



## Agreement between self-reported and central cancer registry-recorded prevalence of cancer in the Alaska EARTH study

Sarah H. Nash<sup>a</sup>, Gretchen Day <sup>b</sup>, Vanessa Y. Hiratsuka<sup>c</sup>, Garrett L. Zimpelman<sup>a</sup> and Kathryn R. Koller <sup>b</sup>

<sup>a</sup>Alaska Native Tumor Registry, Alaska Native Epidemiology Center, Community Health Services, Alaska Native Tribal Health Consortium, Anchorage, AK, USA; <sup>b</sup>Clinical and Research Services, Community Health Services, Alaska Native Tribal Health Consortium, Anchorage, AK, USA; <sup>c</sup>Research Department, Southcentral Foundation, Anchorage, AK, USA

### ABSTRACT

Reliance on self-reported health status information as a measure of population health can be challenging due to errors associated with participant recall. We sought to determine agreement between self-reported and registry-recorded site-specific cancer diagnoses in a cohort of Alaska Native people. We linked cancer history information from the Alaska Education and Research Towards Health (EARTH) cohort and the Alaska Native Tumor Registry (ANTR), and calculated validity measures (sensitivity, specificity, positive predictive value, negative predictive value, kappa). Multiple logistic regression models were used to assess independent associations of demographic variables with incorrect reporting. We found that among Alaska EARTH participants, 140 self-reported a history of cancer, and 99 matched the ANTR. Sensitivity ranged from 79% (colorectal cancer) to 100% (prostate cancer); specificity was over 98% for all-sites examined. Kappa was higher among prostate and female breast cancers ( $\kappa=0.86$ ) than colorectal cancers ( $\kappa=0.63$ ). Women (odds ratio [OR] (95% confidence interval [CI]): 2.8 (1.49–5.31)) and participants who were older than 50 years (OR (95% CI): 2.8 (1.53–4.12)) were more likely to report incorrectly. These data showed good agreement between self-reported and registry-recorded cancer history. This may be attributed to the high quality of care within the Alaska Tribal Health System, which strongly values patient-provider relationships and the provision of culturally appropriate care.

### ARTICLE HISTORY

Received 9 August 2018  
Revised 14 December 2018  
Accepted 10 January 2019

### KEYWORDS

Native American; cohort study; Alaska Native cancer; self-report; tumour registry; validity; health literacy

### Introduction

Cancer-specific health literacy may affect how and what cancer information a person accesses, and their adherence to cancer prevention guidelines [1–3], including screening [4–6]. Among those with a cancer diagnosis, health literacy may affect understanding of the diagnosis and associated treatment [2], as well as lead to poorer outcomes, including reduced quality of life and survival [7,8]. Comparison of self-reported health conditions with those indicated in the medical record is one way to understand how patients perceive their medical diagnoses and assess the validity of self-reported information [9–12]. Reliability of self-reported health conditions is influenced by factors including age, sex, education and health condition of interest [10,13–17]. While extensive research has been conducted to investigate and improve other self-reported measures in health research, such as self-reported diet [18], smoking [19] and physical activity [20] histories, the validity of measures of self-reported health conditions has been relatively underexplored.

While several studies have compared self-reported health history to the medical record [13,21–23], for certain diseases, such as cancer, a central registry provides another source of data for comparison to self-report. These data have been collected from multiple healthcare facilities, and compiled and curated by trained professionals [16,17,24,25]; thus, data are complete and of high quality. In this study, we were interested specifically in accuracy of self-reported cancer history among Alaska Native (AN) people. We linked data from a population-based central cancer registry, the Alaska Native Tumor Registry (ANTR), with self-reported medical history information from the Alaska Education and Research Towards Health (EARTH) cohort study. Our objective was to assess agreement with, and describe the validity of, the self-report measures. A recent study of the Alaska EARTH cohort suggested that cancer was among the most accurately self-reported health conditions among AN people, but found differences by age, sex and education level [10]. Here, we expand the previous study to examine

**CONTACT** Sarah H. Nash  [shnash@anthc.org](mailto:shnash@anthc.org)  Alaska Native Tumor Registry, Alaska Native Epidemiology Center, Community Health Services, Alaska Native Tribal Health Consortium, 3900 Ambassador Drive, Anchorage, AK, USA  
 Supplementary material for this article can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

accuracy by site, as well as by the demographic variables rural/urban residence, and primary language spoken at home. We anticipate that the findings of this study will be of importance to those interested in health literacy among AN and American Indian (AI) peoples, as well clinicians that provide cancer care within the tribal health system.

