Ministry of Health, 2016.
2. Disney G, McDonald A, Atkinson J, Blakely T. New Zealand Census Mortality and CancerTrends Study Data Explorer. Available at: <http://www.uow.otago.ac.nz/NZCMS-CT-dataexplorer>. Accessed 27/5/18.
3. Colonna M, Mitton N, Bossard N, et al. Total and partial cancer prevalence in the adult French population in 2008. *BMC Cancer*. 2015; 15 (153).
4. Costilla R, Tobias M, Blakely T. The burden of cancer in New Zealand: a comparison of incidence and DALY metrics and its relevance

- for ethnic disparities. *Aust N Z J Public Health*. 2013; 37:218–25.
5. Ministry of Health. *Urban–Rural Health Comparisons: Key results of the 2002/03 New Zealand Health Survey*. Wellington, New Zealand: Ministry of Health, 2007.
 6. Ferlay J, Ervik M, Lam F, et al. *Global Cancer Observatory: Cancer Today*. Available at: <http://gco.iarc.fr/today>. Accessed 11 March 2019.
 7. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. 2019; 144:1941–53.
 8. Statistics New Zealand. *Linking methodology used by Statistics New Zealand in the Integrated Data Infrastructure project*. Available from www.stats.govt.nz. Accessed 23/4/2018.
 9. Ministry of Health. *Cancer in New Zealand: Trends and projections*. Wellington, New Zealand: Ministry of Health, 2002.
 10. Ministry of Health. *Cancer: New registrations and deaths 2005*. Wellington, New Zealand: Ministry of Health, 2008.
 11. Statistics New Zealand. *Experimental population estimates from linked administrative data: methods and results*. Retrieved from www.stats.govt.nz, 2016.
 12. Atkinson J, Salmond C, Crampton P. *NZDep2013 Index of Deprivation*. Available at: <http://www.otago.ac.nz/wellington/otago069936.pdf>. Wellington, New Zealand: University of Otago, 2014.
 13. Ministry of Health. *Presenting ethnicity: comparing prioritised and total response ethnicity in descriptive analyses of New Zealand Health Monitor surveys*. Wellington, New Zealand: Ministry of Health, 2008.
 14. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2008.
 15. Sharp L, Deady S, Gallagher P, et al. The magnitude and characteristics of the population of cancer survivors: using population-based estimates of cancer prevalence to inform service planning for survivorship care. *BMC Cancer*. 2014; 14(767).
 16. Central Statistics Office. *This is Ireland - Highlights from Census 2011, Part 1*. Dublin, Ireland: The Stationery Office, 2012.
 17. Maxwell S, O’Leary P, Slevin T, Moorin R. The increase in cancer prevalence and hospital burden in Western Australia, 1992–2011. *Popul Health Metr*. 2014; 12(33).
 18. Teng A, Atkinson J, Disney G, et al. Changing socioeconomic inequalities in cancer incidence and mortality: Cohort study with 54 million person-years follow-up 1981–2011. *Int J Cancer*. 2017; 140:1306–16.
 19. Ministry of Health. *Cancer patient survival: 1994 to 2011*. Wellington, New Zealand: Ministry of Health, 2015.
 20. Disney G, McDonald A, Atkinson J, Blakely T. *New Zealand Census Mortality and Cancer Trends Study Data Explorer*. Available at: <http://www.uow.otago.ac.nz/> NZCMS-CT-dataexplorer. Accessed 20/12/2019.
 21. Francisci S, Guzzinati S, Dal Maso L, et al. An estimate of the number of people in Italy living after a childhood cancer. *Int J Cancer*. 2017; 140:2444–50.
 22. Macmillan Cancer Support. *Cancer’s unequal burden. The reality behind improving cancer survival rates*. Available at: http://www.macmillan.org.uk/Documents/CancersUnequalBurden_2014.pdf. Accessed 22/03/2016.
 23. Rugbjerg K, Olsen J. Long-term Risk of Hospitalization for Somatic Diseases in Survivors of Adolescent or Young Adult Cancer. *JAMA Oncol*. 2016; 2:193–200.
 24. Black A. *The IDI prototype spine’s creation and coverage*. In: *Statistics New Zealand*, (ed) Working Paper No 16-03, 2016. Available from www.stats.govt.nz.
 25. Milne B, Atkinson J, Blakely T, et al. *Data resource profile: The New Zealand Integrated Data Infrastructure (IDI)*. *Int J Epidemiol*. 2019; 48:1027.
 26. Siegel R, DeSantis C, Virgo K, et al. *Cancer Treatment and Survivorship Statistics, 2012*. *CA Cancer J Clin*. 2012; 62:220–41.
 27. Sarfati D, Gurney J, Lim B, et al. Identifying important comorbidity among cancer populations using administrative data: Prevalence and impact on survival. *Asia Pac J Clin Oncol*. 2016; 12:e47–e56.
 28. Gibb S, Bycroft C, Matheson-Dunning N. Identifying the New Zealand resident population in the Integrated Data Infrastructure (IDI). Retrieved from www.stats.govt.nz, 2016.