## Methods

### AN people

Approximately 144,274 AI/AN people reside in Alaska [26] (individuals reporting AI/AN identity alone or in combination with another racial identity), comprising 19.5% of the Alaskan population and representing 229 federally recognised tribes. Almost 90% of AI/AN people living in Alaska identify as AN [27]; therefore, hereafter we will refer to all AI/AN people resident in Alaska as “Alaska Native (AN) people”. Healthcare for AN people residing in Alaska is provided by 32 regional tribal health organisations, as well as the Alaska Native Tribal Health Consortium, which provides statewide services. There is one tribally managed tertiary healthcare facility in the state, located in Anchorage, which provides the majority of cancer diagnosis and treatment services to AN people.

### The Alaska EARTH study

The Alaska EARTH study was part of a multisite cohort study conducted to understand chronic disease risk, including cancer, among AI/AN communities; full details of study purpose and procedures are given elsewhere [28]. Briefly, for the Alaska EARTH study, participants from three Alaska regions were recruited during 2004–2006 ( $n=3,821$ ). Residents of 25 rural communities and one urban centre who self-identified as AI/AN were invited to participate. Participants completed written informed consent, demographic, lifestyle and medical history questionnaires, and anthropometric data were collected. The EARTH study was approved by the Alaska Area Institutional Review Board [protocol number 2000-03-008]. Tribal approval for this analysis was granted by the Alaska Native Tribal Health Consortium, Southcentral Foundation, Southeast Alaska Regional Health Consortium and the Yukon-Kuskokwim Health Corporation, and each of these tribal health organisations also reviewed and approved this manuscript.

### Study questionnaires

Demographic data collected from participants included age, sex, education level, address and primary language

spoken at home. Self-reported cancer information was obtained using questions prefaced by “Did a doctor or healthcare provider ever tell you that you had cancer?” Response choices were yes, no, skip, refuse or don’t know. If participants answered “yes”, then they were asked to give information on the site of diagnosis. Urban residence was defined by residence in the one Alaska urban centre; all other locations were considered rural. Primary language spoken at home was dichotomised as “English” versus either “Native language” or “Both English and Native language”; the latter two were combined into one group for analysis. Participants self-reported their highest level of education completed; this was dichotomised as high school or less ( $\leq 12$  years) or greater than high school ( $>12$  years).

### Tumour registry

Cancer diagnoses were recorded by the ANTR, a population-based registry that records information on cancers diagnosed among AI/AN people since 1969, who meet eligibility requirements for Indian Health Service benefits and who are Alaskan residents at the time of cancer diagnosis. Cancer information is used to understand the unique burden of cancer among AN people and to support research to understand cancer in this population. The ANTR has been collecting cancer information according to National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) Program standards since its inception, and has been a full member of the SEER Program since 1999. As part of the ANTR’s standard surveillance process, cases were identified in multiple ways, including tumour registry and pathology files of the Alaska Native Medical Center and other Native and non-Native healthcare facilities throughout the state; linkage to the Alaska State Cancer Registry and the Washington State Cancer Registry; and death certificates ( $<1\%$  cases). Only cancers diagnosed prior to EARTH study enrolment were included (i.e. cancers that were prevalent at time of consent). Classification of cancer site of origin was completed according to the *International Classification of Diseases for Oncology*, second and third editions [29]. Linkage of Alaska EARTH participants to the ANTR database was performed based on combinations of key identifying information including first and last names, date of birth, sex and social security number.

### Statistical analysis

Of the 3,821 participants recruited into the Alaska EARTH cohort, 3,747 were included in this study. Participants were excluded for missing information necessary to link with the

registry (n=6) for incomplete questionnaires (n=3) or with missing information on self-reported cancer history (n=65). Descriptive statistics (mean, standard deviation, frequencies) are given for demographic characteristics of EARTH study participants. Comparisons between groups were made using one-way ANOVA for continuous variables (age), using Tukey's *post hoc* tests to determine which groups were responsible for significant differences, and chi-squared for categorical variables (sex, educational attainment, rural/urban residence and primary language spoken at home). Agreement between self-reported and registry-recorded cancer diagnoses was calculated using sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). In addition, kappa values were calculated to differentiate between true agreement, and agreement that may be expected due to chance. We present this comprehensive set of agreement measures to foster comparison with previous studies, including in the present study population [10]. Furthermore, while results are presented using the ANTR (registry-recorded) diagnoses as the reference, as described in Koller et al [10], the tumour registry-recorded PPV and NPV, respectively, reflect sensitivity and specificity if self-report is substituted for the referent. Thus, these analyses enable assessment of agreement independent of the assumption of a gold standard metric. For self-reported diagnoses, a "skip" (n=0), "refuse" (n=1) or "don't know" (n=4) response was coded as "no". Metrics were calculated for cancer (all-sites), as well as those cancer sites with the highest case counts: female breast, colorectal and prostate cancers. Exact methods were used to resolve zero count cells in site-specific stratified analyses.

Multiple logistic regression models were used to estimate independent associations of demographic variables (age, sex, educational attainment, urban vs. rural residence and language spoken at home) with reporting outcome (correct/incorrect reporting). In these models, "correct reporting" was defined as self-reporting that

was verified by the tumour registry. "Incorrect reporting" was defined as either reporting a history of cancer when none was recorded in the registry or reporting no history of cancer when one was recorded in the registry. All statistical tests were two-sided with  $\alpha=0.05$ . All statistical analyses were performed with SAS version 9.4 (SAS Institute, Cary, NC).

## Results

Of the 3,747 participants included in this analysis, 140 (4%) self-reported a history of cancer and 99 (3%) matched to the ANTR as having a registry-recorded history of cancer prior to study enrolment. The mean age of study participants was 40.3 years (SD=15.0), and 61% of the study sample was female. Table 1 provides the demographic characteristics of participants who correctly self-reported a history of cancer (n=86) or no cancer (n=3,594), compared to those who either reported a cancer not recorded in the registry (n=54) or those who failed to report a cancer that was recorded in the registry (n=13). Participants who incorrectly reported a cancer diagnosis (n=67 total) were older ( $F=50.05$ ,  $p<0.0001$ ) and more likely to be female ( $\chi^2=20.8$ ,  $p=0.0001$ ). There was a higher proportion of participants with 12th grade education ( $\chi^2=9.6$ ,  $p=0.0218$ ) and rural residence ( $\chi^2=16.6$ ,  $p=0.0008$ ) among those who had a registry-recorded cancer but did not report it, relative to those who either correctly self-reported their cancer status or who self-reported a history of cancer that was not recorded in the registry. The proportion who spoke English as the primary language was slightly lower among those who had a registry-recorded cancer but did not report it ( $\chi^2=12.6$ ,  $p=0.0056$ ).

To provide more detail on participants who reported incorrectly, we examined independent associations of demographic characteristics with reporting outcome

**Table 1.** Demographic characteristics of Alaska EARTH study participants (n=3680) who correctly self-reported a history of cancer, compared to 67 who incorrectly reported either that they had a cancer not recorded in the tumour registry (n=54) or who failed to report a cancer that was recorded in the registry (n=13)<sup>†</sup>.

	"Correct" reporting			"Incorrect" reporting		p-Value
	Alaska EARTH (all) n=3,747	Alaska EARTH (no cancer and reported no cancer) n=3,594	Alaska EARTH (had cancer and reported it) n=86	Alaska EARTH (Reported cancer, but didn't have it) n=54	Alaska EARTH (had cancer but didn't report it) n=13	
Age (years, mean [SD])*	40.3 (15.0)	39.8 (14.7) <sup>a</sup>	57.0 (13.5) <sup>b</sup>	47.6 (13.3) <sup>c</sup>	58.4 (13.3) <sup>b,c</sup>	<0.0001
Sex (% female)	2,181 (60.9)	2,162 (60.2)	64 (74.4)	43 (79.6)	12 (92.3)	0.0001
Educational attainment (% ≤12 years)	2,193 (59.4)	2,112 (59.6)	49 (57.7)	22 (40.7)	10 (76.9)	0.0218
Residence (% rural)	1,910 (51.0)	1,852 (51.5)	29 (33.7)	20 (37.0)	9 (69.2)	0.0008
Language (% English as the primary language at home)	2,504 (67.1)	2,405 (67.1)	56 (66.7)	40 (74.1)	3 (23.1)	0.0056

\*For continuous variables, statistically significant differences between groups assessed using Tukey's test, and indicated by different letters associated.

<sup>†</sup>Details for the full Alaska EARTH cohort (n=3,821) were previously reported by Redwood and colleagues [39].

**Table 2.** Independent associations of demographic characteristics with incorrect reporting of cancer history<sup>a</sup> (n=3,612<sup>b</sup> who reported their cancer history correctly; n=67 who reported their cancer history incorrectly).

	Odds ratio	95% CI	p-Value
Sex (male vs. female)	2.8	1.49–5.31	0.0014
Age group (<50 years vs. >50 years)	2.5	1.53–4.12	0.0003
Education (<12 vs. >12)	1.4	0.85–2.38	0.1829
Rural (rural vs. urban)	0.8	0.49–1.43	0.5217
Language (English vs. non-English)	1.2	0.69–1.99	0.5698

<sup>a</sup>Outcome = reported cancer status correctly (yes/no).

<sup>b</sup>Sample size varied between 361 and 3,680 due to missing values for covariate data.

(Table 2). Women and participants older than 50 years at study enrolment were more likely to have reported their cancer history incorrectly. Of those who reported a cancer that was not recorded in the ANTR (n=54), 43 (80%) were women. Among cancers women incorrectly reported, 56% were cervical cancers, and a further 21% were other female-specific cancers (e.g. breast (<5%), ovarian (7%), endometrial (9%)). Women also incorrectly reported colorectal (12%) and other (<15%) cancers. Among men who incorrectly reported having cancer, 82% were other or unknown

cancer sites, <10% were prostate cancers and <10% were kidney cancers.

Agreement measures were assessed for cancer (all-sites), as well as among the most common sites observed among Alaska EARTH participants: female breast, colorectal and prostate cancers (Table 3). For all-sites, as well as each common site, self-report was more specific than sensitive. Specificity was more than 98% for all cancer sites, whereas sensitivity ranged from 78.6% (colorectal cancer) to 100.0% (prostate cancer). PPV for self-reported cancers ranged between 52.4% (colorectal cancer) and 84.8% (female breast cancers). NPV was more than 99.6% for all cancer sites examined. Kappa values also varied by cancer site: values were high for female breast and prostate cancers ( $\kappa=0.86$  for both sites), and moderate for colorectal cancer ( $\kappa=0.63$ ).

We examined agreement measures in strata of demographic characteristic (Table 4). For cancer (all-sites), sensitivity was greater among males, those aged 18–50 years at study enrolment, those living in an urban area and those who spoke English as their primary language at home. Neither specificity nor NPV varied substantially by

**Table 3.** Agreement of self-reported and tumour registry for cancer (all-sites), as well as the three leading prevalent cancers among the Alaska EARTH participants (female breast, colorectal and prostate cancers).<sup>a</sup>

	Prevalence		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Kappa
	ANTR	Self-report					
Cancer (all-sites)	99	140	86.9 (80.2–93.5)	98.5 (98.1–98.9)	61.4 (53.4–69.5)	99.6 (99.4–99.8)	0.71 (0.64–0.78)
Breast (F)	34	33	85.3 (68.9–95.1)	99.8 (99.7–99.9)	87.9 (71.8–96.6)	99.8 (99.7–99.9)	0.86 (0.78–0.95)
Colorectal	14	21	78.6 (49.2–95.3)	99.7 (99.5–99.9)	52.4 (29.8–74.3)	99.9 (99.8–99.9)	0.63 (0.44–0.81)
Prostate (M)	7	8	100 (54.1–100)	99.9 (99.8–99.9)	75.0 (34.9–96.8)	100 (99.9–100)	0.86 (0.66–1.0)

<sup>a</sup>For the purposes of these comparisons, ANTR was treated as the reference. However, PPV and NPV, respectively, reflect sensitivity and specificity if self-report is substituted for the reference.

**Table 4.** Agreement between self-reported and tumour registry for cancer (all-sites), among Alaska EARTH participants, stratified by demographic and clinical characteristics.<sup>a</sup>

	Sensitivity		Specificity		PPV		NPV		Kappa	
	(%)	95% CI <sup>a</sup>	(%)	95% CI	(%)	95% CI	(%)	95% CI	(%)	95% CI
<b>Cancer (all-sites)</b>										
<i>Sex</i>										
Male	95.6	78.1–99.9	99.3	98.6–99.6	66.7	48.2–82.0	99.9	99.6–100.0	0.78	0.66–0.90
Female	84.2	74.0–91.6	98.1	97.4–98.6	59.8	49.9–69.2	99.5	99.0–99.7	0.68	0.61–0.77
<i>Age at enrolment (years)</i>										
18–50	90	74.3–98.0	98.8	98.3–99.2	45.9	33.1–59.2	99.9	99.7–99.9	0.6	0.48–0.72
50+	85.3	74.6–92.7	97.7	96.5–98.6	73.4	62.3–82.7	98.9	98.0–99.5	0.77	0.69–0.85
<i>Education (years)</i>										
<12	86.1	70.5–95.3	98.9	98.0–99.7	79.5	63.5–90.7	99.4	98.5–99.8	0.81	0.72–0.91
≥12	89.4	76.9–96.5	98.1	97.3–98.7	54.6	42.8–65.9	99.7	99.1–99.8	0.67	0.57–0.76
<i>Residence</i>										
Rural	76.3	59.8–88.6	98.9	98.4–99.4	59.2	44.2–73.0	99.5	99.1–99.8	0.66	0.54–0.78
Urban	93.4	84.1–98.2	98.1	97.3–98.7	62.6	51.9–72.6	99.8	99.4–99.9	0.74	0.66–0.82
<i>Primary language</i>										
English	94.8	85.9–98.9	98.3	97.8–98.8	58.3	47.8–68.3	99.9	99.6–99.9	0.71	0.63–0.80
Non-English/both	73.7	56.9–86.6	98.8	98.0–99.4	66.7	52.4–80.4	99.2	98.5–99.6	0.69	0.57–0.81
<i>EARTH study region</i>										
Southcentral	93.8	82.8–98.7	97.8	97.0–98.6	61.6	49.5–72.8	99.8	99.3–99.9	0.73	0.64–0.82
Southeast	83.9	66.3–94.6	98.4	97.3–99.1	65	48.3–79.4	99.4	98.6–99.8	0.72	0.60–0.84
Southwest	75	50.9–91.3	99.2	98.6–99.6	55.6	35.3–74.5	99.7	99.2–99.9	0.63	0.47–0.79

<sup>a</sup>For the purposes of these comparisons, ANTR was treated as the reference. However, PPV and NPV, respectively, reflect sensitivity and specificity if self-report is substituted for the reference.



demographic characteristic. In contrast, higher PPV was observed among males, those aged 50+ years at study enrolment, those residing in an urban area and those reporting non-English or both as the primary language(s) spoken at home. The pattern was similar for kappa, where we observed greater values among males, those aged 50+ years at study enrolment and those residing in an urban area.

Finally, we also examined whether site-specific cancer agreement measures varied by demographic characteristic (Supplementary Tables 1 and 2). Due to low case counts within strata, these measures were calculated for female breast and colorectal cancers only. Strata-specific agreement values differed between female breast and colorectal cancers, but for both sites, patterns were similar to those observed for cancer (all-sites). Of note, sensitivity and PPV varied substantially by education level, rural/urban residence and primary language for female breast cancer. For colorectal cancer, sensitivity and PPV also showed variations with age. For female breast cancer, kappa values were substantially higher among women who were 18–50 years at diagnosis and those who spoke English as their primary language at home. For colorectal cancer, kappa values were substantially higher among those who were older than 50 years at time of diagnosis, as well as those with <12 years of education.

## Discussion

Among Alaska EARTH participants, measures of agreement indicated generally high concurrence between self-reported and registry-recorded cancer history. This may reflect the quality of cancer care within the Alaska Tribal Health System (ATHS), which places high value on patient–provider relationships and delivery of culturally appropriate care. Agreement varied by cancer site and demographic characteristic: for all cancer sites, agreement was highest among males, those older than 50 years at study enrolment and those residing in an urban area. Women’s cancers, particularly cervical cancers, were among those more likely to be self-reported by individuals without a registry-recorded history of cancer.

Our results may have important implications for health literacy and cancer communication among AN people. Previous studies have indicated that cancer health literacy varies by age, income and education level [30,31], and that cancer-specific health literacy may affect how a patient accesses cancer information, their adherence to cancer prevention guidelines [1–3] and their understanding of cancer diagnosis and treatment [2]. Our findings suggest there may be several population subgroups, including women and older patients that could benefit from tailored explanations from their providers regarding

cancer diagnoses, or interventions to improve cancer-related health literacy. In particular, care should be taken to ensure that patients fully understand the difference between precancerous (non-malignant) and malignant findings of cancer screening tests.

Our results are in general agreement with the work of Koller and colleagues, who recently demonstrated that cancer was among the best self-reported chronic diseases by AN people [10]. Of note, whereas other chronic conditions, including heart disease and diabetes, were more likely to be underreported, cancer was over-reported in both the Koller et al. study and the present analysis. We observed higher over-reporting for colorectal and cervical cancers, both screenable cancers which we speculate may be more likely than some other sites to be confused with non-malignant diagnoses (e.g. precancerous polyps or cervical dysplasia). It is also possible that precancer may have been interpreted by the patient as early cancer. Since precancerous conditions require treatment to prevent cancer, it is further possible that such treatment could be confused with or interpreted as treatment of cancer. Despite this finding, agreement between self-reported and tumour registry-recorded cancer history was generally high. We observed the highest agreement between self-report and the tumour registry for cancers of the prostate and female breast ( $\kappa=0.86$  for both sites). This concurs with findings from other studies in the U.S. and Europe, where breast cancer in particular has been shown to be well-recalled [11,16,24,32]. In part, this may be attributed to greater public awareness and de-stigmatisation of breast cancer, as well as the association with one’s sense of self identity and body image. Additionally, the ATHS has actively promoted breast cancer screening as a cancer prevention strategy.

Agreement also varied by demographic characteristic, with the highest agreement (cancer, all-sites) being among males, and those aged >50 years at baseline. These findings seem contrary to some previous reports, which have observed higher validity of self-report among women and younger persons [17]. Some of these differences may be methodological, for example, differences in the age category definitions. An alternative explanation may be the young age of the Alaska EARTH cohort at baseline. Interestingly, all-sites agreement was very similar among categories of educational attainment (greater or less than high school), rural/urban residence and primary language spoken at home (English or non-English/both). Again, this may speak to the importance placed upon delivering culturally competent care to all AI/AN people within the ATHS, regardless of potential communication or distance barriers.

We also examine site-specific agreement measures for colorectal and female breast cancers. For female

breast cancers, agreement was higher among younger women. Breast cancer diagnoses at younger ages are likely to be of greater severity, requiring more extensive treatment courses, and associated with genetic risk factors such as BRCA [33–35], which may heighten accuracy of recall. However, genetic testing was not available during the time period of diagnosis for the cancers examined herein. For colorectal cancers, agreement was lower among those aged 18–50 years at study enrolment and those who had greater than a high-school education. These findings may be related to screening; previous research suggests that individuals with higher education are more likely to receive colorectal cancer screening [36,37], and individuals are substantially more likely to receive screening if they meet age recommendations. In turn, it is possible that individuals with a screening history may be more likely to confuse a non-malignant finding with a cancer diagnosis.

There are strengths and limitations of this study that warrant consideration. A key strength was the use of the registry to identify cancer cases. The previous examination by Koller et al. compared to medical records, which were available only from within the ATHS. In contrast, the ANTR is a central cancer registry, collecting information from healthcare facilities throughout Alaska, as well as linkages with the Alaska and Washington state cancer registries, to ensure completeness. While it is possible that there were some missing cases, the tumour registry records are likely to provide a more complete record of cancer cases among AN people than ATHS medical records alone. Another strength was our examination of multiple measures of agreement. As described in detail by Koller et al. [10], provision of sensitivity and specificity as well as PPV and NPV enables assessment of the self-reported record assuming the tumour registry as the referent (sensitivity/specificity), as well as the registry information assuming self-reported information as the referent (PPV/NPV). Thus, our results enable assessment of agreement independent of the assumption of a gold standard metric. The primary limitation of this study was the small sample size of Alaska EARTH participants with a history of cancer diagnosis; this restricted our ability to conduct certain stratified analyses and resulted in wide confidence intervals around some of our validity measures. Second, the Alaska EARTH study employed a community-based recruitment system producing a convenience sample, which may limit the generalisability of these findings. However, the cohort was representative of its three regions by several demographic variables, including age [38]. Initial recruitment for the Alaska EARTH study, including the self-reported questionnaires discussed herein, occurred in 2004–2006. While

institutional efforts have addressed patient–provider communication and patient health literacy, none of these efforts have focused specifically on cancer. Therefore, we anticipate that these results will remain valid into 2018 and beyond. Furthermore, it is possible that cases may have been missed due to incomplete matching; however, the use of a series of identifiers, including full name, date of birth and social security number, increased the likelihood of accurate matching.

The findings of this study reaffirm previous reports of the validity of self-reported cancer history within the ATHS [10]. However, since agreement was imperfect, it also supports the concept that self-reported public health surveys may provide an incomplete picture of health and disease prevalence. Alternatively, a discrepancy between patient and health record/registry may be a missed cancer diagnosis or a cancer diagnosed in another area or facility beyond the catchment area of the registry; therefore, it may be important for clinicians to continue such medical history discussions with their patients. Finally, this report provides evidence that accurate self-report of a cancer diagnosis varies by cancer site, as well as demographic factors including sex and age. This information provides insight regarding subgroups who may benefit from health literacy interventions to improve understanding of their cancer diagnoses. It also suggests that clinicians within the ATHS may need to tailor their explanations about the cancer screening, prevention, diagnosis, treatment and follow-up to specific population characteristics in order to improve understanding. Further research may be needed to evaluate reasons for discordance by demographic factors, including the exploration of health literacy and patient–provider communication among AN people.

## Conflict of Interest

No potential conflict of interest was reported by the authors.

## ORCID

Gretchen Day  <http://orcid.org/0000-0003-4697-1884>

Kathryn R. Koller  <http://orcid.org/0000-0002-4066-0930>

## References

- [1] Morris NS, Field TS, Wagner JL, et al. The association between health literacy and cancer-related attitudes, behaviors, and knowledge. *J Health Commun.* 2013 Dec;18(Suppl1):223–241. PubMed PMID: 24093358; PubMed Central PMCID: PMC3815140. eng.
- [2] Davis TC, Williams MV, Marin E, et al. Health literacy and cancer communication. *CA Cancer J Clin.* 2002;52(3):134–149.

- [3] Peterson NB, Dwyer KA, Mulvaney SA, et al. The influence of health literacy on colorectal cancer screening knowledge, beliefs and behavior. *J Natl Med Assoc.* 2007;99(10):1105.
- [4] Kim K, Han HR. Potential links between health literacy and cervical cancer screening behaviors: a systematic review. *Psychooncology.* 2016 Feb;25(2):122–130. PubMed PMID: 26086119.
- [5] Oldach BR, Katz ML. Health literacy and cancer screening: a systematic review. *Patient Educ Couns.* 2014 Feb;94(2):149–157. PubMed PMID: 24207115; PubMed Central PMCID: PMCPCMC3946869.
- [6] van der Heide I, Uiters E, Jantine Schuit A, et al. Health literacy and informed decision making regarding colorectal cancer screening: a systematic review. *Eur J Public Health.* 2015 Aug;25(4):575–582. PubMed PMID: 25733553.
- [7] Song L, Mishel M, Bensen JT, et al. How does health literacy affect quality of life among men with newly diagnosed clinically localized prostate cancer? Findings from the North Carolina-Louisiana Prostate Cancer Project (PCaP). *Cancer.* 2012 Aug 1;118(15):3842–3851. PubMed PMID: 22180041; eng.
- [8] Halverson JL, Martinez-Donate AP, Palta M, et al. Health literacy and health-related quality of life among a population-based sample of cancer patients. *J Health Commun.* 2015;20(11):1320–1329. PubMed PMID: 26161549; PubMed Central PMCID: PMCPCMC4751057. eng.
- [9] Bokhof B, Eisele L, Erbel R, et al. Agreement between different survey instruments to assess incident and prevalent tumors and medical records – results of the Heinz Nixdorf Recall Study. *Cancer Epidemiol.* 2014 Apr;38(2):181–192. PubMed PMID: 24534296; eng.
- [10] Koller KR, Wilson AS, Asay ED, et al. Agreement between self-report and medical record prevalence of 16 chronic conditions in the Alaska EARTH study. *J Prim Care Community Health.* 2014 Jul;5(3):160–165. PubMed PMID: 24399443; eng.
- [11] Phillips KA, Milne RL, Buys S, et al. Agreement between self-reported breast cancer treatment and medical records in a population-based Breast Cancer Family Registry. *J Clin Oncol.* 2005 Jul 20;23(21):4679–4686. PubMed PMID: 15851764; eng.
- [12] D’Aloisio AA, Nichols HB, Hodgson ME, et al. Validity of self-reported breast cancer characteristics in a nationwide cohort of women with a family history of breast cancer. *BMC Cancer.* 2017;17(1):692.
- [13] Bergmann MM, Byers T, Freedman DS, et al. Validity of self-reported diagnoses leading to hospitalization: a comparison of self-reports with hospital records in a prospective study of American adults. *Am J Epidemiol.* 1998 May 15;147(10):969–977. PubMed PMID: 9596475; eng.
- [14] Bergmann MM, Calle EE, Mervis CA, et al. Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries. *Am J Epidemiol.* 1998 Mar 15;147(6):556–562. PubMed PMID: 9521182; eng.
- [15] Colditz GA, Martin P, Stampfer MJ, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol.* 1986 May;123(5):894–900. PubMed PMID: 3962971; eng.
- [16] Navarro C, Chirlaque MD, Mj T, et al. Validity of self reported diagnoses of cancer in a major Spanish prospective cohort study. *J Epidemiol Community Health.* 2006 Jul;60(7):593–599. PubMed PMID: 16790831; PubMed Central PMCID: PMCPCMC2566235.
- [17] Desai MM, Bruce ML, Desai RA, et al. Validity of self-reported cancer history: a comparison of health interview data and cancer registry records. *Am J Epidemiol.* 2001;153(3):299–306.
- [18] Schoeller DE, Westerterp M, editors. *Advances in the assessment of dietary intake.* Boca Raton: CRC Press; 2018.
- [19] Mattes W, Yang X, Orr MS, et al. Biomarkers of tobacco smoke exposure. *Adv Clin Chem.* 2014;67:1–45. PubMed PMID: 25735858; eng.
- [20] Hills AP, Mokhtar N, Byrne NM. Assessment of physical activity and energy expenditure: an overview of objective measures. *Front Nutr.* 2014;1:5. PubMed PMID: 25988109; PubMed Central PMCID: PMCPCMC4428382. eng.
- [21] Fowles JB, Fowler EJ, Craft C. Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. *J Ambul Care Manage.* 1998 Jan;21(1):24–34. PubMed PMID: 10181337; eng.
- [22] Machon M, Arriola L, Larranaga N, et al. Validity of self-reported prevalent cases of stroke and acute myocardial infarction in the Spanish cohort of the EPIC study. *J Epidemiol Community Health.* 2013 Jan;67(1):71–75. PubMed PMID: 22577182; eng.
- [23] Simpson CF, Boyd CM, Carlson MC, et al. Agreement between self-report of disease diagnoses and medical record validation in disabled older women: factors that modify agreement. *J Am Geriatr Soc.* 2004 Jan;52(1):123–127. PubMed PMID: 14687326; eng.
- [24] Gentry-Maharaj A, Fourkala EO, Burnell M, et al. Concordance of National Cancer Registration with self-reported breast, bowel and lung cancer in England and Wales: a prospective cohort study within the UK Collaborative Trial of Ovarian Cancer Screening. *Br J Cancer.* 2013 Nov 26;109(11):2875–2879. PubMed PMID: 24129231; PubMed Central PMCID: PMCPCMC3844906. eng.
- [25] Li J, Cone JE, Alt AK, et al. Performance of self-report to establish cancer diagnoses in disaster responders and survivors, World Trade Center Health Registry, New York, 2001–2007. 433021. 2016 May–Jun;131(3):420–429. PubMed PMID: 27252562; PubMed Central PMCID: PMCPCMC4869085.
- [26] Alaska Population by Age, Sex, Race (Alone or in Combination) and Hispanic Origin [Internet]. 2015 Jul [cited 2017 Mar 23]. Available from: <http://live.laborstats.alaska.gov/pop/index.cfm>
- [27] Bureau USC. Census Summary File 1. 2010 [cited 2017 Mar 23]. Available from: <https://factfinder/census.gov>
- [28] Slattery ML, Schumacher MC, Lanier AP, et al. A prospective cohort of American Indian and Alaska Native people: study design, methods, and implementation. *Am J Epidemiol.* 2007 Sep 1;166(5):606–615. PubMed PMID: 17586578; PubMed Central PMCID: PMC2556228. eng.
- [29] International Classification of Diseases for Oncology (ICD-O). First revision. 3rd ed. Geneva: World Health Organization; 2013.
- [30] Molisani A, Dumenci L, Matsuyama RK. Influences of patient sociodemographics on cancer information received through the first 9 months of treatment. *J Cancer Educ.* 2014 Mar;29(1):158–166. PubMed PMID: 24113903; eng.

- [31] Gansler T, Henley SJ, Stein K, et al. Sociodemographic determinants of cancer treatment health literacy. *Cancer*. 2005;104(3):653–660.
- [32] Berthier F, Grosclaude P, Bocquet H, et al. Prevalence of cancer in the elderly: discrepancies between self-reported and registry data. *Br J Cancer*. 1997;75(3):445–447. PubMed PMID: 9020495; PubMed Central PMCID: PMC2063375. eng.
- [33] Parazzini F, Villa A, Polverino G, et al. Epidemiology of breast cancer in young women. In: Biglia N, Peccatori FA, editors. *Breast cancer, fertility preservation and reproduction*. Cham: Springer International Publishing; 2015. p. 1–9.
- [34] Gabriel CA, Domchek SM. Breast cancer in young women [journal article]. *Breast Cancer Res*. 2010 Oct 28;12(5):212.
- [35] Bleyer A, Barr R, Hayes-Lattin B, et al. The distinctive biology of cancer in adolescents and young adults [review article]. *Nat Rev Cancer*. 2008 Apr 1;8:288. online.
- [36] Meissner HI, Breen N, Klabunde CN, et al. Patterns of colorectal cancer screening uptake among men and women in the USA. *Cancer Epidemiol Biomarkers Prev*. 2006;15(2):389–394.
- [37] Shapiro JA, Klabunde CN, Thompson TD, et al. Patterns of colorectal cancer test use, including CT colonography, in the 2010 National Health Interview Survey. *Cancer Epidemiol Biomarkers Prev*. 2012;21(6):895–904.
- [38] Lanier AP, Redwood DG, Kelly JJ. The Alaska Education and Research Towards Health (EARTH) Study: cancer risk factors. *J Cancer Educ*. 2012 Apr;27(1 Suppl):S80–S85. PubMed PMID: 22298198; eng.
- [39] Redwood DG, Lanier AP, Johnston JM, et al. Chronic disease risk factors among Alaska Native and American Indian people, Alaska, 2004–2006. *Prev Chronic Dis*. 2010 Jul;7(4):A85. PubMed PMID: 20550843; PubMed Central PMCID: PMC2901583. eng.